Karlsruhe Institute of Technology

# Cost-Optimal Service Selection Based on Incident Patterns

## A Service Level Engineering Approach

Zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

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

genehmigte Dissertation

von

Dipl.-Inform.Wirt Axel Christoph Kieninger

| | |
|---|---|
| Tag der mündlichen Prüfung: | 11.11.2015 |
| Referent: | Prof. Dr. Gerhard Satzger |
| Korreferent: | Prof. Dr. Thomas Setzer |

# Abstract

In spite of the steadily increasing significance of IT services, there are still serious concerns with regard to the management of information technology. In this work, we particularly focus on the issue of business productivity, which reflects the need to leverage IT to reduce a company's overall costs and to foster close cooperation between business units and IT departments. More specifically, we address a service customer's problem to monetarily assess the adverse business impact of imperfect service quality in terms of 'total business costs' and, based thereon, to select the cost-optimal IT service among different service offers which minimizes the sum of service price and this monetarily quantified business impact. First, we describe the service customer's optimization problem from an economic point of view. We analyze service quality measures commonly used in industry with regard to their suitability to describe the adverse business impact that a service causes. We exclusively consider service quality measures which characterize services with regard to the incidents these entail. We find that a service customer will usually not be able to precisely determine the business impact caused by specific service levels if aggregating service quality indicators are applied which 'merge' information about the incident behavior of services. Since the assessment of services' adverse business impact is a prerequisite for the determination of cost-optimal service offers, we suggest 'service incident patterns' as an advanced form of service quality measures. In contrast to a service quality indicator, a service incident pattern describes a service's characteristic incident behavior through a distribution function instead of a scalar. We recognize that the knowledge about both, service incident patterns and the business impact induced by individual service incidents, is required to determine the total business costs a service induces. Therefore, we develop the method for Cost-Optimal Service Selection (COSS) to support service customers in solving their optimization problem. COSS leverages discrete-event simulation to monetarily assess the business impact of service incidents on business operations and applies auction theory to facilitate and structure the contract negotiation process between the service customer and potential service providers. Finally, we evaluate and demonstrate the applicability of COSS in a case study and suggest directions for further research.

# Acknowledgements

At this point, I would like to express my sincere gratitude to everyone who contributed to the completion of this thesis. First of all, I want to thank my supervisor Prof. Dr. Gerhard Satzger for his ongoing encouragement, advice, and support, giving me the opportunity to pursue this research in an exciting setting which interlinks research and practice. Furthermore, I am grateful to my co-advisor Prof. Dr. Thomas Setzer for his positive, constructive feedback and for fruitful discussions on possible extensions to this work. Additionally, I would like to thank Prof. Dr. Hansjörg Fromm for making me look beyond my own field of research and for insightful discussions on terminology. I am also thankful to Prof. Dr. York Sure-Vetter and Prof. Dr. Andreas Geyer-Schulz for serving on my thesis committee.

It is a pleasure to thank my former colleagues from the Karlsruhe Service Research Institute and, especially, the Service Innovation and Management team who accompanied the development of this thesis. Their valuable comments and the inspiring environment significantly improved my work. In particular, I owe my gratitude to Björn Schmitz who supported my research for several years through profound questions and thought-provoking impulses. Additionally, I am thankful to Detlef Straeten, Martin Koffler and Dr. Carsten Holtmann from IBM with whom I had numerous discussions on the practical applications of my work. Moreover, I would like to thank the diploma and master students whose theses I supervized and contributed to this work, specifically Jens Westernhagen, Dian Baltadzhiev and Florian Berghoff.

Furthermore, I would like to give a special thank you to Björn Schmitz, Dr. Christian Haas, Stefan Momma and Dr. Andreas Neus for proofreading parts of this work and, especially, for providing me with constructive comments and critical questions. Furthermore, I am grateful for the valuable advice of Dr. Christian Haas, Dr. Margeret Hall and Dr. Robert Kern during the preparation of my disputation.

Finally, a very special thank you to my most precious support. I owe my deepest gratitude to my parents, Ernst and Irmgard Kieninger, for the environment in which I grew up and for always encouraging me to pursue my goals. Likewise, I am thankful to my brother, Andreas Kieninger, for giving me confidence at all times. Lastly and foremost, I am grateful to my wife, Alina, for always being there for me, for her patience and her love. Without you this work would not have been possible.

Axel Christoph Kieninger

# Contents

# List of Abbreviations

ASL ........... Actual Service Level
BDIM ......... Business-Driven IT Management
BIA ........... Business Impact Analysis
COSS ......... method for Cost-Optimal Service Selection
ibid. ........... ibidem, meaning 'in the same place'
RfP ........... Request for Proposal
SLA ........... Service Level Agreement
SLE ........... Service Level Engineering
SLO ........... Service Level Objective
WMS .......... Warehouse Management System

# List of Figures

# List of Tables

# Part I.

# Foundations for Cost-Optimal Service Selection

# 1. Introduction

In 2015, the worldwide yearly growth (in constant currency) in the IT service market is forecasted to reach 3.7 percent compared to 2014 (Hale *et al.*, 2015). The revenue is expected to total nearly $942 billion in 2015 and $1.1 trillion in 2019. The IT outsourcing market, as part of the IT service market, is supposed to grow by more than $79 billion until 2019 (ibid.). These forecasts highlight that organizations increasingly depend on external service and solution providers (e.g., Ambrose & Huntley, 2012). Experts anticipate companies to increase their focus on core IT competencies and to outsource "commodities", such as application operation, application maintenance and infrastructure services (Meyer & Eul, 2013, p.82).

Meanwhile, we observe several trends which will influence the IT outsourcing service market in the years to come (see Meyer & Eul, 2013): Experts anticipate customer companies to conclude contracts with further decreasing deal sizes (with regard to volume and number of services covered), shorter contract durations and, consequently, more frequent negotiations. Furthermore, they expect increasing competition among providers, which is induced through customer companies purchasing IT services from multiple providers (multi-sourcing, e.g., Wiener & Saunders, 2014; Goldberg *et al.*, 2015a; Goldberg & Satzger, 2015) as well as through enhanced tendering and benchmarking procedures applied by customers (Meyer & Eul, 2013). Additionally, customer companies increasingly demand usage-bound pricing and dynamic adaptation of IT services to changing business requirements (Thorenz & Zacher, 2013).

Yet, in spite of the steadily increasing significance of IT services, there are still serious concerns with regard to the management challenges of information technology (e.g., Luftman & Derksen, 2012; Luftman, 2014). In this work, we will particularly address the issue of *"business productivity"*, which reflects the need to leverage IT to reduce a company's overall costs[1] and to foster close cooperation between business

[1] In contrast to Luftman & Derksen (2012, p.209) we regard a company's overall costs instead of "overall company expenses".

units and IT departments (Luftman & Derksen, 2012, p.209). Furthermore, this objective implies the claim to use IT as an enabler and driver (Luftman *et al.*, 2012) for the improvement of business productivity and operational efficiency (Luftman & Ben-Zvi, 2010).[2] More specifically, we aim to monetarily assess the business impact of imperfect service quality in terms of 'total business costs' and, based thereon, to select the cost-optimal IT service among different service offers which minimizes the sum of service price and this monetarily quantified business impact.[3] The next section describes the research problem, which is addressed in this thesis, in detail.

## 1.1. Research Problem

IT departments form the interface between internal business departments and external service providers (e.g., Gewald & Helbig, 2006).[4] Consequently, Luftman & Derksen (2012, p.209) state that by purchasing IT services to support specific business operations, it is the IT department's responsibility to minimize "overall company expenses". We regard a company's payments resulting from the procurement of a *single* IT service as the sum of service price and monetarily quantified, adverse business impact (total business costs) which the particular IT service induces on business operations (Kieninger *et al.*, 2011a). Thus, in order to identify the optimal IT service among different provider offers, IT departments have to be able to determine the business impact a specific IT service causes.

IT departments, however, only have "a modicum of visibility into the business that they are supporting" (Sauvé *et al.*, 2009, p.48). Results of our own empirical studies show that 37% of the IT departments seem to have problems with the determination of business requirements (Goldberg *et al.*, 2015b), which significantly depends on the assessment of services' impact on business operations.[5] In the context of this work, we will focus on a specific aspect of this issue: the challenge to assess the influence of imperfect service quality on business results (e.g., Moura *et al.*, 2006).

Furthermore, focusing on the field of Service Level Management, we find that the "definition of [appropriate] IT service quality [objectives]" — which ought to be done in consideration of services' business impact — represents an urgent challenge for IT departments (e.g., Unterharnscheidt & Kieninger, 2010, p.4). Our studies reveal that 51% of the IT departments seem to struggle with translating business requirements into service requests (Goldberg *et al.*, 2015b). These findings are confirmed by literature stating that service quality goals are often stipulated inappropriately

[2] This topic has been among the top five concerns of Chief Information Officers since 2009, and has been their major concern in 2009, 2010 and 2012 (Luftman, 2014).

[3] For reasons of clarity, we do not consider the impact of interests in this work. Consequently, we may regard cash-effective costs to correspond to cash flows and weigh a service's price against the total business costs it causes.

[4] In the context of IT outsourcing those parts of a service customer's IT department which are not outsourced are also often denoted as 'retained organizations' (e.g., Gewald & Helbig, 2006).

[5] More precisely, we find that 46% of retained organizations with an outsourcing degree between 20% and 50% and even 70% with an outsourcing degree above 50% have difficulties to determine business requirements. According to Lacity *et al.* (2009), we measure the outsourcing degree by the percentage of the IT budget outsourced.

resulting in overly high costs compared to the benefit for business (Taylor & Tofts, 2005). Moreover, today's service quality measures and service quality objectives are often defined in an ad-hoc and heuristic manner (ibid.) — which is sometimes even denoted as "pure guesswork" (Sauvé *et al.*, 2005, p.73).

To summarize, there is no systematic engineering approach today which allows the selection of cost-optimal service solutions among different service offers through the definition of service quality measures that reflect services' adverse impact on business operations.

Within this work, we focus on four facets of an IT department's problem to procure cost-optimal service solutions: The description of the decision problem in economic terms, the definition of appropriate service quality measures, the monetary assessment of the business impact induced by specific service quality levels in terms of total business costs and the negotiation of service offers with service providers. We explicitly consider the trade-off between service price and service quality.

## 1.2. Research Outline

In the scope of this thesis, we address four concrete research questions, each referring to one of the previously outlined facets of an IT department's problem. The overall research question is:

**Which approach is suitable for a customer IT department to select the cost-optimal service solution among different provider offers?**

In order to answer this main research question, we first describe the IT department's optimization problem from an economic point of view. Furthermore, we elaborate on challenges IT departments face when trying to solve this optimization problem. This leads us to our first research question:

**Research Question 1: How can the decision problem of an IT department be described in economic terms?**

Aiming to leverage IT to optimize total service costs, we suggest IT departments to follow a Service Level Engineering approach (see Chapter 3) to define business-relevant service quality measures and, based thereon, identify the optimal[6] service solution — by quantitatively considering the business impact of the quality objectives it comprises (Kieninger *et al.*, 2011a). Thus, we recommend to select the particular service offer for which the sum of service price and monetarily quantified, adverse business impact — induced through a certain service quality level — reaches its minimum.

Considering the trade-off between services' adverse business impact and service price, IT departments are enabled to achieve cost-optimal support of business operations.

[6] We aim to identify the service solution, which is cost-optimal from a customer company point of view.

Due to the struggle of IT departments to determine the total business costs induced by imperfect service quality, this is, however, only rarely achieved nowadays. Therefore, we argue that imperfect service quality has to be described by service quality measures, which allow the assessment of services' adverse business impact.

In the scope of this work, we *exclusively* consider service quality measures which describe the incident behavior of services. That is, we focus on service quality measures characterizing services with regard to the incidents these entail (e.g., measures of availability (outage incidents) and response time (response time incidents)).

After the general discussion of the IT department's optimization problem, we analyze service quality indicators[7] commonly used in industry with regard to their suitability to describe the business impact that a service causes. As stated before, the definition of service quality measures which allow the assessment of services' adverse business impact is a prerequisite for the determination of cost-optimal service offers. This leads us to our second research question:

**Research Question 2: Which characteristics of a service and of the business environment which it is to support does an IT department have to consider in order to identify the cost-optimal service solution with regard to service incident behavior?**

To answer this research question, we first elaborate on service quality indicators currently applied in IT outsourcing practice. We state that the quality indicators we encounter 'merge' information about the incident behavior of services. In general, indicators measure facts quantitatively and in an aggregated manner (Kütz, 2011, p.4), as the following simple example illustrates:

*An IT service is unavailable for three times during each reference period. Deterministically, there are outage incidents with durations of 10, 20, 30 and 40 minutes. The provider describes the unavailability of services through the service quality indicator 'total outage duration', which is calculated by adding the durations of all outage incidents to occur within a reference period. Consequently, it offers the service assuring a total outage duration of 100 minutes. Due to this aggregation of information, a service customer is not able to reconstruct the number and duration of single outage incidents any more.*

We find that, if the quality of services is described through indicators and corresponding objectives, only providers know about services' specific incident behavior. An external, independent IT department, which is interested in purchasing a service, however, usually has no information about the characteristic 'service incident patterns' of this service. We state that the information concealed by aggregation through indicators may be crucial for the determination of a service's business impact. Continuing our example, we briefly clarify this observation:

*An IT department aiming to purchase the service is well-informed about the adverse business impact potential outage incidents would have on business operations: It*

[7] These measures are also frequently denoted as 'service level indicators' (e.g., Sturm *et al.*, 2000, p.66).

*is aware that outage incidents with durations of 10, 20, 30 and 40 minutes cause business impacts in the amount of €1,000, €8,000, €27,000 and €64,000. In a service offer a provider assures a total outage duration of 100 minutes. It is common knowledge that this quality objective will always be exactly met.*

*The IT department now attempts to determine the adverse business impact the service offered would cause. It calculates that the total business costs would amount to €10,000 in case of ten outage incidents with a duration of 10 minutes each (best case). Furthermore, it figures out that the total business costs would amount to €136,000 in case of two 40-minute outage incidents and one 20-minute outage incident (worst case). The actual adverse business impact would amount to €100,000, however.*

A lack of knowledge about the characteristic incident patterns of a service, i.e., the combinations of incidents which realize the assured service quality levels, may inhibit the determination of the actual amount of total business costs and, thus, impede the selection of cost-optimal service solutions. We define the concept of 'business cost functions' to describe the dependence of a single incident's business impact, which we measure in terms of 'business costs', on its specific properties (e.g., its duration).[8]

We recognize that 'business cost functions' and 'service incident patterns' are two essential concepts to be considered when addressing the IT department's optimization problem.[9] Since service quality indicators which are commonly applied in practice today may leave IT departments uninformed about a service's impact on business operations, we suggest to use incident patterns to describe, to measure and to negotiate on service quality.

In the context of this thesis, we refer to the cash-effective equivalent of a service's adverse impact on business operations as total business costs.[10] Thus, total business costs represent the financial impact business departments incur due to imperfect service quality (Kieninger *et al.*, 2012a). A prerequisite for the calculation of total business costs is the knowledge about the adverse monetary impact of single service incidents, which is subject of our third research question:

**Research Question 3: How can IT departments quantify the impact of service incidents on business operations in case of well-defined business processes?**

IT departments have to understand the effects of service incidents on business operations in order to be able to weigh service quality against service price. Therefore, they have to work closely with business departments and support these in determining business cost functions. The identification of business consequences, which may result from service incidents and can be assessed in terms of business costs, is a first step towards this goal.

8 In the example above, business costs of outage incidents amount to the third power of their duration in minutes (business cost function).

9 In Chapter 4, we discuss further important concepts related to the ones introduced.

10 As stated above, we do not consider the impact of interests in this work and, consequently, may regard cash-effective costs to correspond to cash flows.

In order to support this step, we provide an overview on classifications of business consequences and corresponding costs, which we identify via an extensive literature review. The classifications discussed may serve business and IT departments as a means to determine business consequences which have to be considered in a specific scenario. Supporting the identification of business consequences in a systematic way, these classifications can, for example, be used as a structure for interviews with experts from business departments.

Furthermore, we develop a simulation-based procedure to support the estimation of business cost functions. Our approach consists of four steps: First, we identify the business consequences which are affected by service incidents in the scenario under consideration. Then, we determine business process-related metrics which are associated with these business consequences. Afterwards, we simulate the influence of single service incidents on those business process-related metrics. Thus, we link the occurrence of a service incident with specific properties (e.g., a duration of ten minutes) at a certain time (e.g., at 01:30 p.m.) to changes in the values of business process-related metrics (e.g., the difference in the number of transactions processed). Finally, based thereon, we estimate the particular business cost functions for the different types of service incidents to be considered at different time periods.

We address the scenario in which one or more business processes depend on a specific service. Using discrete-event simulation, we determine the impact of service incidents on business operations through the adoption and usage of an established, formal technique for business process analysis. Thus, we are able to consider stochastic as well as dynamic aspects of business processes.

Considering research question 2, we discuss the suitability of service quality indicators to suggest the business impact which a service causes. We recall that this is a prerequisite for the determination of cost-optimal service offers. At this point, we want to discuss how IT departments can ensure that providers do not have an incentive to state incident patterns[11] untruthfully in their service offers.[12] In other words, we aim to ensure that providers, which we assume to be risk-neutral and to act rationally, state actual service incident patterns and are not tempted to define service incident behavior in a strategic manner.

**Research Question 4: Which strategy should an IT department follow in a negotiation with risk-neutral external service providers when service incident patterns are used as quality measures?**

We develop two auction-based, cost-optimal mechanisms (e.g., Krishna, 2010, p.67), which enable IT departments to select the cost-optimal service solution among different external providers' offers. Both approaches address the last phase of a typical contract negotiation process in IT outsourcing[13], in which buyers negotiate

[11] Service incident patterns fulfill the prerequisite and indicate the business impact a service causes.

[12] So far, we implicitly assumed service quality levels offered to correspond to actual service quality levels.

[13] The preceding phases are typically denoted as 'request for information' and 'request for proposal'.

contract details with selected providers (e.g., Bräutigam, 2009, p.797). Since we aim to weigh up incident-related service quality against service price, we consider the case of the IT department and potential providers to exclusively discuss service incident patterns and service price at this stage.

We assume the IT department to have specified contractual conditions in a 'request for proposal', which providers necessarily have to accept when submitting an offer. This request for proposals stipulates all functional and non-functional service properties except for service incident patterns and service price. It also defines the monetary amount which a provider has to pay in case of service level breaches (penalties), or receives in case of service level overfulfillment (rewards).

We find that an IT department will not be able to select the cost-optimal service solution among different offers in any case if penalty and reward functions are not identical with business cost functions. Furthermore, we demonstrate that service providers will have an incentive to define service offers strategically if they know that the values of penalty and reward functions do not correspond to those of business cost functions. That is, they might try to pretend that their service solutions are cheaper in terms of total service costs. Moreover, we show that this incentive does not exist if penalty and reward functions are identical with business cost functions. Consequently, defining our mechanisms we assume that IT departments reveal business cost functions to potential providers.[14] We use a multi-attribute[15] or single-attribute[16], first-price procurement auction to incentivize providers to offer services at market-oriented prices and to "compete for the right to sell" (Krishna, 2010, p.1).

Based on the answers to the research questions, we develop our method for 'Cost-Optimal Service Selection' (COSS[17]). COSS addresses the previously introduced challenges: First, it supports IT and business departments in discussing and assessing the business impact of IT services. Our 'discrete-event simulation-based method' (step 1) allows the determination of single service incidents' impact on business process measures' values.[18] It refers to the classifications of business consequences and corresponding costs, which support the decision about which business consequences to consider in a specific business setting. Second, using our cost-optimal procurement auction mechanism, COSS enables IT departments to select the optimal service solution with regard to total service costs among different provider offers (step 2).[19]

14 We show that the statement of untrue or distorted functions may lead to suboptimal customer decisions with regard to total service costs.

15 In our multi-attribute auctions providers have to bid tuples of service incident patterns and service price.

16 In contrast, in our single-attribute auctions providers are allowed to state a service price only.

17 The method supports the determination of a service solution's business impact, which today is an unknown variable in the IT department's optimization problem. 'Coß' was the common name for the unknown variable in the German Middle Ages and the name of the renowned algebra textbook by Adam Riese (see Berlet & Riese, 1860).

18 Step 1 of COSS addresses research question 3.

19 Step 2 of COSS aims at research question 4.

Figure 4.8 illustrates the two steps of COSS, its foundation on the Service Level Engineering approach and the idea to apply service incident patterns as an advanced form of service quality measures.[20]

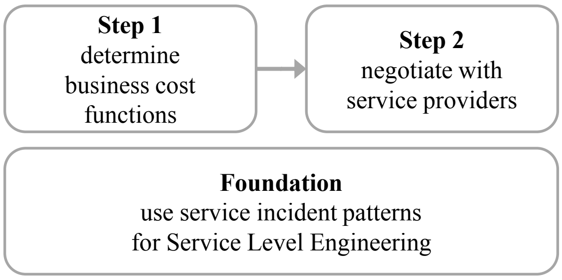


Figure 1.1.: COSS — A Method for Cost-Optimal Service Selection

The following section describes the structure of the work at hand and briefly summarizes the content of its different parts and chapters, as depicted in Figure 1.2.

## 1.3. Structure of this Thesis

Part I (*Foundations for Cost-Optimal Service Selection*) motivates the work at hand and introduces concepts and methods, which it is based on.

- Chapter 1 (*Introduction*) describes the context of this work, specifies the research problem, and details the addressed research questions. Furthermore, it depicts the development of this thesis with regard to its refinement and extension.
- Chapter 2 (*Towards Service Level Engineering*) presents foundations and preliminaries which form the basis of the method for Cost-Optimal Service Selection. It first discusses characteristics of IT services. In particular, Chapter 2 elaborates on IT service quality and quality measures. Additionally, it provides an overview on the fields of simulation and auction theory which are relevant for the implementation of steps 1 and 2 of COSS.
- Chapter 3 (*Service Level Engineering*) explains the Service Level Engineering approach in detail. Moreover, it defines the business setting addressed in subsequent chapters and further specifies the customer's optimization problem to select the cost-optimal IT service. Finally, Chapter 3 refers to fields of research which are related to Service Level Engineering.

Part II (*Development of the Method for Cost-Optimal Service Selection*) elaborates on the method for Cost-Optimal Service Selection proposed in this thesis. It establishes the means and methods which allow an IT department to select the cost-optimal service offer — with regard to incident-related service quality and service price — among different provider offers.

[20] The Service Level Engineering approach tackles the research questions 1 & 2. The suggestion of service incident patterns as service quality measures addresses research question 2.

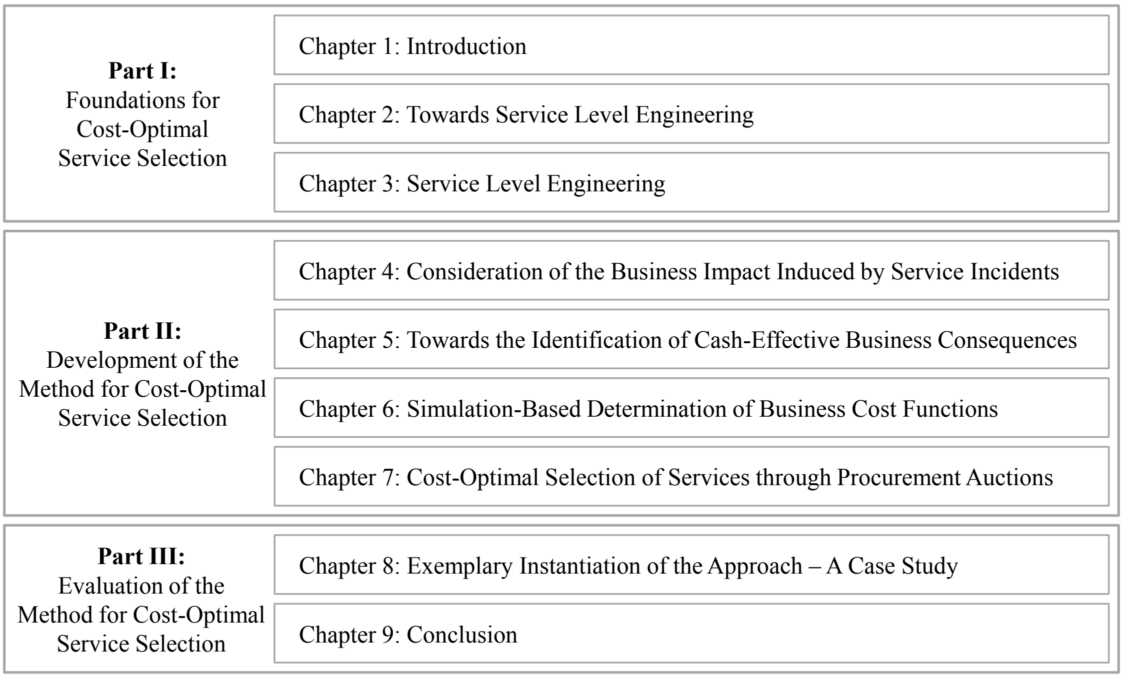


Figure 1.2.: Structure of this Thesis

- Chapter 4 (*Consideration of the Business Impact Induced by Service Incidents*) demonstrates and discusses an IT department's need for new service quality measures in order to enable the selection of cost-optimal service solutions. It introduces a set of incident-related concepts which are essential for the definition of meaningful service quality measures. Furthermore, the chapter briefly describes the steps of the method for Cost-Optimal Service Selection, which builds on the introduced concepts, and refers to related work. Finally, as an excursus, Chapter 4 identifies challenges for service providers emerging from the application of service incident patterns as an advanced form of service quality measures.
- Chapter 5 (*Identification of Cash-Effective Business Consequences*) provides an overview on classifications of business consequences and corresponding costs. These classifications may serve business and IT departments as a means to identify and categorize business consequences when analyzing a specific business setting. Chapter 5, thus, provides the basis for the first step of COSS. Additionally, this chapter discusses why IT departments should focus on cash-effective business consequences when making service investment decisions.
- Chapter 6 (*Simulation-Based Determination of Business Cost Functions*) develops the simulation-based method to guide IT and business departments in jointly determining business cost functions. It supports the assessment of the business costs induced by single service incidents with specific properties which occur at a certain time in case of well-defined business processes. This approach, which builds on business impact analysis, implements the first step of COSS. Furthermore, Chapter 6 discusses limitations and challenges with

regard to the application of the simulation-based method. Finally, it reviews related work which focusses on the assessment of the adverse business impact of imperfect service quality.

- Chapter 7 (*Cost-Optimal Selection of Services through Procurement Auctions*) specifies two auction-based approaches which enable an IT department to select the cost-optimal service solution among different service offers. These cost-optimal mechanisms, which structure the last phase of contract negotiation between the IT department and potential service providers, implement the second step of COSS. Furthermore, Chapter 7 discusses how IT departments can use penalty and reward rules to determine the total service costs which a service offer entails. Finally, this chapter elaborates on limitations and application challenges and briefly refers to related work.

Part III (*Evaluation of the Method for Cost-Optimal Service Selection*) provides an exemplary evaluation of the method for Cost-Optimal Service Selection and comprises the concluding discussion of this work.

- Chapter 8 (*Exemplary Instantiation of the Approach — A Case Study*) demonstrates the applicability of the COSS-method using the example of an order-picking process (customer business process), which is to be supported by a warehouse management system (IT service). In the case considered, the IT department faces the challenge of purchasing the cost-optimal IT service from among different provider offers and, therefore, applies the COSS-method. First, the IT department applies the simulation-based method (step 1; see Chapter 6) in order to determine business cost functions with regard to service outage incidents. Afterwards, the IT department prepares and conducts a multi-attribute procurement auction (step 2; see Chapter 7) with two external service providers regarding the business service required.
- Chapter 9 (*Conclusion*) discusses the contributions and findings of this thesis. Based thereon, this chapter derives managerial implications for representatives of IT departments, business departments and IT providers. Finally, Chapter 9 addresses limitations of this work and suggests directions for further research in the field of Service Level Engineering.

## 1.4. Research Development

Excerpts of the work at hand have been accepted and presented at European and international academic conferences or published as a journal article. This section provides a brief overview on which parts have been discussed in which research community. Furthermore, it depicts the development of this work with regard to its refinement and extension.

The Service Level Engineering paradigm was presented at the 44th Hawaii International Conference on System Sciences (HICSS '44) (Kieninger *et al.*, 2011a). This publication forms the basis for Chapter 3.1. The paper discusses the challenge to realize the cost-optimal service quality level from the perspective of a service system, which is formed by a customer and a provider. It distinguishes different Service Level Engineering scenarios with regard to the way of collaboration between customer and provider.

Whereas the aforementioned contribution addresses an optimization problem from a service system point of view, Kieninger *et al.* (2012a) elaborate on the Service Level Engineering challenge of a customer aiming to select the cost-optimal service solution among the offers by an external provider. This paper was published in the International Journal of Service Science, Management, Engineering, and Technology (IJSSMET). It examines the deficiency of 'aggregating' service level indicators with regard to the description of services' adverse business impact and, thus, serves as groundwork for Chapter 4.1 of this thesis. Furthermore, it suggests a set of incident-related concepts, which should be used to address these shortcomings. A refined version of this set of concepts is discussed in Section 4.3.

Kieninger *et al.* (2012b) extends the approach addressing a setting with one customer and several providers. Additionally, this contribution considers penalty rules (as discussed in Section 7.1), which commit providers to compensate the customer for the business cost caused by every single service incident which turns up during service delivery. Moreover, the paper introduces an auction-based method (applying a single-attribute procurement auction) to structure the last phase of contract negotiation, which corresponds to Section 7.3.2 of this thesis. It was published in the Proceedings of the 18th Americas Conference on Information Systems (AMCIS 2012).

The above-mentioned work is complemented by Kieninger *et al.* (2013a), which discusses the stipulation of reward rules (to be applied in case of providers performing 'better than promised') and penalty rules. The negotiation approach developed uses a multi-attribute procurement auction, which invites providers to state tuples of incident patterns and a price in their service offers. This auction-based method corresponds to the cost-optimal multi-attribute mechanism presented in Section 7.3.1 of the work at hand. The article (Kieninger *et al.*, 2013a) was published in the proceedings of the 11th International Conference on Wirtschaftsinformatik (WI 2013).

Kieninger *et al.* (2013b) presents the simulation-based method for the determination of business cost functions. In particular, it focusses on the quantification of single service incident's adverse impact on business operations. The content of this contribution corresponds to the one of Chapter 6 of the work at hand. The paper was presented at the Fourth International Conference on Exploring Service Science (IESS 1.3).

Taking a provider point of view, Schmitz *et al.* (2014) elaborates on the challenge to design service offers in a business setting in which future incident behavior of service solutions is uncertain and in which a negative or positive deviation from assured service incident patterns is penalized or rewarded. In particular, the paper addresses the question of how a service provider can mitigate risks due to uncertain service incident behavior by incorporating risk premiums in service prices. The paper was published in the proceedings of the Fifth International Conference on Exploring Service Science (IESS 1.4). More generally, Section 4.6 of this thesis discusses how a provider can proceed to determine characteristic service incident patterns of a service solution, and which challenges a provider faces due to the application of service incident patterns as service quality measures.

Additionally, the following works in the fields of Service Level Management and IT Outsourcing, which are related to this thesis, were published between 2010 and 2014:

The results of a qualitative, empirical study on current challenges of Service Level Management from a service customer point of view were presented at the 16th Americas Conference on Information Systems (AMCIS 2010) (Unterharnscheidt & Kieninger, 2010). The study identifies the definition of IT service quality as the major concern.

A model analyzing under which conditions a customer and a provider would conclude a risk-reward sharing service contract was presented at the 44th Hawaii International Conference on System Sciences (HICSS '44) (Satzger & Kieninger, 2011). The work discusses the case of a provider offering price discounts in return for the customer sharing its uncertain revenue. Different risk-attitudes of customer and provider are considered as well.

A study on how IT services are perceived from an end-user point of view was published in the proceedings of the 2011 Service Research Innovation Initiative Global Conference (SRII 2011) (Kieninger *et al.*, 2011b). The paper compares groups of end-user services, which were identified through a series of interviews, with groups of user-related IT services as defined in literature.

Furthermore, a quantitative, empirical study on major shortcomings regarding the design of customer IT departments and key problems with respect to the operational management of an outsourcing relationship is currently under review for publication (Goldberg *et al.*, 2015b). The contribution analyzes correlations between design shortcomings and delivery problems in order to identify potential reasons for problems in outsourcing relationships. Thus, it contributes to improving the design of IT departments.

Complementary to the previously mentioned work and based on the same empirical study, Goldberg *et al.* (2014) elaborates on major challenges which arise in the transition and the delivery phase of IT outsourcing.[21] The contribution was presented at the Fifth International Conference on Exploring Services Science (IESS 1.4).

[21] The transition phase executes the decisions and plans which were made during the design phase.

# 2. Towards Service Level Engineering

In order to achieve the objective of this thesis, developing the method for Cost-Optimal Service Selection, it is important to understand the relevant foundations and preliminaries. Therefore, this chapter presents concepts and methods which form the basis of COSS.

In the following section, we briefly discuss characteristics of IT services. In particular, we elaborate on IT service quality and quality measures. Afterwards, we give an overview on 'simulation', which we apply to determine the adverse business impact of service incidents (see Chapter 6). Additionally, we present foundations of 'auction theory', which we leverage to support the contract negotiation process between the IT department and potential external service providers (see Chapter 7).

## 2.1. IT Services

In the literature there is a multitude of definitions of the term 'service'. ITIL, the "recognized best practice standard" for IT Service Management (Addy, 2007, p.VIII), defines a service as a "means of delivering value to Customers by facilitating Outcomes Customers want to achieve without the ownership of specific Costs and Risks" (Taylor *et al.*, 2007a, p.309).[1] Thus, ITIL implicitly assumes a provider to bear the risks of service delivery.

Building on the above-mentioned definition, ITIL characterizes an 'IT service' as a service, which is "based on the use of Information Technology and supports the Customer's Business Processes" (Taylor *et al.*, 2007a, p.301). Likewise, Berger (2007,

[1] Definitions of the term 'service' (in general) are discussed, for instance, by Hill (1977) and Gadrey (2000).

p.17) states that an IT service is a service which requires the involvement of an IT system.[2]

### 2.1.1. Foundations

ITIL states that an IT service is "made up from a combination of people, processes and technology, and should be defined in a Service Level Agreement" (Taylor *et al.*, 2007a, p.301). It considers the allocation of resources (people and technology), the activities of providers and customers (processes) as well as their results.[3] Thus, ITIL covers the three perspectives of value creation which are used in most service definitions, namely potential-, process- and outcome-orientation (Engelhardt *et al.*, 1993, pp.398-404). Similar to ITIL, Addy (2007, p.80) defines an IT service as "a defined capability or set of deliverables aimed at satisfying a defined requirement, using resources (people, things and tools) and following a defined delivery process."

Zarnekow *et al.* (2006, p.17) emphasize that IT services create value through the support of "customer business processes and business products". They argue that IT services have to be defined from a customer point of view and that IT services may be realized as combinations of 'basic' IT services. ITIL denotes IT services which are visible to service customers as "business services" (Taylor *et al.*, 2007a, p.62). In contrast, ITIL refers to IT services which are essential components of business services (service components that are invisible from a business point of view) as "supporting services".[4] In IT outsourcing,[5] IT services and IT service components are usually grouped into different categories (e.g., Huston, 2013, pp.32-35; Küchler, 2004, pp.85-128). Examples of such categories are desktop services, network services, server services, storage services, system software services and application provisioning services.

In the following, we denote a specific combination of organizational and technical service components that form a certain IT service as a 'service delivery environment.' Providers may be able to implement a particular IT service through different service delivery environments, which vary in at least one service component (e.g., a technical resource, an employee with specific skills or the service delivery process). Moreover, we refer to a service delivery environment which a provider defines in order to offer a certain IT service (i.e., to fulfill specific business requirements) as a 'service solution'.

[2] Further definitions of the term 'IT service', as it is specified in different IT Service Management frameworks, are discussed, for instance, in Mora *et al.* (2014, pp.85-88).

[3] SLAs are used to stipulate service quality objectives, i.e., the outcomes to be achieved.

[4] Similarly, Setzer *et al.* (2010, p.61) differentiate between "composite services" and "atomic services". Blau *et al.* (2008, p.3) distinguish three service layers, namely "utility services", "composite services" and "complex services."

[5] IT outsourcing represents the transfer of responsibilities from a company's IT department to an external service provider (Küchler, 2004, p.54). This third party supplier "provides and manages [...] services for an agreed fee over an agreed time period" (Kern & Willcocks, 2002, p.3). Reviews of IT outsourcing literature are provided, for instance, by Dibbern *et al.* (2004), Gonzales *et al.* (2006) and Lacity *et al.* (2009).

In general, characteristics of services may be classified into the groups of functional and non-functional properties. Building on Sommerville (2007, p.119), functional properties can be regarded as the characteristics of a service which define how it reacts to specific inputs and how it behaves in specific situations. In other words, functional properties describe what a service does.[6]

On the other hand, following O'Sullivan (2006, p.18),[7] we consider non-functional properties as "constraints over [a service's] functionality." O'Sullivan (2006) examines non-functional properties of services (in general) and proposes nine categories — namely availability, price, payment terms, discounts and penalties, obligations, rights, quality, security, and trust. In this work, we explicitly concentrate on IT service quality and price and assume all other functional and non-functional properties to be defined according to business needs.[8]

### 2.1.2. IT Service Quality

In general, quality is regarded as the "degree to which a set of inherent characteristics [...] fulfils requirements" (International Organization for Standardization, 2005, 3.1.1). In IT outsourcing, the quality of IT services is usually specified in 'service level agreements' (SLAs). According to Berger (2007, p.21), an SLA is an agreement which is formally negotiated and set out in writing. It is concluded between two partners that are independent of one another and documents the responsibility to provide IT services (provider) and (monetary) equivalents (customer). An SLA is valid for a fixed period of time and defines rights and duties of the partners. Furthermore, it specifies service processes and stipulates the quality of IT services by the use of well-defined service quality measures, measurement methods and appropriate service quality objectives. Finally, an SLA states the price of the covered IT services and defines consequences (such as penalties) to be borne in case of service level breaches.[9]

Service quality should ideally be measured end-to-end (e.g., Sturm *et al.*, 2000, p.153), i.e., from an end-user point of view. This is to ensure that service quality measures are meaningful to business (Taylor *et al.*, 2007a, p.42). If "technical metrics" (Marques et al., 2009, p.3) are used, which refer to individual service components only, the achievement of quality targets ('service quality objectives') may not necessarily coincide with the satisfaction of business needs.

Comprehensive discussions of service quality measures which are typically used for IT Service Management in practice are provided, for instance, in Berger (2007, pp.113-160) and Brooks (2006). In this work, we concentrate on quality measures which describe the incident behavior of services. Building on the definition of an

6 Glinz (2007), for instance, compare different definitions of the term "functional requirement" from the field of requirements engineering.

7 O'Sullivan (2006) bases his definition of non-functional properties on the definition of non-functional requirements in Chung (1991, p.6).

8 In contrast to O'Sullivan (2006), we regard service quality to also comprise service availability.

9 Berger (2007) identifies this set of characteristic properties of service level agreements in publications by various authors (e.g., Burr, 2003; Karten, 1997; Lewis, 1999; Niessen & Oldenburg, 1997; Pantry & Griffiths, 1997).

'incident' in ITIL[10] (Taylor *et al.*, 2007a, p.299; Hunnebeck *et al.*, 2013, p.454), we define a 'service incident' as follows:

> **Service incident (based on ITIL)**: A service incident is an "unplanned interruption to an IT Service or reduction in the Quality of an IT Service" which is observable from a business point of view.

We note that this definition supports an end-to-end view on services. In other words, it exclusively considers the interruptions and reductions in the quality of a service which have an influence on business operations.[11]

Usually, 'service level indicators' are defined in order to measure and stipulate service quality (e.g., Sturm *et al.*, 2000, pp.66-67). Indicators (in general) measure facts quantitatively and in an aggregated manner (Kütz, 2011, p.4). Accordingly, service level indicators which describe service incident behavior aggregate the attribute values of service incidents on which they are based into a single value, namely the actual service level (ASL) using a specific calculation formula. Besides, a service level indicator usually refers to a certain reference period. Hence, its ASL is calculated considering only the service incidents which occur during the corresponding reference period and belong to this particular service level indicator.

*For example, the attribute 'outage duration' could describe a service incident of the type 'outage incident'. The ASL of the corresponding service level indicator 'service availability' (e.g., Berger, 2007, p.138) could then be calculated summing up the values of the attribute 'outage duration' of all outage incidents (which occurred during the reference period) and, afterwards, subtracting the resulting 'total outage time' (aggregated value) from the 'agreed service time'.[12]. Finally, this difference is divided by the agreed service time and multiplied by 100 in order to obtain the service availability (ASL) in percent (see Equation 2.1).*

$$\textit{service availability} := \frac{\textit{agreed service time - total outage time}}{\textit{agreed service time}} \cdot \textit{100} \qquad (2.1)$$

*Let us assume a month to have 30 days and let us suppose an agreed service time of 43,200 minutes (twenty-four-seven service). Furthermore, let us assume that in this particular month there were three outages with durations of 216, 108 and 54 minutes. Then, the total outage time amounts to 378 minutes which, in turn, results in a service availability of 99.125%.*

We refer to the target value of a service level indicator as 'service level objective' (SLO).[13] At the end of a reference period a service level objective is usually compared with the corresponding actual service level in order to determine if the stipulated

[10] ITIL defines an incident as an "unplanned interruption to an IT Service or reduction in the Quality of an IT Service" (Taylor *et al.*, 2007a, p.299; Hunnebeck *et al.*, 2013, p.454).

[11] The definition of an incident in ITIL also includes incidents, which cannot be perceived by end-users (e.g., "failure of one disk from a mirror set" (Taylor *et al.*, 2007a, p.299)).

[12] The agreed service time is the period of time within the reference period during which the service should ideally be available.

[13] Sturm *et al.* (2000, pp.63-66) argue that SLOs, which are stipulated in SLAs, should meet

service quality was achieved. In some cases, a service level objective is complemented by another target value which aims to define the maximum or minimum attribute value of a single service incident which is still acceptable ('threshold value'). We denote such a target value as 'service incident objective'.

*A service incident objective might define, for instance, that there may not be any outage incidents with a duration of more than 120 minutes.*

In this work, we exclusively focus on service quality measures which describe the incident behavior of services. Accordingly, we differentiate between incident-based and not incident-based service quality measures (see Figure 2.1).

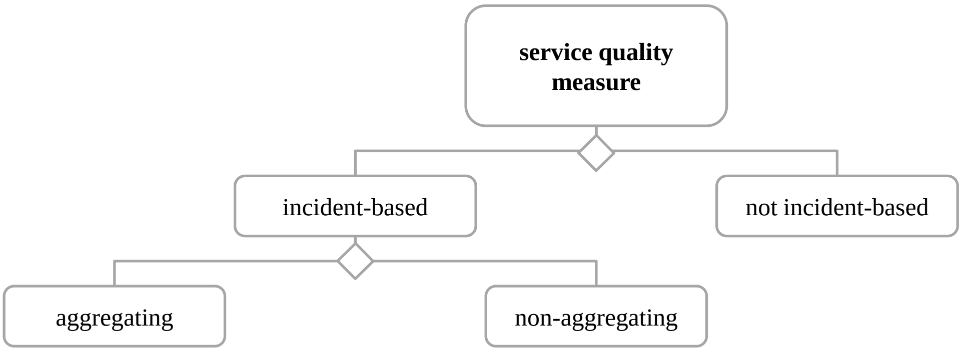


Figure 2.1.: Classification of Service Quality Measures

Examples of service quality measures which are not based on attribute values of service incidents are 'agreed service time', 'percentage of service problems resolved within stipulated time' (problem-based service quality measure; e.g., Taylor *et al.*, 2007b, p.67) and 'number of service changes implemented within stipulated time' (change-based service quality measure; e.g., Taylor *et al.*, 2007c, p.64). Furthermore, we distinguish aggregating, incident-based service quality measures from non-aggregating, incident-based service quality measures. An aggregating service quality measure combines the values of a particular service incident attribute into a single value. Thus, it 'merges' information about the incident behavior of the corresponding service. A non-aggregating service quality measure, in turn, does not merge and conceal information about the individual values of a service incident attribute. In Chapter 4, we discuss this difference in more detail.

## 2.2. Techniques for Cost-Optimal Service Selection

This section provides an overview on the methodologies and techniques we apply in subsequent chapters of this work. Presenting essential definitions and referring to further reading it serves as a brief reminder for experts and as a basic introduction for other interested readers.

seven criteria: They should be attainable and controllable for the service provider, meaningful and understandable to customer and provider, measurable, affordable for the customer and acceptable for all parties involved.

### 2.2.1. Simulation

Banks *et al.* (2010, p.21) define a simulation as the "imitation of the operation of a real-world process or system over time." A simulation model, which mimics the considered real-world process or system, is developed in order to generate an "artificial history" of process or system behavior. This data is analyzed in order to "draw inferences concerning the operating characteristic of the real system."

Generally, simulation models can be classified as being deterministic or stochastic, static or dynamic and discrete or continuous (e.g., Kelton *et al.*, 2004, pp.9-10; Banks *et al.*, 2010, pp.33-34; Leemis & Park, 2006, p.3):

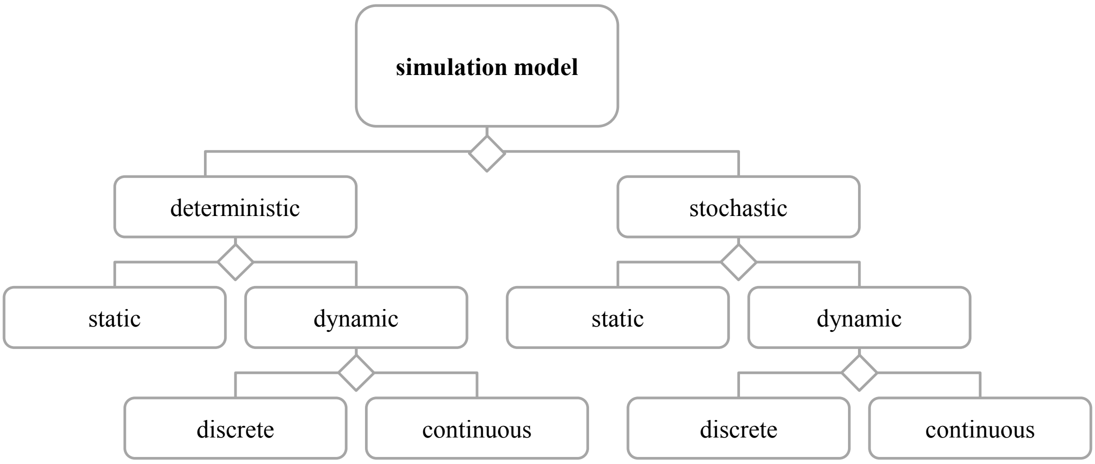


Figure 2.2.: Classification of Simulation Models (Leemis & Park, 2006, p.3)

First, deterministic simulation models are distinguished from stochastic simulation models. Deterministic simulation models do not contain random variables. They have a predefined input which induces a unique output (Banks *et al.*, 2010, p.33). In contrast, stochastic simulation models contain stochastic components (Leemis & Park, 2006, p.2). A stochastic business process model, for instance, might comprise random variables (which follow specific probability distributions) that represent the arrival times of customers or processing times of activities.

Second, a distinction is made between static and dynamic simulation models. Static simulation models "represent a system [or process] at a particular point in time" (Banks *et al.*, 2010, p.33). In contrast, dynamic simulation models regard the development of systems and processes over time. A dynamic simulation model is required if the "passage of time [...] plays a significant role" (Leemis & Park, 2006, p.2). If it is of importance, for instance, at what time a customer enters a business process, a dynamic model should be used.

Finally, two subtypes of dynamic simulation models are differentiated. A dynamic simulation model may be discrete or continuous.[14] In a discrete dynamic simulation

[14] We note that there may also be "mixed continuous-discrete [simulation] models" which comprise both, model elements that change their state at certain points of time or continuously (Kelton *et al.*, 2004, p.9).

model, the state[15] of the model can change only at those discrete points in time at which an event occurs (Kelton *et al.*, 2004, p.9). In contrast, in a continuous dynamic simulation model the state of the model evolves continuously over time (ibid.). The state of a business process may, for instance, depend on the value of the variable 'number of customers in the process'. Consequently, the state of the business process changes whenever a customer enters or leaves the process.

According to the above classification, a discrete-event simulation model is defined as follows (Leemis & Park, 2006, p.3):

> **Discrete-event simulation model**: A discrete-event simulation model is a stochastic and dynamic simulation model which changes its state at discrete points in time only.

Simulation models which are stochastic and static are denoted as Monte Carlo simulation models (Leemis & Park, 2006, p.3). As opposed to discrete-event simulation models, this type of simulation models does not consider the evolution of a model's state variables over time.

Comprehensive overviews on the field of simulation are provided, for instance, by Banks *et al.* (2010) and Law (2011). Banks *et al.* (2010, pp.22-24) also present a list of ten rules which state when simulation is not an appropriate approach. Furthermore, they elaborate on advantages and disadvantages of simulation.

### 2.2.2. Auction Theory

McAfee & McMillan (1987, p.701) describe an auction as a "market institution with an explicit set of rules determining resource allocation and prices on the basis of bids from the market participants." The set of rules defines "who wins what and pays how much" (Krishna, 2010, p.6). Auctions are applied, for instance, in order to determine the price of an object or a service which has "no standard value" (Cassady, 1967, p.20). Comprehensive overviews on the field of auction theory are provided, for instance, by Klemperer (2004) and Krishna (2010).

Auctions are a means to elicit private information from bidders, such as their willingness to pay (Krishna, 2010, p.6). Furthermore, they can be applied to sell or procure various types of objects or services — that is, they are "universal" (ibid.). Moreover, auctions are "anonymous" in the sense that it is not of interest who a particular bidder is, when the winner and the price are determined (ibid.).

Diverse aspects may be considered in order to distinguish different types of auctions. Following Wurman *et al.* (2001, pp.306-307), we first differentiate between "single-sided" and "double-sided" auctions depending on the number of buyers and sellers that take part (see Figure 2.3). In double-sided auctions there are multiple buyers as well as multiple sellers (van Dinther, 2007, p.22).[16] In contrast, in single-sided

[15] The state of a model may be defined as the "collection of [the values of those] variables necessary to describe the system [or process regarded] at any time [(state variables)], relative to the objectives of the study" (Banks *et al.*, 2010, p.30).

[16] Stock markets are an example of double-sided auctions (Berninghaus *et al.*, 2010, p.232).

auctions there either are a single buyer and multiple sellers (procurement auctions) or multiple buyers and a single seller (sales auctions; Berninghaus *et al.*, 2010, p.232). In this work, we concentrate on procurement auctions which may be used to incentivize providers to offer services at market-oriented prices and to "compete for the right to sell" (Krishna, 2010, p.1).

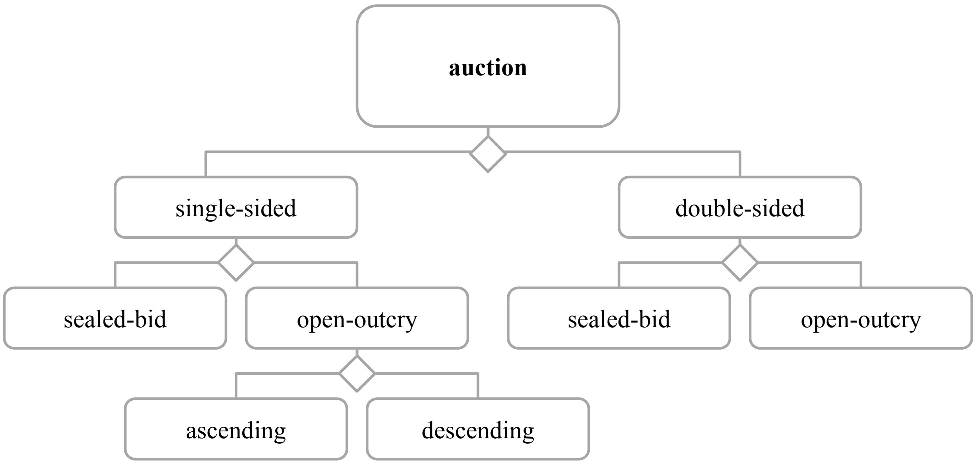


Figure 2.3.: Classification of Auctions (based on Wurman *et al.*, 2001, p.307)

We further distinguish between "sealed-bid" and "open-outcry" procurement auctions (e.g., Wurman *et al.*, 2001, pp.306-307). In open-outcry auctions a participant's bids are "publicly announced" (van Dinther, 2007, p.23) and, thus, known to all bidders. In contrast, the content of a participant's bid is not disclosed to other bidders in sealed-bid auctions. In a first-price, sealed-bid procurement auction the bidder stating the lowest price wins the contract and has to provide the service in question at the quoted price. The bidder that quotes the lowest price wins the contract in a second-price, sealed-bid procurement auction as well. In this case, however, it has to offer the service at the price of the second-lowest bid.[17]

Additionally, we differentiate between "descending" and "ascending" single-sided, open-outcry procurement auctions (based on Wurman *et al.*, 2001, pp.306-307). Initially, an upper or a lower price limit is defined, respectively. Afterwards, in a descending procurement auction the price is successively lowered "until only one bidder remains, and that bidder wins" (Klemperer, 2004, p.11) the auction and has to provide the service at the final price.[18] In an ascending open-outcry procurement auction the price is successively raised until a bidder states that he is willing to offer the service at the current price.[19]

[17] First-price and second-price, sealed-bid sales auctions are described, for instance, in Klemperer (2004, pp.11-12), Berninghaus *et al.* (2010, p.234) and Krishna (2010, p.2). They are regarded as standard forms of auctions (ibid.).

[18] A single-sided, ascending open-outcry sales auction is usually denoted as an "English auction" (e.g., Klemperer, 2004, pp.11-12; Berninghaus *et al.*, 2010, p.234; Krishna, 2010, p.2).

[19] A single-sided, descending open-outcry sales auction is usually denoted as a 'Dutch auction' (e.g., Klemperer, 2004, pp.11-12; Berninghaus *et al.*, 2010, p.234; Krishna, 2010, p.2).

Finally, we differentiate between single-attribute and multi-attribute procurement auctions (e.g., Bichler, 2000, pp.253-254). In a single-attribute procurement auction a seller's bid consists of a single attribute value only (e.g., a price quote). In multi-attribute procurement auctions, on the other hand, a bid is a set of values — one value for each attribute to be considered (e.g., specifications of the price and the offered quality).[20] In order to assess a bid a "scoring rule" (i.e., a multi-dimensional function) is applied (Klemperer, 2004, p.34).

Different auction forms are usually compared in order to identify the particular type which either minimizes a buyer's procurement costs[21] or maximizes social welfare (efficiency) from the perspective of all participants as a whole (Krishna, 2010, p.5). In this thesis, we propose two auction mechanisms which enable IT departments to minimize their service procurement costs (see Chapter 7).

In this chapter, we elaborated on characteristics of IT services and gave brief overviews on the fields of simulation and auction theory. Thus, we laid the foundation for subsequent chapters of this work. Next, we introduce Service Level Engineering as a field of research, which aims to determine business-relevant service quality measures and, based thereon, select — with regard to service quality — cost-optimal service solutions.

[20] In literature multi-attribute procurement auctions are also denoted as "multidimensional auctions" (e.g., Che, 1993; Klemperer, 2004, p.34).

[21] In case of sales auctions the aim is to find the particular auction type which maximizes a seller's sales revenue.

# 3. Service Level Engineering

The quality of a service should always be defined to match the requirements of the business operations it supports. In recent years, however, the business perspective was frequently not considered sufficiently when defining service level indicators (e.g., Mason, 2002; Marques *et al.*, 2007). That is, service quality was usually not specified in an end-to-end manner.[1] Instead, often technical service level indicators were applied, which refer to single components of a service only and, thus, do not reflect the impact of imperfect service quality on business operations as a whole (e.g., Marques *et al.*, 2007).[2]

As a consequence, when technical service level indicators are applied, service customers struggle to understand the business impact which the selection of specific service quality targets entails. Accordingly, service level objectives are often stipulated inappropriately, resulting, for instance, in overly high total service costs (Taylor & Tofts, 2005). Some authors state that SLOs are commonly defined in an ad-hoc and heuristic manner (ibid.), which is sometimes even denoted as "pure guesswork" (Sauvé *et al.*, 2005, p.73).

Therefore, we suggest service customers and providers to follow a systematic Service Level Engineering approach when selecting service solutions. The overall goal of Service Level Engineering is to choose — with regard to service quality — the cost-optimal service solution among different service offers. That is, to solve a quantitative optimization problem to select the particular service offer for which the sum of 'service delivery costs' and monetarily quantified 'adverse business impact' — resulting from imperfect service quality — reaches its minimum.[3]

---

1 The term 'end-to-end' is briefly explained in Section 2.1.2.

2 We regard a service as imperfect, if at least one of its actual service levels is lower than those of a perfect service which achieves the maximum actual service level with regard to each service quality measure.

3 Since we do not consider the impact of interests in this work, we regard cash-effective costs to correspond to cash flows.

The following sections of this chapter provide an overview on the field of Service Level Engineering.[4] First, we discuss its foundations. Second, we introduce the business setting which is addressed throughout the following chapters and describe the optimization problem which an IT department faces in a typical IT outsourcing scenario. Finally, we discuss fields of research which are related to Service Level Engineering.

## 3.1. Foundations of Service Level Engineering

Service Level Engineering pursues two objectives in order to achieve its overall goal to select the cost-optimal service solution among different service offers:

- For a required service, it aims to determine the service quality measures which are relevant from a business point of view and which should hence be used to define the quality of this service in service level agreements.
- It intends to use the service quality measures determined, in order to identify the cost-optimal service solution by quantitatively considering the adverse business impact which its service quality objectives entail.

According to these objectives, we define Service Level Engineering as follows:

> **Service Level Engineering (SLE):** Service Level Engineering is a systematic engineering approach to determine business-relevant service quality measures and, based thereon, select the cost-optimal service solution — by quantitatively considering the adverse business impact of imperfect service quality — among different service offers.

We base our SLE approach on financial metrics, since investment opportunities are usually assessed using capital budgeting techniques (cf., Baker & English, 2011). The definition of service quality measures which are meaningful to business is a challenging task. Service quality is ideally measured end-to-end (e.g., Sturm *et al.*, 2000, p.153), i.e., from a business point of view, in order to ensure it fulfills business needs (Taylor *et al.*, 2007a, p.42). If "technical metrics" (Marques *et al.*, 2009, p.3) — which describe service quality with regard to individual service components — are used, their values have to be aggregated in order to assess the adverse business impact of a service. Our empirical studies show that, in this case, providers usually strive to achieve quality targets only and do not consider business requirements sufficiently (Unterharnscheidt & Kieninger, 2010).

The understanding of business processes plays an important role with regard to the definition of service quality objectives. Without that knowledge, the monetary quantification of specific service quality levels' impact on business operations is not possible. This assessment, in turn, is indispensable to weigh service delivery costs

[4] A previous version of Section 3.1 was published in the proceedings of the Forty-Fourth Hawaii International Conference on System Sciences (see Kieninger *et al.*, 2011a). Furthermore, Section 3.3.1 is based on Cardoso *et al.* (2016), which will be published in the textbook 'Fundamentals of Service Systems'.

against the costs of poor service quality — and, thus, to identify the cost-optimal service solution.

In Service Level Engineering we generally assume all properties of a service required to be already determined by business needs and, thus, to be fix — except for service quality and service delivery costs (but including 'base sizing', i.e., a specification of number of users, traffic, intervals for demand, etc.). Consequently, from an SLE point of view, a service offer is represented by a tuple of service quality objectives and service delivery costs — since these are the only variables specifying offers which are subject to optimization.

Based on the fundamental concepts introduced, we formally describe the general SLE optimization problem (see Equation 3.1):

$$find : s^* := \underset{s \in S}{\operatorname{argmin}} \Big( c_s + i_s \Big) \tag{3.1}$$

For a service required, Service Level Engineering aims to select the cost-optimal service offer $s^*$ (from a set of service offers $S$), which minimizes total service costs. The total service costs of a distinct service offer $s$ are the sum of service delivery costs $c_s$ and monetarily quantified, adverse business impact $i_s$.

### 3.1.1. Characteristics of Service Level Engineering Approaches

Service Level Engineering can be applied in different settings.[5] First, we distinguish between 'internal' and 'external' SLE approaches — since services can be offered by internal or external providers ('organizational situation'). Internal SLE approaches exclusively consider in-house service providers, which belong to the same company as service requesters (e.g., business departments) do. In contrast, external SLE approaches are characterized by their exclusive applicability on settings with external service providers only. We note that 'hybrid' forms of internal and external SLE approaches are possible as well.

We further differentiate between 'cooperative' and 'unilateral' external SLE approaches, thus, considering the 'competitive situation' regarded. Approaches aiming at cooperative settings assume external providers and customers to work closely together in order to identify the cost-optimal service offer. On the other hand, unilateral approaches assume that a single organization strives to identify the cost-optimal service solution on its own. In general, due to common corporate goals of providers and customers, internal SLE approaches should be cooperative.

Additionally, we categorize SLE approaches with regard to the 'number of service providers and customers' considered. An approach could, for instance, be developed to address a setting with one customer and one provider. We denote this form as 'one-to-one' setting. Accordingly, another approach could target a setting with one

[5] A setting is defined by "the time, place, and circumstances in which something occurs or develops" (Merriam-Webster Online Dictionary, 2014).

customer negotiating with multiple potential service providers at once — which we would call a 'one-to-many' setting.

Moreover, SLE approaches can be classified according to the 'perspective' these take — i.e., the target group which should apply them. An approach could, for example, be designed for implementation by a 'customer' only. Others might be of interest for a 'provider' or a (hypothetical) 'omniscient analyst'.

Finally, we consider 'methods applied' to identify a cost-optimal service solution in order to further distinguish SLE approaches. A business department using a service could, for example, apply analytical approaches on historical financial data in order to monetarily assess the business impact of past service outages ('data analysis'). This information could then be used to quantify the expected influence of different service quality objectives on business operations and, thus, to identify the cost-optimal service solution. Another method to address the task described is to conduct 'expert interviews' (ask experts for their educated guess). Table 3.1 depicts the characteristics of SLE approaches discussed.[6]

Table 3.1.: Characteristics of Service Level Engineering Approaches

| characteristic | exemplary values |
|---|---|
| - organizational situation addressed | internal, external, hybrid |
| - competitive situation regarded | cooperative, unilateral |
| - number of customers & providers considered | one-to-one, one-to-many |
| - perspective taken | customer, provider, omniscient analyst |
| - methods applied | data analysis, expert interviews |

We also use these characteristics to describe business settings in which Service Level Engineering is applied[7] (see, for instance, Sections 3.1.3 and 3.2).

### 3.1.2. Service Delivery Costs and Total Business Costs

We recall that the overall goal of Service Level Engineering is to find the particular service solution among different service offers which minimizes total service costs.[8] Furthermore, we regard total service costs as the sum of 'service delivery costs' and 'monetarily quantified, adverse business impact', which is induced by imperfect service quality. In the following we denote the latter component of total service costs as 'total business costs'.[9]

[6] Since it is not possible to provide a comprehensive list of values a characteristic may take in any case (e.g., regarding the characteristic 'methods applied'), we only state exemplary values.

[7] The characteristic 'methods applied', however, is required for the classification of SLE approaches only.

[8] At the end of Section 3.1.3 we briefly discuss a business setting in which a provider should apply SLE in order to increase its chance to conclude a contract.

[9] Business costs represent a particular component of "quality costs", which are usually regarded as the costs that are incurred to avoid quality problems, the costs induced by evaluating the achieved quality level, and the costs resulting from imperfect quality (Yang, 2008, p.177; see also Section 5.1).

> **Total business costs:** Total business costs are the cash-effective equivalent of a service's adverse impact on business operations, which results from imperfect service quality.

For reasons of clarity, we do not consider the impact of interests in this work. Consequently, we may regard cash-effective costs to correspond to cash flows. Total business costs represent the financial impact an organization incurs due to the application of a service — with specific service quality objectives — on its business side (i.e., in departments utilizing the service).[10] Using the term total *business* costs we aim to explicitly distinguish these from the costs which accrue through service provision on the *delivery* side. Therefore, we define service delivery costs as the cash-effective costs arising from the provision of a service — again with distinct service quality objectives.[11] In case of an IT service, service delivery costs comprise, inter alia, cash-effective costs of IT personnel, infrastructure and software.

> **Service delivery costs:** Service delivery costs are the cash-effective costs for delivering a service with specific service quality objectives.

In general, internal and external providers may have to increase efforts and invest more (e.g., on better educated employees and more reliable infrastructure) in order to achieve 'better' service quality. Thus, it can be assumed that service delivery costs typically rise with increasing service quality objectives. In contrast, total business costs can normally be expected to decrease with increasing service quality objectives. That is, the adverse impact of a service on business operations should usually decline with 'less imperfect' service quality (e.g., due to fewer service outages).

Figure 3.1 depicts an idealized[12] illustration of the SLE optimization problem for a particular service[13] considering one service quality measure only. Thus, a service offer $s$ is fully specified by a single service quality objective $x_s$ (with regard to the service quality measure) and the corresponding service delivery costs $c_s$. We assume the value range of service quality objectives to be limited to the pre-defined closed interval $X := [x_{min}; x_{max}]$. Reasons for such limitations could, for instance, be business requirements or technical constraints.[14]

Furthermore, two continuous cost curves describe service delivery costs and total business costs as functions of the service quality objective $x$ in Figure 3.1. This assumes there is a single service offer for each service quality objective. A third continuous curve in Figure 3.1 represents total service costs as the sum of service delivery costs and total business costs. Finally, the service offer assuring the service

[10] We elaborate on different detail levels of total business costs in Section 4.3.1.

[11] In general, costs are regarded as "the cash or cash equivalent value necessary to attain an objective such as acquiring goods and services, complying with a contract, performing a function, or producing and distributing a product" (Kinney & Raiborn, 2012, p.809).

[12] Service delivery costs rise and total business costs decrease with increasing service quality, as discussed above.

[13] That is, the goal is to select a single service offer from among different service offers.

[14] Business requirements or technical constraints may determine the minimum or maximum service quality level which is required or achievable.

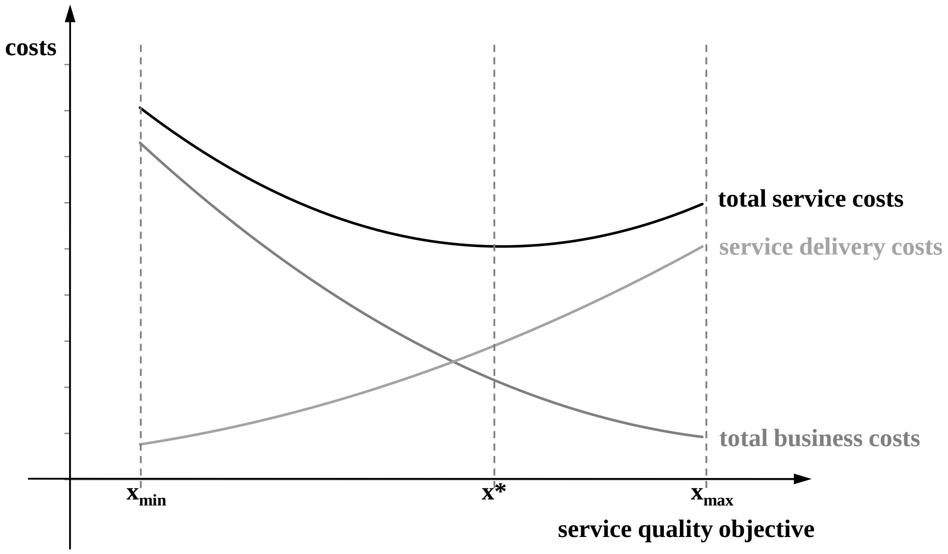


Figure 3.1.: The Generic SLE Optimization Problem

quality objective $x^*$, which induces the minimum amount of total service costs, is selected.

In this work, we assume business departments to use services to support business operations. We recall that a service provider may or may not belong to the same company as business departments do. Consequently, we distinguish two cases:

A provider usually has detailed knowledge about its own costs for delivering a specific service, which we denote as the provider's 'service provision costs'. Consequently, the provider assigns this private information to the variable 'service delivery costs' in Equation 3.1 in order to solve its SLE optimization problem.

> **Service provision costs:** Service provision costs are the actual, cash-effective costs which a provider incurs for delivering a service with specific service quality objectives. This private information is usually not disclosed to third parties but treated as confidential.

On the other hand, an external provider does normally not disclose this private information to its customers, since this mostly collides with its goal to maximize its own utility (Blau, 2009, p.68). That is, it might, for instance, reduce its profit.[15] Instead, providers typically quote a service price which comprises a profit mark-up.

> **Service price:** The service price is the amount of money a service customer has to pay an external provider in exchange for the delivery of a service with specific service quality targets.

From an external customer point of view, the cost-optimal service minimizes the sum of service price and total business costs (i.e., cash outflows). Therefore, addressing

[15] In some cases providers might even pretend to incur higher service delivery costs in negotiations in order to increase their profits.

its SLE optimization problem, the customer substitutes the variable 'service delivery costs' in Equation 3.1 by the service price.

Analogous to external providers treating service delivery costs as confidential, service customers may have concerns about disclosing information about total business costs to external providers as well. Third parties might use this information when negotiating service prices and, thus, potentially reduce service customers' profits. If sensitive information needs to be disclosed measures to avoid third parties to take advantage thereof have to be defined (as, for instance, the application of procurement auctions (see Chapter 7). In conformity with microeconomic theory, we denote the case of one party having more detailed information about something than another one (e.g., about the quality of a service) as a case of "asymmetric information" (e.g., Varian, 2010, p.718).

In Service Level Engineering we generally assume (internal and external) providers as well as business departments to be able to obtain appropriate information about their own service delivery costs or total business costs.

### 3.1.3. Selected Types of Service Level Engineering

In internal SLE settings we assume business departments and internal providers to work closely together in order to realize the cost-optimal service solution.[16] Therefore, business departments and providers should disclose total business costs and service delivery costs to each other. Thus, they would be able to jointly solve the SLE optimization problem.

In contrast, in an external setting the case of either party 'simply' disclosing its private information to the other party — i.e., without taking any precautions to avoid disadvantages — is rather hypothetical. Long-term service relationships between both parties might foster such close cooperation.[17] This proceeding would achieve the 'cost-efficient' solution of the SLE optimization problem, which maximizes social welfare.[18]

Thus, most likely, in external settings either external providers or service customers apply Service Level Engineering unilaterally. We briefly outline an external, unilateral, one-to-many setting, in which a service provider[19] aims to select the cost-optimal service solution:

We take the perspective of a single provider which competes with other providers for a service contract. We assume this provider to be able to offer different service solutions to meet the customer's business requirements. Thus, the provider applies Service Level Engineering in order to identify the particular service solution — among the potential service offers it could make — which maximizes its chance to close a

[16] We call this an internal, cooperative one-to-one SLE setting.

[17] This situation corresponds to an external, cooperative, one-to-one SLE setting.

[18] This would be the cost-optimal solution from a system point of view, since it minimizes the sum of total business costs and service delivery costs.

[19] In Section 3.2 we will change the 'perspective taken' and focus on a service customer's SLE optimization problem.

contract. Consequently, it regards a service offer as a combination of service price and service quality objectives.

The provider strives to identify the service solution which minimizes the sum of service price and total business costs — subject to the constraint that it achieves the intended profit margin.[20] Since the provider does usually not have detailed information about the customer's total business costs it has to estimate these for each potential service offer. Information about a customer's industrial sector and business processes to be supported should help with this task.

In an external, unilateral, one-to-one setting (e.g., based on a long-term service contract) the provider could apply SLE in order to satisfy its customer and make sure its exclusive contract will be renewed.

## 3.2. Business Setting: Selection of Cost-Optimal IT Service Solutions from a Service Customer Point of View

Service Level Engineering is applicable to a wide variety of settings. In this section we specify the particular IT outsourcing business setting which is addressed in the following chapters. We take the perspective of an IT department within a company, which aims to purchase a 'business service' (see 2.1.1) from an external provider. It is asked to provide this end-to-end IT service in order to support a particular business process. We assume the service to measurably contribute to the business value created.

Moreover, we assume the IT department and the business departments requesting the service to have defined all general properties of the service required[21] — except for service quality objectives and service price — according to business needs. Therefore, the business process to be supported has been sufficiently analyzed.

Various service offers by different competing providers meet the general business requirements but differ significantly in service quality objectives and service price. The service price given a distinct set of service quality objectives is subject to negotiation.[22]

We suppose that the rational IT department decides to apply a Service Level Engineering approach in order to minimize total service costs. Thus, the IT department needs to solve the following SLE optimization problem in the external, unilateral, one-to-many business setting described:

$$find : s^* := \operatorname*{argmin}_{s \in S} \Big( p_s + i(\vec{o_s}) \Big) \tag{3.2}$$

[20] This might possibly be a barrier to make an offer at all.

[21] General properties include the base sizing of the service required, i.e., a specification of the expected number of users, data volume etc.

[22] That is, due to the heterogeneity of service solutions (as to their implementation), a uniform market price for the service does not exist.

The IT department aims to select the cost-optimal service offer $s^*$ (from a set of service offers $S$), which minimizes total service costs. Each service offer $s$ by an external provider is a combination of a service price $p_s$ and a vector of service quality objectives[23] $\vec{o_s}$. The total service costs of a distinct service offer are the sum of service price and total business costs $i(.)$, where the IT department has to calculate total business costs based on the service quality objectives $\vec{o_s}$ of the particular service offer considered.

The approach needed to solve this SLE optimization problem should ideally support the determination of total service costs given a service offer. That is, it should provide guidance for IT and business departments on how to jointly assess the adverse impact of imperfect service quality on business operations. Furthermore, it should offer a method which facilitates and structures negotiations between the IT department and external providers — meanwhile enabling the selection of the cost-optimal service solution.

## 3.3. Related Fields of Research

We have defined Service Level Engineering as a field of research, which aims to determine business-relevant service quality measures and, based thereon, select cost-optimal service solutions. In order to achieve these main objectives, SLE can draw upon approaches from various fields of research. Additionally, there are areas of research which pursue goals similar to those of SLE. In this section we refer to selected fields of research which enable or complement Service Level Engineering.

### 3.3.1. Selected Fields of Research Enabling Service Level Engineering

There are numerous fields of research which offer models, methods and tools to support Service Level Engineering. Methods from analytics and statistics, for instance, enable providers to determine, describe and predict the service quality to be expected from a specific service solution. This information is essential with regard to the definition of service offers.

Similarly, business departments may apply approaches from the fields of business process modelling and simulation (e.g. discrete-event simulation) in order to formally describe and analyze their business processes. Thus, they can, inter alia, gain a more thorough understanding of the influence of imperfect service quality on business operations. This, in turn, supports the derivation of total business costs which a certain service solution induces.

Furthermore, approaches from the field of operations research can be applied to formally model and solve SLE optimization problems and, thus, select cost-optimal service solutions. The complexity of SLE optimization problems increases, inter alia, with the number of service quality measures to be considered.

[23] The vector of service quality objectives specifies one service quality target regarding each service quality measure.

In external Service Level Engineering settings service providers and customers usually negotiate service offers. Approaches from the field of game theory (e.g., auctions) can be leveraged to structure and conduct this process. Furthermore, these may be used to influence or forecast results of negotiations.

Finally, service providers frequently have to state service prices knowing that they have to compensate customers for potential non-achievement of service quality targets in the future. Approaches from the fields of decision and insurance theory can be applied to address this challenge.

### 3.3.2. Selected Complementary Fields of Research

Fields of research which are complementary to SLE (as to the goals they pursue) may address multiple areas of application (e.g., IT Management) and strive to achieve additional objectives (e.g., provider-internal improvements of IT service delivery processes).

**Business-Driven IT Management**

In recent years a group of authors around Sauvé, Moura and Bartolini established the field of 'Business-Driven IT Management' (e.g., Sauvé *et al.* (2006a), Moura *et al.* (2007) and Moura *et al.* (2008)). They characterize Business-Driven IT Management (BDIM) as a further development of IT service management, which applies "business metrics [...] such as cost or revenue" in order to assess the business impact of distinct levels of service quality (Sauvé *et al.*, 2006a, p.1).

More precisely, Sauvé *et al.* (2006a) define BDIM as "the application of a set of models, practices, techniques and tools to map and to quantitatively evaluate dependencies between IT solutions and business performance and using the quantified evaluation to improve the IT solutions' quality of service and related business results" (Sauvé *et al.*, 2006a, p.2).

BDIM encompasses all phases of the Plan, Do, Check, Act (PDCA) cycle with regard to IT solutions. That is, it considers the (technical) development and selection of an IT solution as well as its implementation, evaluation and, if needed, further adjustment. In contrast, Service Level Engineering focuses on the selection of the cost-optimal service solution with regard to service quality targets among existing service solutions. Thus, SLE particularly supports the PDCA phase 'plan' while abstracting from technical implementations and taking a 'service view'.

Therefore, we argue that Service Level Engineering is an area of research within the field of Business-Driven IT Management, which concentrates on the definition of business-relevant service quality metrics and, based thereon, the identification and selection of cost-optimal service solutions.

**Service Level Management**

As a "service design process" (Taylor *et al.*, 2007a, p.60), Service Level Management encompasses all tasks concerning the stipulation and adaptation of service quality measures and their target values (service quality objectives), the monitoring and reporting of achieved service quality and the control of associated service operation processes (e.g., Taylor *et al.*, 2007a, p.60, Unterharnscheidt & Kieninger, 2010).

Service Level Management mostly uses "technical metrics" to determine service quality (Sauvé *et al.*, 2006a, p.1). Consequently, service quality is usually not measured and defined from a business point of view. Furthermore, the influence of imperfect service quality on business operations is hardly considered in a quantitative manner.

In contrast, Service Level Engineering explicitly aims at the definition of business-relevant service quality measures. Furthermore, it particularly strives to quantify the business impact which service quality objectives entail. What is more, SLE, as opposed to SLM, does not comprise the control of (provider-internal) IT service management processes, which are related to Service Level Management.

Therefore, we state that SLE extends the field of SLM with regard to business-orientation. At the same time, SLE considers service quality resulting from specific implementations of "service operation processes" (Taylor *et al.*, 2007b, pp.33-77) but does not comprise their management.

**Service Engineering**

The German Institute for Standardization (1998) defines Service Engineering as a field of research which concerns the systematic development and design of services (through the application of appropriate models, methods and tools) and the management of processes required for this purpose. Service Engineering follows a "technical-methodological approach" and aims to apply knowledge from product development in order to create new kinds of services (Bullinger *et al.*, 2003, p.2).

Bullinger *et al.* (2003) differentiate three dimensions defining a service: First, the structure dimension addresses the allocation of resources. Second, the process dimension specifies the activities of service creation. Third, the outcome dimension describes services' results and effects. Thus, SLE particularly supports the outcome dimension.

Whereas Service Engineering focuses on the development of service solutions, SLE concentrates on the selection of cost-optimal service solutions with regard to service quality. Therefore, we regard Service Engineering and SLE as complementary fields of research.

**Quality Management**

The International Organization for Standardization (ISO) defines Quality Management to comprise all activities which aim to "direct and control an organization [...] with regard to quality" (International Organization for Standardization, 2005, p.21). It states that Quality Management encompasses, inter alia, the planning, definition, assurance and control of quality.

SLE, in turn, strives to identify and select cost-optimal service solutions with regard to service quality. Therefore, we argue that SLE supports Quality Management.

**Business Continuity Management**

Business Continuity Management is a "holistic management process that identifies potential threats[24] to an organization and the impacts to business operations those threats, if realized, might cause" (International Organization for Standardization, 2012a, p.9). Additionally, Business Continuity Management "provides the means for coping with consequences" of threats and takes "measures to reduce the likelihood and impact of incidents" (Cornish, 2011, p.122).

Similarly, SLE quantitatively assesses the adverse impact of imperfect service quality on business operations. SLE represents a 'measure' to manage the adverse impact of service quality on business operations (during service selection) and, thus, supports efficient Business Continuity Management.

**Operational Risk Management**

Operational risk management is regarded to address the risk of loss induced by "inadequate [...] internal processes, people and systems or [...] external events" (Basel Committee on Banking Supervision, 2006, p.144). It encompasses legal risks (e.g., due to pending litigation) but does not include "strategic and reputational risk" (ibid.). Exemplary activities of operational risk management are the identification of risks, the definition of risk control systems, the documentation and analysis of risk events as well as the (qualitative or quantitative) assessment of risks and their mitigation (Kenett & Raanan, 2011, pp.22-29).[25]

As opposed to operational risk management, SLE concentrates on the business impact resulting from imperfect service quality and does not take into account other customer-internal or -external risks. Furthermore, SLE also considers strategic and reputational consequences which are induced by service quality issues.

The relationship between SLE and the fields of research discussed above becomes obvious, since Service Level Engineering represents a business-oriented engineering approach to manage the quality levels of services. SLE considers 'service-induced risks' by analyzing the effects of imperfect service quality on the continuity of business operations.

## 3.4. Summary

In this chapter we first introduced Service Level Engineering as a field of research, which aims to determine business-relevant service quality measures and, based thereon, select — with regard to service quality — cost-optimal service solutions. In other words, SLE strives to solve a quantitative optimization problem and choose the

[24] In this context, a threat is regarded as a "potential cause of an unwanted incident, which can result in harm to individuals, a system or organization [...], the environment or the community" (International Organization for Standardization, 2012b, p.5).

[25] Generally, risk management is considered to comprise all "coordinated activities to direct and control an organization with regard to risk" (International Organization for Standardization, 2009, 2.2).

particular service offer for which the sum of service delivery costs and monetarily quantified, adverse business impact — resulting from imperfect service quality — reaches its minimum.

We denominate the cash-effective equivalent of this adverse business impact as 'total business costs'. Additionally, we refer to the cash-effective costs for delivering a service with specific service quality objectives as service delivery costs.

Furthermore, we discussed characteristics of SLE approaches, which can be used to classify and distinguish these from one another. These comprise, inter alia, the number of customers and providers considered, the competitive situation regarded and the perspective taken (e.g., a customer point of view). Afterwards, we briefly outlined selected types of Service Level Engineering.

Moreover, we discussed fields of research which enable or complement Service Level Engineering. Examples for areas of research which provide complementary approaches are Business-Driven IT Management, Service Level Management and Operational Risk Management. Enabling fields of research comprise, inter alia, statistics, business process modelling and simulation as well as operations research.

Finally, we specified the business setting, which is addressed throughout the following chapters. We take the perspective of an IT department within a company, which aims to purchase an IT service from one of several external providers. Business departments ask the IT department to provide this end-to-end IT service in order to support a particular business process. The rational IT department decides to apply Service Level Engineering in order to minimize total service costs and select the cost-optimal service offer. In the following chapter we further elaborate on the IT department's optimization problem.

# Part II.

# Development of the Method for Cost-Optimal Service Selection

# 4. Consideration of the Business Impact Induced by Service Incidents

Service Level Engineering suggests to weigh a service's price against its quality in order to achieve cost-optimal support of business operations. This, however, is only rarely done using structured approaches by IT departments today since they struggle to determine the total business costs induced by imperfect service quality.

After the general discussion of the IT departments' optimization problem in Chapter 3, we analyze service quality measures commonly used in practice as to their suitability to describe the business impact a service causes. In the scope of this work and when discussing the IT department's optimization problem, we exclusively consider service quality measures which describe the incident behavior of services. That is, we focus on measures characterizing service quality with regard to service incidents it entails (see Section 2.1.2).[1]

We demonstrate and discuss the IT department's need for new service quality measures in order to enable it to select cost-optimal service solutions. To address this issue, we suggest an advanced form of service quality measures, which we denote as 'service incident patterns'. Afterwards, we introduce a set of service incident-related concepts, which are essential for the definition of meaningful service quality measures. Furthermore, we describe the steps of the method for Cost-Optimal Service Selection, which builds on the concepts introduced, and refer to related work. Finally, as an excursus, we identify challenges for service providers emerging from the application of service incident patterns.

This chapter forms the basis of the method for Cost-Optimal Service Selection, which addresses the business setting described in Section 3.2.[2]

---

[1] Examples are measures of service availability (outage incidents) and of service response time (response time incidents).

[2] Section 4.1 and Section 4.3.1 are based on Kieninger *et al.* (2012a) and Kieninger *et al.* (2013a).

## 4.1. A Deficiency of Current Service Quality Measures

In practice service level indicators are defined in order to measure and stipulate service quality (e.g., Berger, 2007, pp.113-160, Sturm *et al.*, 2000, pp.66-67). Service level indicators which describe service incident behavior typically aggregate the attribute values of service incidents they are based on into a single value (see Section 2.1.2), namely the actual service level. From a service customer point of view this aggregation results in a loss of information, which may have undesirable consequences regarding the selection of service solutions — as we will see in the following.

The application of aggregating service level indicators implies that IT departments do not have any knowledge of the probabilities of a service to achieve, fall below or even exceed the service level objectives defined. As a consequence, IT departments cannot appropriately assess the total business costs, which the selection of a particular service offer entails. But, what is more, IT departments may even be unable to assess total business costs of a service offer if SLOs are exactly met. This is due to the fact that adverse impact on business operations depends on the particular set of service incidents which occurs.[3] Each service incident induces a specific business impact.

Therefore, we assume business departments to incur business costs for each single service incident that occurs. These business costs strongly depend on the values of a service incident's attributes (e.g., its duration) as well as on the service incident's time of occurrence.[4]

We introduce the concept of 'business cost functions', which describe the business costs induced by a single service incident with respect to its attribute values and time of occurrence. We define that a business cost function refers to a specific period of time and a single type of service incidents only.

> **Business cost function:** A business cost function describes the cash-effective adverse impact on business operations induced by a single service incident with respect to its attribute values and time of occurrence. Each business cost function refers to a specific period of time and a single type of service incidents only.

We note that business cost functions may develop disproportionally with a service incident's attribute values — they may, for instance, decrease or increase non-linearly. The following example, which refers to a single time period and disregards that business costs may vary over time, illustrates this consideration.

*A management system (provided as an IT service) is used in a warehouse to locate goods, which are to be shipped by truck. Trucks arrive at five-minute intervals in order to pick up goods. This shipping process is interrupted in case of outages of the warehouse management system, i.e., if service (outage) incidents occur. In a specific*

[3] For the definition of the concept 'service incident' see Section 2.1.2.
[4] The influence of a service incident may, e.g., be higher on a Monday morning than on a Friday evening.

*time period, every minute a truck has to stay longer due to delays in loading induces business costs of €1. Figure 4.1 shows the corresponding business cost function.*[5]

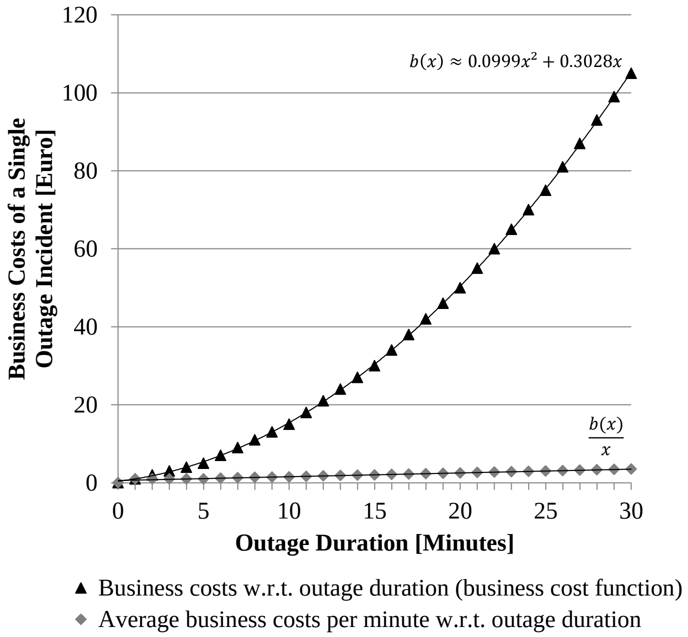


Figure 4.1.: Business Costs of a Single Outage Incident w.r.t. Outage Duration

*The black triangles in Figure 4.1 each describe the business costs which the logistics company incurs in case of a single service incident with a certain duration (business cost function).*[6] *The grey diamonds represent the 'average business costs per minute' which a single outage incident with a distinct duration induces.*[7] *An outage incident with a duration of 20 minutes, for instance, causes €50 of business costs. This corresponds to average business costs of €2.5 per minute of outage.*

Service level indicators aggregate the values of a particular service incident attribute (with respect to a specific type of service incident) into a single value, namely the actual service level (see Section 2.1.2). A certain actual service level (ASL) can, thus, result from different sets of attribute values.

In case of business cost functions developing proportionally with service incidents' attribute values, total business costs can be assessed without detailed information about the set of attribute values realizing a certain ASL. We illustrate this continuing our example.

*We assume the IT service to be available for exactly 99% of a month,*[8] *i.e., to be unavailable for exactly 432 minutes (total outage time) within this reference period.*

5 For reasons of simplicity, we assume that all trucks waiting can be loaded at once as soon as the IT service is available again (that is, without incurring further business costs).

6 The continuous business cost function $b(x) = 0.0999x^2 + 0.3028x$ approximates the discrete business cost values (black triangles).

7 The continuous function $a(x) = 0.0975x + 0.4786 \approx \frac{b(x)}{x}$ approximates the average business costs per minute (grey diamonds).

8 We assume a month to have 30 days and suppose an agreed service time of 43,200 minutes

*Furthermore, we suppose that no single service outage incident lasts longer than 30 minutes.*[9]

*Let $T$ be the length of the reference period and $x_{max}$ be the maximum length of a single outage incident. Furthermore, let $x_i$ be the length of outage $i$ (measured in minutes) during $T$ and let $b(x_i) = x_i \cdot 1\frac{Euro}{min}$ be the proportional business cost function. Now assume $x_i \leq x_{max}$ and $x_i > 0$, $\forall i$, and, $x_i = x_{fix}, \forall i$, i.e., all outage incidents to have the same duration. Then different values of $x_{fix}$ entail the same total business costs (see Table 4.1).*

Table 4.1.: Total Business Costs in Case of a Proportional Business Cost Function

| $x_{fix}$ | $b(x_{fix})$ [Euro] | number of incidents within T | total business costs caused within T [Euro] |
|---|---|---|---|
| 4 | 4.00 | 108 | 432.00 |
| 8 | 8.00 | 54 | 432.00 |
| 16 | 16.00 | 27 | 432.00 |

We recognize, however, that a precise determination of business costs resulting from the achievement of a specific ASL is not possible if business cost functions develop disproportionally — unless the combination of service incidents' attribute values (e.g., their individual outage durations) is known:

*Let $b(x_i) = 0.0999x_i^2 \cdot 1\frac{Euro}{min^2} + 0.3028x_i \cdot 1\frac{Euro}{min}$ now be the disproportional business cost function depicted in Figure 4.1. Then different values of $x_{fix}$ induce significantly different total business costs (see Table 4.2).*

Table 4.2.: Total Business Costs in Case of a Non-Linear Business Cost Function

| $x_{fix}$ | $b(x_{fix})$ [Euro] | number of incidents within T | total business costs caused within T [Euro] |
|---|---|---|---|
| 4 | 2.81 | 108 | 303.44 |
| 8 | 8.82 | 54 | 476.06 |
| 16 | 30.42 | 27 | 821.32 |

This simple example already shows the influence of different outage durations on business costs. We now consider a more realistic setting and drop the assumption that all outages which occur within $T$ are of equal length:

*Figure 4.2 depicts the probability densities of different outage durations $x$ for selected generalized beta-distributions.*[10] *The parameters $\alpha$ and $\beta$ define the shape of the probability density functions $w(x, \alpha, \beta, 0, 30)$. We limit the probability density functions*

(twenty-four-seven service). Section 2.1.2 provides an example calculation formula for a service level indicator to measure 'service availability', which we use in this example.

[9] This threshold value represents a service incident objective (see Section 2.1.2).

[10] We use the generalized beta-distribution to be able to model a large variety of combinations of outage incident durations. In addition, the generalized beta-distribution can be defined on finite intervals, is easily adaptable and is analytically tractable.

*to the interval* $(0; 30]$ *(see the fourth and the fifth parameter of w(.)) since we have limited the maximum duration of single outage incidents to 30 minutes.*

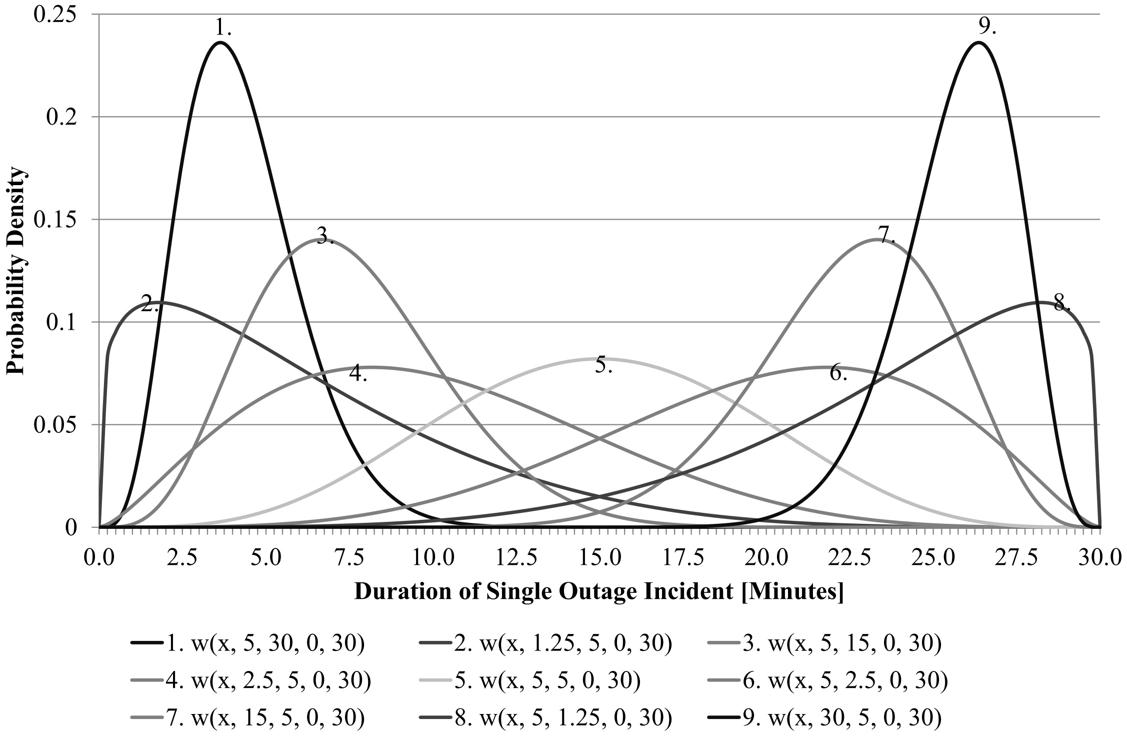


Figure 4.2.: Probability Densities of Beta-Distributed Outage Incident Durations

*Having assumed the service to be unavailable for exactly 432 minutes we calculate the monetarily quantified business impact i(.) which the different generalized beta-distributions (of single outage durations) cause. Therefore, for each beta-distribution (specified through its parameter values* $\alpha$ *and* $\beta$*) we compute the integral defined in Equation 4.1.*[11]

$$i(\alpha, \beta) = \int_{0}^{30} w(x, \alpha, \beta, 0, 30) \cdot 432 \cdot \frac{b(x)}{x}\, dx \tag{4.1}$$

*In Equation 4.1, the term* $\frac{b(x)}{x}$ *represents the average business costs per minute (delineated as grey diamonds in Figure 4.1). Multiplying the probability density function of a beta-distribution with the total outage duration (i.e., 432 minutes), we calculate the 'time slice'*[12] *which is spent on a distinct outage duration. Afterwards, we weigh this time slice with the corresponding average*[13] *business costs per minute.*

[11] For reasons of clarity, units of measurement are omitted in the following.

[12] That is, we consider the limiting behavior of the distribution.

[13] The time slice represents the number of minutes which is spent on outages with a certain duration but does not indicate the number of outage incidents with a specific duration. Therefore, we use the *average* business costs per minute of outage (given an outage incident with a certain duration) in order to calculate the business costs the time slice induces.

*Figure 4.3 depicts the business costs caused by each beta-distribution.*

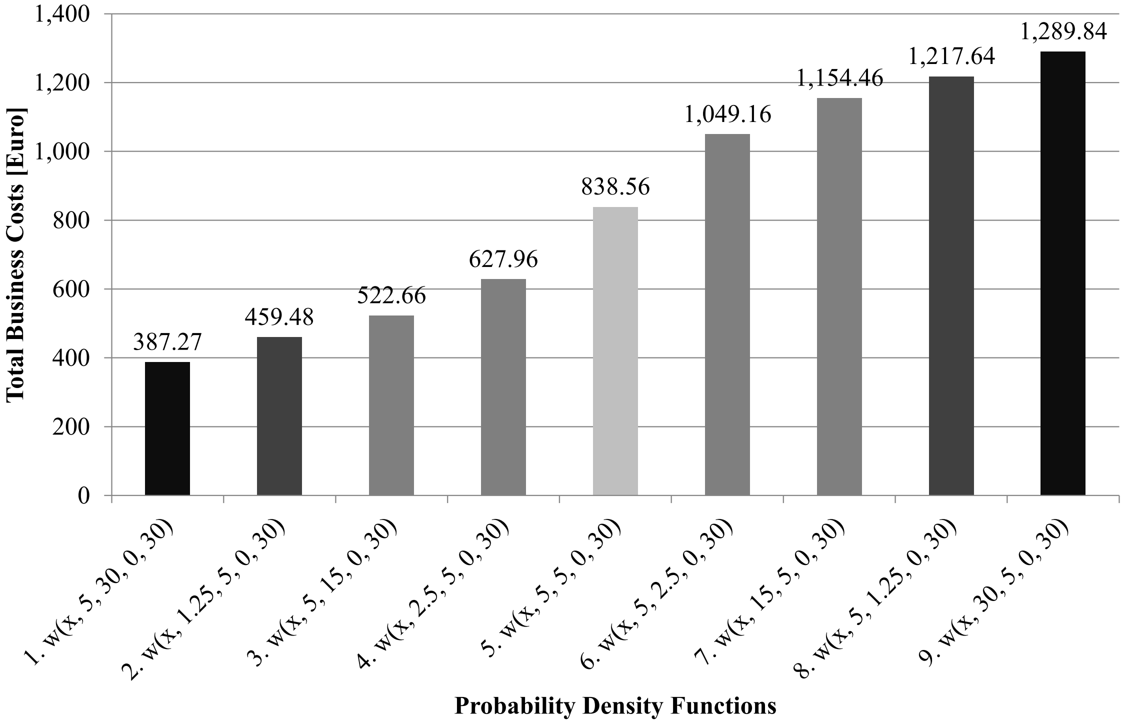


Figure 4.3.: Expected Business Costs given Different Probability Density Functions

Our examples show that already in case of moderately non-linear business cost functions the distribution of incident attribute values has to be considered in order to determine business costs. Different distributions of incident attribute values may induce significantly different total business costs. We further illustrate this finding:

*Let us assume we have to assess a service offer with regard to its adverse business impact and we know that single incident durations follow a generalized beta-distribution. Even if we have this information, the assumption of incorrect beta-distribution parameter values may lead to significant discrepancies between the total business costs we expect and the total business costs we actually incur (see Table 4.3).*

Table 4.3.: Consequences of Inaccurate Estimation of Service Incident Behavior

| | | Assumed Probability Density Function | | |
|---|---|---|---|---|
| | | w(x,5,30,0,30) | w(x,5,5,0,30) | w(x,30,5,0,30) |
| Actual Probability Density Function | w(x,5,30,0,30) | 0.0% | 116.5% | 233.1% |
| | w(x,5,15,0,30) | -25.9% | 60.4% | 146.7% |
| | w(x,5,5,0,30) | -53.8% | 0.0% | 53.8% |
| | w(x,15,5,0,30) | -66.5% | -27.4% | 11.7% |
| | w(x,30,5,0,30) | -70.0% | -35.0% | 0.0% |

Consequently, we state that service level indicators are only appropriate to select cost-optimal IT services in case of business impact developing proportionally with all attribute values of single service incidents (e.g., their duration). This is the case, if the first derivatives of business cost functions (with respect to their incident attributes) are constant values. We suppose such constant 'marginal business cost functions' not to be the standard case in practice, however. They would, for instance, require that every further minute of duration of an ongoing outage incident induces the same business impact (as in our simple example above; see Table 4.1).

Thus, service level indicators commonly applied in practice today may leave IT departments uninformed about a service's influence on business operations — even if we assume service level objectives to be exactly met. The fact that services also frequently under- or even outperform stipulated SLOs in practice makes it even more difficult to estimate total business costs.

Consequently, the usage of service level indicators may impede the selection of cost-optimal service offers. Service quality measures have to allow a precise assessment of the business impact resulting from imperfect service quality and may not aggregate incident attribute values only. Service availability, for instance, may not be regarded as a "mere percentage" of total service time only (Franke, 2012, p.23). Henceforth, we explicitly concentrate on the case of non-constant marginal business cost functions.

## 4.2. Service Incident Patterns — An Advanced Form of Service Quality Measures

In order to address the deficiency of service level indicators discussed, we recommend the application of 'service incident distributions' to precisely describe the combinations of incident attribute values which are characteristic for a service — i.e., to specify its characteristic 'service incident patterns'. First, we develop the concepts of service incident distribution and service incident pattern. Afterwards, we discuss prerequisites for their application.

### 4.2.1. Characteristics of Service Incident Patterns

In this work, we suggest service incident patterns as an advanced type of incident-based, non-aggregating service quality measures.[14] Service incident patterns may refer to specific time periods within a reference period only. Therefore, we define a separate service incident distribution and a corresponding reference value (e.g., the 432 minutes of total outage duration in our example above) for each combination of service incident type and time period.

> **Service incident pattern:** A service incident pattern represents the characteristic combination of a service's incident attribute values with respect to a distinct type of service incidents (i.e., they specify a service's incident behavior). A service incident pattern can be precisely described through a service incident distribution and its corresponding reference value and refers to a specific time period.

[14] For a classification of service quality measures see Section 2.1.2.

Service incident distribution functions (as well as business cost functions) can either be discrete or continuous. If the value range of a service incident attribute (characterizing a specific type of service incident) is divided into a set of disjoint intervals we use discrete distribution functions and discrete business cost functions.[15]

In Equation 4.2 we describe the IT department's SLE optimization problem using service incident distributions and corresponding reference values as service quality measures. The equation considers multiple classes of service incidents[16] $(c_1, ..., c_n)$, which are specified through the values of a single attribute[17] each $(x_1, ..., x_n)$.

We note that Equation 4.2 assumes services to exactly meet the service incident patterns promised. In case an IT department (aiming to optimize total service costs) expects deviations from service incident patterns promised in service offers, it would consider penalty and benefit payments as well. The IT department might, for instance, receive penalty payments if a service underperforms the stated service incident patterns and could have to pay the provider a bonus if a service outperforms these.[18]

$$find : s^* := \operatorname*{argmin}_{s \in S} \Bigg( \Bigg( \int_{x_{1_{min}}}^{x_{1_{max}}} w_{s,c_1}(x_1) \cdot r_{s,c_1} \cdot b_{c_1}(x_1)\ dx_1 + \int_{x_{2_{min}}}^{x_{2_{max}}} w_{s,c_2}(x_2) \cdot r_{s,c_2} \cdot b_{c_2}(x_2)\ dx_2 + ... + \int_{x_{n_{min}}}^{x_{n_{max}}} w_{s,c_n}(x_n) \cdot r_{s,c_n} \cdot b_{c_n}(x_n)\ dx_n \Bigg) + p_s \Bigg) \tag{4.2}$$

The IT department aims to select the cost-optimal service solution $s^*$ (from a set of service offers $S$), which minimizes total service costs. The total service costs of a distinct service offer $s$ are the sum of service price $p_s$ and total business costs. The latter are calculated as follows:

The terms $w_{s,c_1}(x_1), ..., w_{s,c_n}(x_n)$ represent the probability density functions as to the service incident classes $c_1, ...c_n$ of the service $s$. These functions are multiplied with the corresponding reference values $r_{s,c_1}, ..., r_{s,c_n}$ (e.g., the total outage duration in our example above). Afterwards, the result is weighed with the corresponding business cost functions $b_{c_1}(x_1), ..., b_{c_n}(x_n)$. In order to obtain total business costs

15 The categorization of attribute values into intervals reduces the complexity of the approach and, thus, enhances its applicability in practice.

16 Exemplary classes of service incidents are 'service outage incidents', 'reduced response time incidents' and 'reduced throughput incidents'.

17 The duration of a service incident is an example of such an attribute.

18 We discuss how an IT department can use penalty and reward rules to find out about the total business costs a service offer entails in Chapter 7 of this work.

of the service we integrate over all possible attribute values ($x_{i_{min}}$ to $x_{i_{max}}$) of each service incident class ($c_i, i \in [1, 2, ..., n]$).[19]

Therefore, each business cost function has to be defined in a way that its measurement unit matches the one of the corresponding reference value. If, for instance, the measurement unit of the reference value is 'minutes' and the measurement unit of the business cost function is 'Euro per incident', the business cost function has to be normalized accordingly (to 'Euro per minute of incident duration').[20] In the general case described in Equation 4.2, no adaptation of the business cost functions is needed, since the units of measurement of the reference values and the business cost functions, are 'number of incidents' and 'Euro per incident'. That is, they both refer to a number of incidents.

In Appendix A we present different forms of the IT department's optimization problem. We provide equations for the cases of continuous and discrete service incident distribution functions. Furthermore, we consider that service incident patterns and business cost functions may vary over time.[21] Additionally, we elaborate on the case that a certain type of service incident may be specified through multiple attributes.

### 4.2.2. Prerequisites for Cost-Optimal Service Selection

In order to calculate total business costs for a service offer which is based on service incident patterns, an IT department needs knowledge of business cost functions as well as truthful information about services' incident behavior. This implies that service providers must not have any incentive to state service incident patterns incorrectly. We address this challenge when discussing penalty- and reward-rules to be applied in case of deviations from assured service quality in Section 7.1.

Furthermore, in our approach we hitherto supposed service solutions' incident behavior to be stable over time. If service incident behavior is subject to fluctuation, however, IT departments may define penalty- and reward-rules in a way which ensures that they are still able to calculate service solutions' total service costs.[22] Thus, providers are forced to consider expected future penalty- and reward-payments when setting service prices. Consequently, the effect of fluctuations in service quality can be reflected in total service costs as well.

Moreover, we require IT departments to be able to state business cost functions. In Chapter 5 we provide an overview on classifications of business consequences and corresponding costs, which may serve IT and business departments as a means to jointly identify and classify business consequences in practice. Furthermore, in

19 For reasons of clarity, we denote the attribute values of a service incident class $c_i$ as $x_i$ instead of writing $x_{c_i}$.

20 In our example above, we divided the business cost function $b(x)$ by $x$ to obtain the average business costs per minute of outage for every outage duration.

21 Business cost functions may refer to certain time periods within a reference period only.

22 In Section 7.1, we discuss this prerequisite for the selection of cost-optimal service solutions in detail. Usually, the purpose of penalties is only to incentivize providers to achieve the promised service quality levels (e.g., Sturm *et al.*, 2000, p.68).

Chapter 6 we present a business process simulation-based approach which supports the determination of single service incidents' impact on business operations. We note that business cost functions have to be determined at reasonable costs. Otherwise savings achieved through SLE with regard to total service costs might be lost. Thus, a good estimate of business cost functions may prove better than a precise determination from a financial point of view.

Introducing our service incident pattern approach, we implicitly assume service incidents to be independent of one another. This assumption seems reasonable, since single service incidents are normally of short duration compared to a reference period's length. Consequently, the influence of a service incident on business operations has typically already disappeared when the next service incident occurs. Furthermore, the number of service incidents to occur during a reference period is usually small. If we need to consider service incidents which occur simultaneously as well, business costs have to be determined for each tuple of such interrelated service incidents.[23] In this case, a tuple is characterized by the set of service incidents' attribute values. Accordingly, the frequency of occurrence for each set of interrelated service incidents has to be specified. This, however, considerably increases the complexity of our approach.[24]

## 4.3. Concepts to Consider for Cost-Optimal Service Selection

After we have motivated the usage of service incident patterns, we now elaborate on further characteristics of services and of the business environments these shall support, which are essential for solving the IT department's SLE optimization problem.[25] Furthermore, we discuss the interdependencies between these concepts in a more formal way.

First, we have a closer look at the characteristics of service incidents (Section 4.3.1). Afterwards, we examine which concepts have to be considered in order to formally describe the influence of a single service incident on business operations (Section 4.3.2).

### 4.3.1. Characteristics of Service Incidents

We group service incidents which share the same set of attributes into the same 'service incident class' (e.g., outage incidents or reduced throughput incidents). In other words, we define service incident classes in such a way that their elements

[23] This may be the case, if combinations of service incidents cause a business impact which significantly differs from the 'sum of business impacts' that the single service incidents would induce.

[24] In Section 4.6 we will discuss prerequisites for the application of service incident patterns as service quality measures from a provider point of view. We will describe how a provider may proceed to derive service incident patterns of a specific service solution. Furthermore, we will review provider challenges which emerge from the usage of service incident patterns.

[25] Some of these concepts have already been briefly discussed when we introduced service incident patterns and the IT department's SLE optimization problem.

affect business operations in a similar way. This still allows for service incidents of the same service incident class to differ in their attribute values (e.g., in the duration of outage regarding the class 'outage incidents'). Furthermore, we denote a tuple of attribute values of a single service incident as its 'service incident level'.

*For example, we could use the attributes 'reduced throughput incident duration' and 'reduced throughput incident degree' to describe service incidents of the class 'reduced throughput incident'. Then, a pair of attribute values regarding a single service incident would represent its service incident level (e.g., a duration of three minutes and a degree of 900 Mbit/s or a duration of two minutes and a degree of 500 Mbit/s).*

Figure 4.4 illustrates the interdependencies between the concepts introduced in the form of an 'entity relationship diagram' (e.g., Thalheim, 2000).[26] It explains that a single service incident with a specific service incident level which occurs at a certain time and that belongs to a certain service incident class induces business costs.

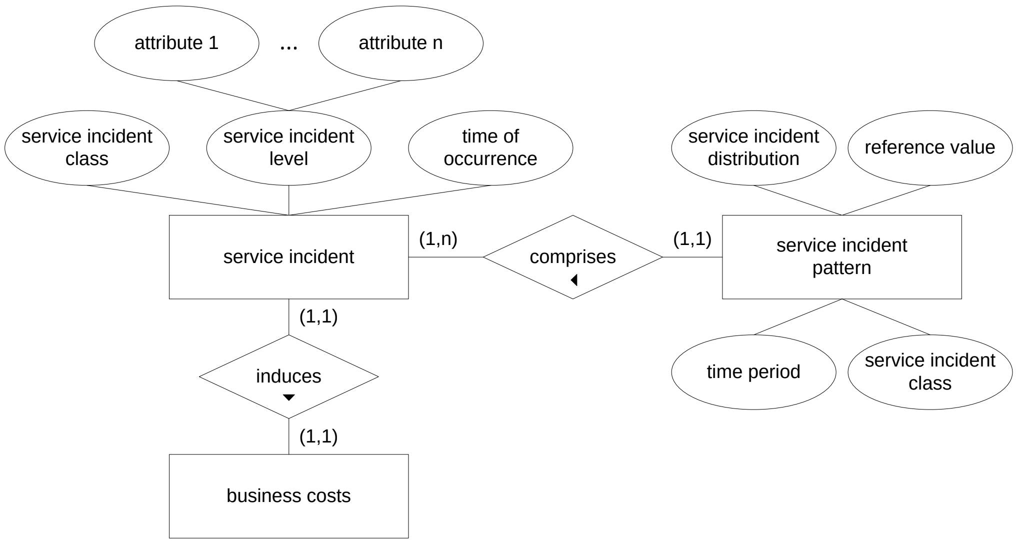


Figure 4.4.: Characteristics of Service Incidents

Each service incident belongs to one service incident class only. Furthermore, all incidents of a certain class share the same set of attributes. The tuple of attribute values of a single service incident represents its particular service incident level. Moreover, the business costs a single service incident induces are determined by its service incident level as well as its time of occurrence. Finally, a service incident pattern is a set of all service incidents of a certain class which occur during the specific time period the pattern refers to. A service incident pattern can be described through a tuple of a (discrete or continuous) service incident distribution and a corresponding reference value.[27]

26 In this work, we use entity relationship diagrams to depict interdependencies between different concepts. We note that these diagrams are provided for illustrative purposes only and may not comply with formal requirements for database design.

27 The attribute 'service incident class' of the entity 'service incident' has the same value range as the attribute 'service incident class' of the entity 'service incident pattern'.

In the following we differentiate between business costs induced by a single service incident, by a service incident pattern and by a set of service incident patterns. Therefore, we denote the business costs induced by a specific service incident pattern — referring to a certain time period — as 'aggregated business costs' . Furthermore, we refer to the business costs induced by the set of all service incident patterns — again referring to the same time period and describing the behavior of a distinct service — as its 'total business costs' . The following entity relationship diagram depicts the interrelations between these concepts (see Figure 4.5). Additionally, it explains how 'service price' and 'total business costs' add up to 'total service costs'.

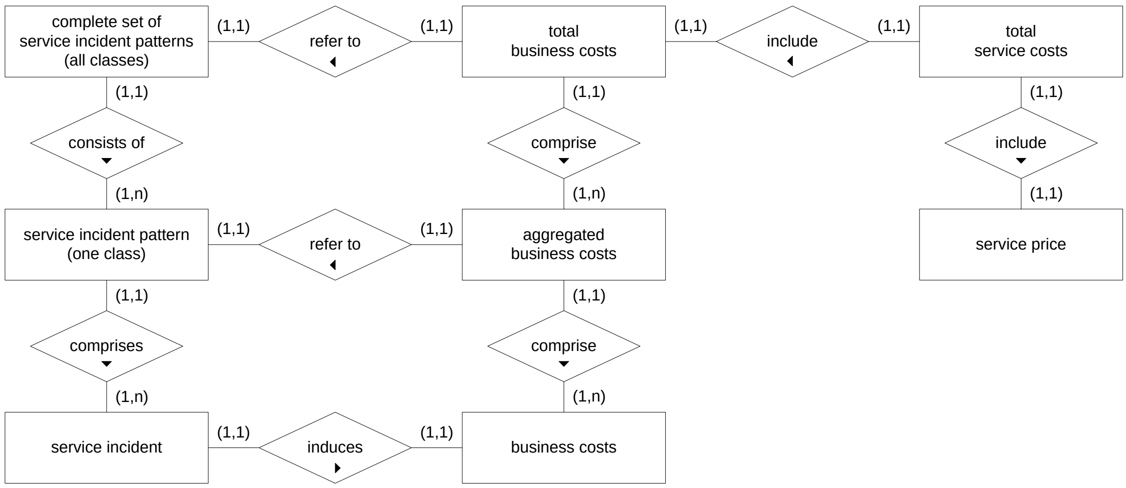


Figure 4.5.: Aggregation of Business Costs

In the following section we further elaborate on the relationship between a single service incident and the business costs it induces.

### 4.3.2. The Relationship between Service Incident and Business Costs

A single service incident may have different effects on business operations,[28] which we denote as 'business consequences' . IT departments have to monetarily assess these business consequences (e.g., in terms of business costs) in order to be able to weigh service quality against service price.

The established 'business impact analysis' (BIA) approach (e.g., Barnes, 2011) supports the (financial) assessment of business consequences.[29] It analyzes "business functions" with regard to the effects that a "business disruption might have upon them" (Business Continuity Institute, 2011, p.12) meanwhile considering the development of these business consequences over time (e.g., Cornish, 2011, p.129).

[28] Accordingly, literature on risk management describes events and resulting effects, whereas an event may be triggered by different causes (e.g., Dowd, 2003, p.31)

[29] Business impact analysis is usually applied in order to identify "most critical business functions" and define business continuity strategies with regard to these (e.g., Taylor *et al.*, 2007a, p.209).

Contrary to other (risk management) approaches, BIA concentrates on business consequences instead of events causing these (e.g., Barnes, 2011). This focus reduces the risk of neglecting crucial business consequences due to the consideration of specific types of events only (von Rössing, 2005, p.77).

Emphasizing the importance of a focus on business processes, Wiedemann (2008) suggests a three step BIA process (see Figure 4.6) which we adapt to the Service Level Engineering terminology in the following. Wiedemann focuses on the application of the BIA process in the field of Business Continuity Management dealing with rare service incidents, which have a severe impact on business operations (Wiedemann, 2008, p.40).

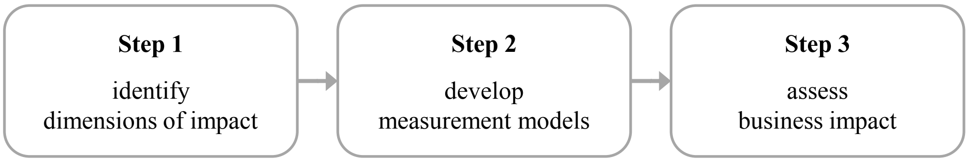


Figure 4.6.: BIA process by Wiedemann (2008, p.122)

First, the business consequences (‘dimensions of impact’), which are induced by a disruption of business processes, are identified. Examples of such business consequences are ‘lost sales’ and ‘dissatisfaction of stakeholders’.[30]

Second, for each business consequence a ‘measurement model’ is developed. It describes the calculation of financial impact (concerning a distinct business consequence) as a function of one or several ‘business parameters’. For this purpose, the relevant business parameters of each business consequence are determined. The loss in revenue might, for instance, depend on the reduction in output due to a service incident or the length of time during which a customer cannot order products.

Third, the values of all business parameters (more precisely, their development during the disruption of business operations, i.e., over time) are determined. These values are obtained through expert interviews and document analyses (for every measurement model). One might, for example, find that the reduction in output amounts to 100 units (value of the business parameter) due to a service incident.

Finally, business costs are calculated based on business parameters’ values as described in the corresponding measurement models.

We use the concepts introduced to precisely describe the link between a service incident (which occurs at a certain time and has a distinct service incident level) and the business costs it induces (see the entity relationship diagram in Figure 4.7). A service incident causes business consequences, which we assess in terms of business costs. Due to the close relationship between business consequence and business costs we also say that a service incident induces business costs.

[30] Chapter 5 provides an overview on comprehensive classifications of business consequences and corresponding business costs.

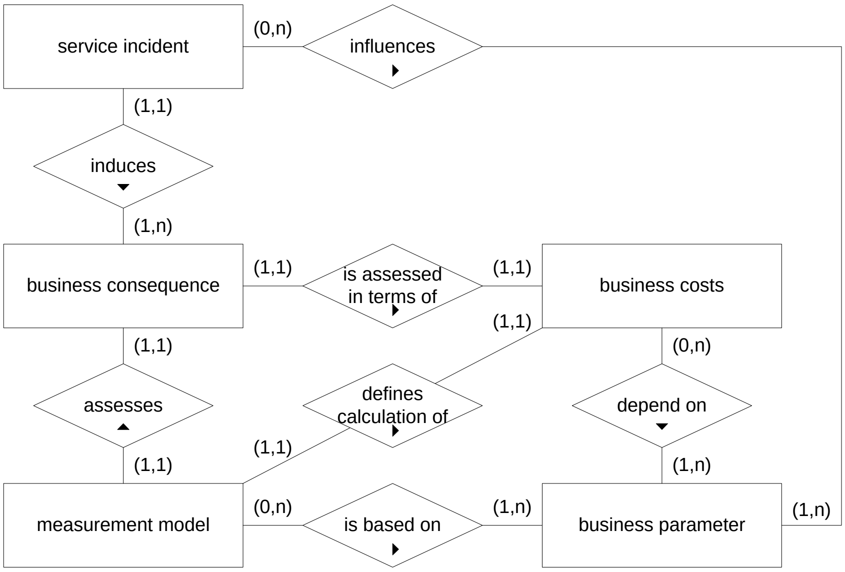


Figure 4.7.: Relationship between Service Incident and Business Costs

The influence of a service incident on the value of a certain business parameter depends on its service incident level as well as on its time of occurrence. We notice that a service incident may affect certain business parameters only. That is, the value of a specific business parameter may be influenced by a service incident (e.g., the value of the business parameter 'output reduction') while the values of others may not (e.g., the value of the business parameter 'product price'). By definition, each service incident at least affects the value of one business parameter.

As described above, a measurement model defines how to calculate the business costs resulting from a certain business consequence. For this purpose, it considers the values of all business parameters which are affected by the particular service incident causing the business consequence. The following simple example demonstrates the meaning of the concepts discussed:

*Let us assume that the disruption of a production process immediately results in lost sales (business consequence) for a service customer. The corresponding measurement model states that the financial impact is calculated by multiplying the values of two business parameters, namely the reduction in output due to the disruption and the product price. The production process depends on an IT service which, in turn, may be disrupted by service incidents. Now, an outage incident with a duration of 10 minutes (service incident level) occurs. This outage incident reduces the output of the production process by 20 units (value of the business parameter 'output reduction') but does not influence the product price, which amounts to €50. Consequently, the service incident induces business costs of €1,000 with regard to the business consequence 'lost sales'.*

Furthermore, we introduce the concept of 'business parameter functions':

> **Business parameter function:** A business parameter function describes the influence of a single service incident on the value of a specific business parameter as a function of the service incident's level and time of occurrence. Each business parameter function refers to a single service incident class only.

The derivation of business parameter functions forms the basis for the determination of business cost functions, since business costs are calculated based on business parameters' values. We note that business cost functions describe the adverse impact of single service incidents of a specific service incident class considering all business consequences. In contrast, measurement models define the calculation of business costs with regard to a single business consequence and all classes of service incidents. Consequently, the information contained in measurement models is used for the definition of business cost functions as well.

## 4.4. COSS - A Procedure to Select Cost-Optimal Service Offers

In this section we present our method for 'Cost-Optimal Service Selection' (COSS), which is based on the Service Level Engineering approach as well as the service incident- and business cost-related concepts described in the previous section. Furthermore, COSS builds on the BIA process by Wiedemann (2008, p.122, see Section 4.3) and consists of two main steps (see Figure 4.8), which we describe in the following.

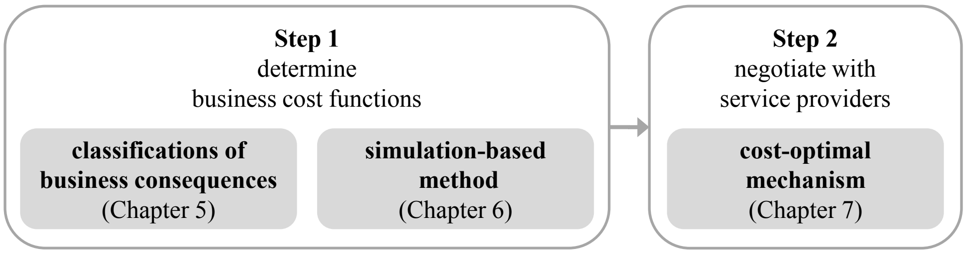


Figure 4.8.: COSS — A Method for Cost-Optimal Service Selection

### Step 1: Determine Business Cost Functions

First, IT and business departments jointly determine the business cost functions for each service incident class. Therefore, they measure the impact of single service incidents[31] with regard to business consequences using a business process simulation-based method, which we present in Chapter 6.

[31] More precisely, they measure the impact of service incidents of *the particular service to be procured.*

Following this method, IT and business departments initially identify business consequences which are potentially induced by a disruption of business processes in the business setting under consideration.[32] They collect this information through expert interviews and document analyses.

In order to facilitate this initial task of our simulation-based method, we provide an overview on classifications of business consequences and corresponding costs (see Chapter 5). Suggesting business consequences which might be of relevance in a specific business setting these classifications serve IT and business departments as a means to identify and classify business consequences in practice.

Afterwards, IT and business departments jointly define a measurement model (see Section 4.3.2) for each relevant business consequence identified and determine (or, if necessary, estimate) values of the corresponding business parameter functions. Finally, they determine the business cost function of each service incident class.[33]

We support the determination of distinct values of business parameter functions through discrete-event business process simulation. Hence, our simulation-based method facilitates the determination of business cost functions in case of well-defined business processes. It complements the approach by Wiedemann, which suggests to assess the values of business parameters through expert interviews and document analyses.

#### Step 2: Negotiate Service Offers with Potential Service Providers

Second, the IT department applies a cost-optimal mechanism which supports the negotiation of service offers with multiple risk-neutral providers (see Chapter 7). This allows the IT department to select the cost-optimal service solution among different provider offers.[34]

Following our mechanism the IT department invites selected providers to take part in a service procurement auction. Providers are asked to bid tuples of service incident patterns (one pattern for each service incident class) and a service price.

We suggest IT departments to define *specific* penalty and reward rules which are identical with business cost functions.[35] These rules should be applied in case a service solution deviates from the assured service quality in the future. Thus, risk-neutral providers do not have an incentive to state service incident patterns incorrectly. Furthermore, our penalty and reward rules ensure that an IT department is still able to calculate service solutions' total service costs in case of fluctuations in service quality.

To summarize, COSS first supports IT and business departments in discussing and assessing the business impact of IT services. The application of classifications of

[32] This task corresponds to the first activity of the BIA process by Wiedemann.
[33] These tasks encompass activities two and three of the BIA process by Wiedemann.
[34] That is, we further extend the BIA process by Wiedemann.
[35] In Section 7.1 we elaborate on the implications of different penalty and reward rules on providers' bidding behavior.

business consequences (see Chapter 5) provides advice about which cash-effective business consequences to consider and business process simulation allows the determination of single service incidents' impact on business parameter values. Second, using our cost-optimal, auction-based mechanism, COSS enables IT departments to select the optimal service solution with regard to total service costs among different provider offers.

COSS offers approaches which support the BIA process activities 1 and 3 (i.e., 'identify dimensions of impact' and 'assess business impact'). Furthermore, COSS extends the BIA process through capturing negotiation with providers as well.

Before we develop the supporting approaches which enable COSS (simulation-based method and auction mechanism), we refer to related work (see Section 4.5) and discuss challenges for service providers emerging from the application of service incident patterns (see Section 4.6).

## 4.5. Related Work

In this section we present related work which concentrates on the selection of optimal service quality objectives, which may thus be attributed to the field of Service Level Management.[36] The articles discussed in the following represent the selection of most closely related publications from the results of an extensive literature review.

We identified these articles following the approaches suggested by vom Brocke *et al.* (2009) and Webster & Watson (2002). We used Google Scholar as primary source and conducted a keyword-based full-text search without temporal restrictions. Search terms applied were 'business impact', 'business loss', 'business cost', 'service interruption', 'service incident', 'service level', 'incident management', 'downtime cost', 'downtime loss', 'interruption cost' and 'operational risk' as well as combinations thereof. In abstract- and, if appropriate, full-text-analyses, we identified peer-reviewed journal and conference papers as well as doctoral theses. We complemented the results obtained through a systematic forward and backward search using the same selection criteria.

**Bartsch (2010): "Modellierung und Simulation von IT-Dienstleistungs-prozessen"**

Bartsch (2010) analyzes the influence of IT service availability on the output of IT service processes. More precisely, he models service processes and services which support specific activities of these as High-level Petri-Nets and runs a number of discrete-event simulation experiments. The author examines different structures of business processes and systematically changes service availability levels, activity processing times, resource capacities, durations of single service outages, and inter-arrival times of jobs in these experiments (Bartsch, 2010, pp.193-237).

[36] Further approaches which are related to the first step of our method for Cost-Optimal Service Selection and which focus on the assessment of the adverse business impact of imperfect service quality are discussed in Section 6.6. Moreover, Section 7.5 refers to articles on service procurement negotiations and auctions, which are associated with the second step of COSS.

In each simulation experiment, the defined service availability level is exactly achieved and all service outage incidents simulated are of equal duration. The time between service outage incidents is variable, however. In order to assess the influence of distinct service availability levels on service process performance, Bartsch (2010) compares the output of a service process devoid of service outage incidents to the output which is achieved when supporting services are disrupted.

In contrast to Bartsch (2010), we elaborate on a deficiency of aggregating service level indicators (such as service availability measures currently used in practice). In order to address this shortcoming, we analyze the impact of single service incidents on business operations and propose service incident patterns as a new form of service quality measures. Furthermore, as opposed to the work at hand, Bartsch (2010) measures the aggregate impact of a number of sequential service outages which are of *equal* duration. We develop a method which provides guidance for customer IT and business departments on how to proceed in order to determine business cost functions. More precisely, we monetarily assess the impact of single service incidents with *different* service incident levels and times of occurrence through the adoption and usage of business process simulation (see Chapter 6).

**Breitgand *et al.* (2007): "Derivation of Response Time Service Level Objectives for Business Services"**

Taking a provider perspective, Breitgand *et al.* (2007) develop a model to derive the cost-optimal combination of response time SLOs for the different components of a specific service delivery environment (one SLO for each service component). The approach takes into account different time periods ("usage windows", Breitgand *et al.*, 2007, p.29) within the total service time agreed and examines all potential combinations of "response time service level objectives" which could be stipulated in a service level agreement.

The overall goal is to maximize service provider's utility, which depends on the number of transactions processed in time ("successful transactions", Breitgand *et al.*, 2007, p.31) within the different time periods. Utility is calculated considering gains from successful transactions and losses due to delayed transactions. The optimization model by Breitgand *et al.* (2007) requires that only a certain percentage of all transactions may breach the response time SLOs defined. This "breach budget" represents another SLO, which is assumed to be fix (Breitgand *et al.*, 2007, p.29).

Whereas the model suggested aims to identify the cost-optimal SLO combination for a given service delivery environment (service provider perspective), our approach supports the selection of the cost-optimal service solution among different service offers (service customer perspective). That is, we consider service quality targets as essential input parameters of our approach.

**Franke (2012): "Optimal IT Service Availability — Shorter Outages, or Fewer?"**

Franke (2012) addresses the derivation of cost-optimal service availability. The paper focusses on the discussion of business costs and their variation over time. In particular, Franke suggests to assess a service's total business costs using expected values and standard deviations of business costs, which are incurred at distinct service incident levels.

The case that a provider and a customer may have different knowledge about business costs or service outage behavior and may refuse to disclose these to one another is not considered. Therefore, information required for optimization is supposed to be available. The article regards service availability and outage incidents but does not examine further service incident classes.

Moreover, Franke (2012) does not elaborate on the significance and effects of service incident distributions with regard to total business costs. Instead, the approach developed assumes all occurring incidents to be of equal duration. We argue that service incident distributions have to be defined for each service incident class in order to allow the determination of total business costs and, thus, the selection of cost-optimal service offers.[37]

**Jakoubi *et al.* (2010): "A Formal Approach Towards Risk-Aware Service Level Analysis and Planning"**

Jakoubi *et al.* (2010) present an approach to assess the influence of imperfect service availability on business operations through process simulation. They model services as resources which support specific business process activities. Furthermore, the authors quantify the financial impact due to disruptions of business operations which result from repeated service outages.

Jakoubi *et al.* (2010) demonstrate the application of their method using the example of a manufacturing process. In order to perform a certain activity of this business process, a supporting service, which is offered by an external provider and specified in an SLA, must be available. Furthermore, the organization regarded has stipulated an SLA with a customer. This SLA defines penalties to be paid in case the organization cannot deliver the agreed quantity of goods.

As opposed to this thesis, Jakoubi *et al.* (2010) consider aggregating service level indicators for service availability only. Similar to Bartsch (2010), they measure the aggregate business impact of a series of service outages which are of equal duration. In contrast, we develop a simulation-based method (see Chapter 6) which provides guidance for customer IT and business departments on how to prepare and conduct simulation experiments in order to determine business cost functions. That is, we quantify the business impact of single service incidents with different service incident levels and times of occurrence.

**Moura *et al.* (2006): "A Quantitative Approach to IT Investment Allocation to Improve Business Results"**

Moura *et al.* (2006) present an approach to rank business processes with regard to investment needs in IT services enabling these. Thus, they support the comparison of

[37] We note that the work by Franke (2012) was published after the first submission of Kieninger *et al.* (2012a). Nevertheless, we were able to refer to this related piece of work in the camera-ready version of Kieninger *et al.* (2012a).

investment opportunities with regard to service quality in case of budget limitations. The authors define performance objectives for business processes and regard deviations from these targets as the adverse impact of service incidents.

Based on historical data Moura *et al.* (2006) estimate a monetary "loss rate" for each point in time, which quantifies the marginal business impact if there is a disruption at that instant. They calculate the overall business costs incurred by a business process through integrating the corresponding loss rates over all disruption time periods. Afterwards, they multiply the resulting overall business costs with the relative weight of a business process (representing the business process's importance compared to other business processes) in order to determine its rank.

In contrast to the work at hand, Moura *et al.* (2006) do not consider the case of non-constant marginal business cost functions. Accordingly, they neither discuss nor address the deficiency of current service quality measures with respect to their appropriateness to assess the business impact of imperfect service quality on business operations.

However, the authors examine the case that loss rates may not be the same in different time periods within a reference period.[38] In this thesis, we capture this aspect through the definition of a separate business cost function for each time period within which a service incident (at a specific service incident level) induces business costs different to those that the same service incident causes in another time period.

**Sauvé *et al.* (2005): "SLA Design from a Business Perspective"** (et al.)

Following a BDIM approach (see Section 3.3.2), a group of authors around Sauvé, Marques and Moura proposes an approach to determine cost-optimal service level objectives for service level agreements in an e-commerce setting (Sauvé *et al.* (2005); Sauvé *et al.* (2006b), Marques *et al.* (2007); Marques *et al.* (2009)).

More precisely, their models aim to minimize the sum of IT infrastructure's "total cost of ownership" and monetarily quantified, adverse business impact it induces ("business loss" due to imperfect availability and performance, e.g., Sauvé *et al.*, 2005, p.74). The approach allows for a detailed description of service components and corresponding service provision costs.

In contrast to our work, it assumes constant marginal business cost functions: When a service outage occurs, the revenue from all e-commerce user transactions which are processed at that time is lost. Furthermore, the revenue from user transactions which exceed a specific service response time (which is constant and the same for all users) is assumed to be lost, too. Our approach, however, could also describe the case that more and more users may defect to a competing e-commerce platform the longer a service outage or performance degradation lasts.[39] Thus, we are able to model the effects of user behavior with respect to business costs more realistically.

[38] The influence of a service incident may, for instance, be higher on a Monday morning than on a Friday evening.

[39] This corresponds to a setting with non-constant marginal business cost functions.

Finally, the authors do not differentiate between private information of provider and customer and consider aggregating service quality measures only.

**Résumé**

The articles reviewed in this section take different perspectives (provider, customer or omniscient analyst point of view) on the selection of optimal service quality objectives and apply different research methods. We note, however, that when determining services' overall business impact, all authors either assume service incidents to be of *equal duration* or suppose marginal business costs to be *constant*. Furthermore, they consider *aggregating* service quality measures only.

Thus, the literature review confirms the novelty and contribution of our approach. In contrast to the work discussed, we suggest service customers to apply non-aggregating service quality measures (i.e., service incident patterns, which reflect services' characteristic combination of service incident levels) and to determine business cost functions in order to be able to select cost-optimal service solutions in case of non-constant marginal business costs.

## 4.6. Excursus: Application of Service Incident Patterns from a Provider Perspective

In this section, we illustrate how a provider might proceed in order to determine characteristic service incident patterns of a service solution. Furthermore, we discuss challenges for service providers, which emerge from the application of service incident patterns as service quality measures.

The set of provider challenges which we present was identified in numerous discussions with subject matter experts from practice and academia. Participants of these sessions have deep knowledge of IT outsourcing and long-standing experience with service delivery.

We recall that a service incident pattern describes a service's characteristic combination of service incident levels with regard to a single service incident class and time period (see Section 4.2).

### 4.6.1. Determination of Service Incident Patterns

Service providers typically have private information about the service incident behavior of service solutions they offer. Since they deliver similar service solutions to different customers, providers can gather data about the characteristic service incidents which occur. That is, they are able to create service incident histories with regard to each service delivery environment used.

Table 4.4 illustrates a potential structure of such a service incident history with regard to a distinct service solution and one specific service incident class. Each data record refers to a certain time period during which the service solution was provided to a particular service customer. A data record describes in schematic form the frequencies with which service incidents at the different service incident levels occurred.

The service incident class regarded has $n$ different attributes.[40] The value range of each attribute is divided into a number of disjoint intervals. The value range of the first attribute is split into $a$ disjoint intervals, the value range of the second attribute is split into $b$ disjoint intervals etc. The terms $l_{1,1,\ldots,1}$ to $l_{a,b,\ldots,n}$ represent all potential service incident levels which a service incident of this class may have (column headings of Table 4.4).[41] The functions $f_1(.)$ to $f_t(.)$ describe the absolute frequencies listed in record number 1 to $t$ for each service incident level (entries of Table 4.4).

Table 4.4.: Service Incident History of a Service Solution regarding a Distinct Service Incident Class

| | absolute frequency of service incidents at service incident level | | | | |
|---|---|---|---|---|---|
| | $l_{1,1,\ldots,1}$ | ... | $l_{1,1,\ldots,n}$ | ... | $l_{a,b,\ldots,n}$ |
| record 1 | $f_1(l_{1,1,\ldots,1})$ | ... | $f_1(l_{1,1,\ldots,n})$ | ... | $f_1(l_{a,b,\ldots,n})$ |
| record 2 | $f_2(l_{1,1,\ldots,1})$ | ... | $f_2(l_{1,1,\ldots,n})$ | ... | $f_2(l_{a,b,\ldots,n})$ |
| ... | ... | ... | ... | ... | ... |
| record t | $f_t(l_{1,1,\ldots,1})$ | ... | $f_t(l_{1,1,\ldots,n})$ | ... | $f_t(l_{a,b,\ldots,n})$ |

*For example, a record could document that there were three reduced response time incidents*[42] *with a duration in the interval* $(0;5)$ *minutes and a latency in the interval* $(0;200)$ *milliseconds (*$f_2(l_{1,1}) = f_2((0;5),(0;200)) = 3$*), five reduced response time incidents with a duration in the interval* $[5;10)$ *minutes and a latency in the interval* $(0;200)$ *milliseconds (*$f_2(l_{2,1}) = f_2([5;10)(0;200)) = 5$*), etc. in a specific time period.*

By analyzing these incident histories, service providers can predict service incident distributions for each service incident class and time period, i.e., determine service incident patterns.[43] In the following, we discuss challenges for service providers, which arise with the application of service incident patterns.

### 4.6.2. Challenges for Service Providers

In order to be able to determine service incident behavior and establish service incident histories, providers may monitor service solutions in an end-to-end manner.

40 Consequently, the service incident level of a single service incident is defined by a tuple of $n$ values — one value regarding each attribute.

41 The term $l_{1,1,\ldots,1}$ represents the particular service incident level at which the value of each attribute lies in the first interval of the corresponding set of disjoint intervals.

42 Reduced response time incidents are described by two service attributes, namely 'reduced response time incident duration' (measured in minutes) and 'reduced response time incident degree' (measured in milliseconds). In this case, the service incident level of a single reduced response time incident is specified by a tuple of two values — one value regarding each attribute.

43 The prediction of service incident patterns may be regarded as a specific form of "service analytics" (e.g., Fromm *et al.*, 2012).

Data collection has to be thorough, reliable and exhaustive in order to allow for accurate predictions. For example, the behavior of an application software (which is provided as a service) can be measured using "robotic means" (Darmawan *et al.*, 2008, p.7), i.e., specific software applications which simulate end-user behavior and record the performance of the service under observation — instead of looking at the behavior of service components only (Darmawan *et al.*, 2008, p.1). Additionally, end-to-end service incident behavior can be calculated or simulated bottom-up (e.g., van Dinther *et al.*, 2010), both based on monitoring data of service components.

Furthermore, providers need to be aware of the particular service components forming a service[44] and standardize service solutions in order to be able to predict service incident patterns. By only applying well-known components of service delivery environments providers can identify similar service solutions and use corresponding service incident histories as basis for the prediction of service incident behavior. The dependencies of a business application (which is provided as a service) on various hardware and software service components, for instance, can be identified using specific "auto-discovery solutions" (Jacob *et al.*, 2008, p.1).

Moreover, service incident behavior of service solutions may be subject to fluctuations. Consequently, providers have to consider variances with respect to frequencies of service incidents at specific service incident levels when defining service offers. Additionally, predictions of service incident behavior require comprehensive service incident histories since estimates which rely on small data sets may be inaccurate. The challenge to predict the incident behavior of servers using incident histories, for instance, is addressed in recent research (e.g., Bogojeska *et al.*, 2013, 2014).

Particularly, providers need to be aware of penalty and bonus payments due to future deviations from stipulated service incident patterns (i.e., service quality targets). Therefore, they will strive to achieve stable service incident behavior. Risk-neutral providers might define service offers based on expected values of service incident frequencies, penalty payments and bonus payments. In contrast, risk averse or risk seeking providers will probably apply specific decision rules to derive favorable combinations of service incident patterns and service delivery costs. In Schmitz *et al.* (2014) we elaborate on the design of service offers in case of uncertain future service incident behavior and service incident-related penalty- and reward-rules. We note that providers might pool such 'service incident behavior'-risks among different services and customers.

In general, service providers have to be able to assure service incident patterns at reasonable (additional) service delivery costs. If they incurred significant extra costs — e.g., due to service incident monitoring, modelling of service architectures and delivery environments as well as prediction efforts — and consequently increased service prices, savings achieved through Service Level Engineering as to total service costs might be lost. Therefore, cost-efficient approaches which support the derivation of service incident patterns are needed.

44 For instance, they could define service-oriented architectures.

Despite the challenges discussed, the application of service incident patterns enables service providers to considerably better understand and meet customers' business requirements. Thus, using service incident patterns, providers may achieve a significant advantage over competitors which are not able to precisely specify service behavior.

## 4.7. Summary

In this chapter we first discussed the deficiency of aggregating, incident-based service quality measures, which are commonly used in practice, with regard to their suitability to describe the business impact a service causes. We argue that IT departments have to consider the impact of single service incidents on business operations in order to be able to assess services' adverse business impact on business operations — which is a prerequisite for the selection of cost-optimal service solutions (see Chapter 3).

Afterwards, we elaborated on the characteristics of service incidents and formally described the relationship between service incidents and the business costs these induce. In order to address the shortcomings of current service quality measures, we suggest an advanced form of incident-based, non-aggregating service quality measures: service incident patterns, which represent services' characteristic combination of service incident attribute values. Furthermore, we recommend service customers to determine business cost functions, which describe the cash-effective adverse business impact a single service incident induces depending on its attribute values and time of occurrence.

Moreover, we introduced our method for Cost-Optimal Service Selection to support IT departments, which is based on the Service Level Engineering approach. COSS consists of two main steps: First, IT and business departments jointly determine business cost functions. Therefore, they measure the influence of single service incidents with regard to business consequences using a business process simulation-based method (see Chapter 6). Second, the IT department applies a cost-optimal mechanism which supports the negotiation of service offers with multiple providers (see Chapter 7). This allows the IT department to select the cost-optimal service solution among different provider offers.

Finally, we referred to related work and we discussed challenges for service providers emerging from the application of service incident patterns as service quality measures. In the following chapter we provide an overview on classifications of business consequences and corresponding costs, which may serve IT and business departments as a means to identify and classify cash-effective business consequences in practice.

# 5. Identification of Cash-Effective Business Consequences

In the previous chapters we argued that the assessment of services' adverse business impact on business operations is a prerequisite for the selection of cost-optimal service solutions. IT departments have to determine business cost functions in order to be able to weigh service quality against service price. Following a BIA approach, the identification of business consequences, which are of "diverse nature" (Salmela, 2007, p.3), is a first step towards this goal (see Section 4.3.2).

Business consequences, which are induced by service incidents, are context-specific and, therefore, have to be determined specifically for each business setting (Dale & Plunkett, 1995, p.47). Some types of business consequences can easily be assessed in terms of costs. Typical examples of such business consequences are 'defective materials, components or products', 'penalties and compensation payments' as well as 'overtime work' (see Table 5.1 for references). But business consequences may also be more difficult to quantify. 'Lost sales', 'dissatisfaction of stakeholders' and 'breach of law' represent rather abstract examples of business consequences. Accordingly, ITIL differentiates between resulting "tangible costs" and "intangible costs" noting that elements of the latter class "may be difficult to measure" (Taylor *et al.*, 2007a, p.106).

This chapter first provides an overview on comprehensive classifications of business consequences and corresponding costs. We identified these classifications via an extensive literature review across different fields of research. Afterwards, we explain why only cash-effective business consequences — which either constitute decreases in cash receipts (cash inflows) or increases in cash payments (cash outflows) — should be taken into account regarding the selection of service solutions.

Table 5.1.: Examples of Business Consequences

| business consequence | exemplary references |
|---|---|
| defective materials, components or products | Taylor *et al.* (2007a, p.105); Hiles (2011, p.190); Sullivan *et al.* (1996); Yang (2008) |
| penalties and compensation payments | Balaouras *et al.* (2008); Gryna (1999, p.8.6); Levitt (1997, p.97); von Rössing (2005, p.94) |
| overtime work | Fox *et al.* (2008); Hiles (2011, p.191); Salmela (2007); Yang (2008) |
| lost sales | Balaouras *et al.* (2008); Gryna (1999, p.8.3); Levitt (1997, p.97); von Rössing (2005, p.91) |
| dissatisfaction of stakeholders | Akkiraju *et al.* (2012); Atkinson *et al.* (1994, p.369); Fox *et al.* (2008); von Rössing (2005, pp.94-95) |
| breach of law | Akkiraju *et al.* (2012); Balaouras *et al.* (2008); von Rössing (2005, p.94); Salmela (2007) |

## 5.1. Classifications of Business Consequences and Corresponding Costs

Classifications of business consequences and corresponding costs may serve IT departments as a means to identify and classify potential adverse effects of service incidents when analyzing a specific business setting. First, categories of business consequences and corresponding cost types described in a classification provide guidance with regard to a systematic determination of adverse effects in practice (e.g., in interviews with experts from business departments and other stakeholders). Second, classifications usually cite examples of business consequences and corresponding costs. These examples foster the detection of relevant business consequences in a concrete business setting as well.

This section provides an overview on classifications of business consequences and corresponding costs from different areas of research: Energy Economics and Production Economics, Business Continuity Management and Maintenance Management, IT Service Management as well as Quality Management. Thus, we bring together "disparate streams of work" as suggested by Webster & Watson (2002, p.xv).

We identified these articles via an extensive literature review, following the approaches suggested by vom Brocke *et al.* (2009) and Webster & Watson (2002). We used Google Scholar as primary source and conducted a keyword-based full-text search without temporal restrictions. Search terms applied were 'business consequence', 'business impact analysis', 'quality cost', 'cost of poor quality', 'business loss', 'business cost', 'financial loss', 'financial impact', 'interruption', 'outage', 'disruption', 'incident', as well as combinations thereof. In abstract- and, if appropriate, full-text-analyses, we identified book chapters, peer-reviewed journal and conference papers as well as doctoral theses. We complemented the results obtained through a systematic forward and backward search using the same selection criteria.

In the following, we refer to a selection of articles which provide the most comprehensive collections of business consequences and corresponding costs in their particular area of research.[1] We note, however, that most of these works do not clearly differentiate between business consequences and associated costs.

### 5.1.1. Energy Economics and Production Economics Literature

**Energy Economics** is an area of research which is "concerned with the economics and econometric modelling and analysis of energy systems and issues" (Energy Economics, 2014, p.1). It examines, inter alia, the "economically efficient use" of energy and the motivation of companies and customers to "supply, convert, transport, [and] use" it (Sickles, 2008, p.842).

**Sullivan *et al.* (1996): "Power Interruption Costs to Industrial and Commercial Consumers of Electricity"**

Sullivan *et al.* (1996) propose a method to determine costs resulting from power interruptions in order to decide on "design and operation alternatives" (Sullivan *et al.*, 1996, p.23) with regard to power supply systems. First, they collect information about a specific energy consumer company in an on-site survey. Afterwards, they use this knowledge to estimate power interruption costs using regression models, which are based on the results of an extensive survey of 210 large commercial and industrial energy consumer companies.

In this comprehensive survey Sullivan *et al.* (1996) identified different categories of costs, each element of which refers to a specific business consequence. In general, they regard power interruption costs to comprise the "value of lost production" and "outage-related costs" less "outage-related savings" (Sullivan *et al.*, 1996, p.25). The authors assume the value of lost production to correspond approximately to the decrease in revenue due to goods or services not being produced or provided (business consequences). Examples of outage-related costs mentioned are "[raw] material damage costs", "equipment damage costs" and "labor costs to restart production" (Sullivan *et al.*, 1996, p.25 & p.28). Outage-related "savings" cited comprise, inter alia, material costs which are not incurred and the "value of scrap" (Sullivan *et al.*, 1996, p.26). In their article, Sullivan *et al.* (1996, p.25) consider different outage durations, degrees of "voltage sags" and times of incident occurrence.

**Production Economics** is a field of research in which "engineering and technology meet the managerial and economic environment" an industry operates in (International Journal of Production Economics, 2014, p.1). It aims to maximize companies' profit given their specific technological opportunities as well as the "input and output prices" these face (Rasmussen, 2013, p.4).

[1] Please see above for the list of research areas considered.

**Fox *et al.* (2008): "Determination of the Financial Impact of Machine Downtime on the Post Large Letters Sorting Process"**

The authors present a classification of costs which may result from machine downtime. They argue that the "costing [machine] downtime" had not been addressed well in literature (Fox *et al.*, 2008, p.738). Therefore, they extracted different types of costs which are incurred due to machine outages from existing generic classifications of costs — using the example of the "automated mail processing equipment" and the "large letters sorting process" at Australia Post (Fox *et al.*, 2008, pp.733-734).

The authors aim to find out how machine downtime affects costs in order to quantify the value of proactive maintenance activities. They distinguish the following categories of costs: costs depending on the duration of single outage events, costs which are contingent on the total machine downtime over a production shift and fixed costs which are induced by each outage.

Examples for the first category of costs are "labour costs" paid for operators who "remain idle" and "equipment hire costs" for renting machines equivalent to those which have broken down (Fox *et al.*, 2008, pp.734-735). The list of cost categories "affected by overall downtime" includes, inter alia, specific types of "lost opportunity costs" such as "lost profit costs", "lost demand costs" and "lost capacity costs" (Fox *et al.*, 2008, pp.735-736). Finally, the category "per event downtime costs" (Fox *et al.*, 2008, p.737) comprises, for instance, the procurement of spares .

### 5.1.2. Business Continuity Management and Maintenance Management Literature

**Business Continuity Management** ensures that an organization is capable of continuing the "delivery of products or services at acceptable predefined levels following [a] disruptive incident" (International Organization for Standardization, 2012b, 2.1.10). An important step in the implementation of effective Business Continuity Management is to understand which activities and resources are critical for the delivery of products and services (e.g., Cornish, 2011, p.125). Usually a Business Impact Analysis is applied in order to examine "business functions" with regard to the effects that a "business disruption might have upon them" (Business Continuity Institute, 2011, p.12).

**Hiles (2011): "Business Impact Analysis — Building a Better Mousetrap"**

Hiles (2011, p.201) argues that traditional approaches for Business Impact Analysis mostly oversimplify the calculation of "financial losses" and considerably underestimate these. Therefore, he suggests a "more realistic" BIA approach which considers, inter alia, different potential levels of incident impact (e.g., most likely, pessimistic and optimistic impact level) and their probabilities.

Furthermore, Hiles (2011, pp.202-205) provides a detailed classification of a wide range of cost types in order to support the calculation of downtime costs. The main categories of this classification are "equipment cost", "production downtime cost", "labour overhead cost", "ICT failure: system restoration cost" as well as "sales &

marketing costs". Additionally, he discusses the cost categories "miscellaneous costs", "knock-on costs" (resulting from cascading effects of service incidents) and "societal costs" (to be borne by society (Duden Online, 2014), such as noise and pollution) as well as "benefits of disruption".

Equipment cost include, for example, costs due to lost "employee productivity", whereas production downtime cost comprise "cost of scrap" and "cost of temporary fix or workaround", for instance. Examples for the category labour overhead cost are "maintenance cost" and costs induced by "downtime management". "Cost of labour for complete ICT restoration" and "cost of working from standby site(s)" pertain to the category of 'ICT Failure: System Restoration Cost'. Finally, sales and marketing costs comprise, for example, costs incurred due to "loss of referrals", "cost of recovering market share & brand value" and "cost of crisis communication".

Hiles (2011) emphasizes that double counting of costs in different categories of a classification is to be avoided and that insurance coverage, which reduces downtime costs, should be considered (Hiles, 2011, p.205). Furthermore, he suggests to identify benefits which might result from disruptions (e.g., benefits of "reorganization, rationalisation, relocation and reduced inventory") and which lower the "true costs of downtime" as well.

**von Rössing (2005): "Betriebliches Kontinuitätsmanagement"**

According to von Rössing (2005, p.86) a business continuity strategy cannot be defined without precise knowledge of the financial impact an incident causes. Therefore, he suggests to monetarily quantify the business impact of potential incidents in a structured way by considering different categories of business consequences (von Rössing, 2005, pp.86-95).

First, von Rössing (2005) discusses 'damage to resources' (e.g., machines) and 'influence on business process performance' (e.g., decrease in output). Second, the categories 'revenue loss' and 'reduction of profit' are introduced. Third, he argues how 'loss of market share' could be valued. Afterwards, von Rössing (2005) elaborates on 'consequences regarding the life cycle of products'. Outages of a software distribution platform may, for instance, result in a delayed launch of a software product and cause futile or additional marketing expenses. Finally, the categories 'legal consequences' and 'loss of reputation' are explained in detail.

**Maintenance Management** aims at the determination of "maintenance objectives, strategies and responsibilities" as well as their implementation (European Committee for Standardization, 2010, p.6). In this context, 'maintenance' is regarded as the sum of all activities needed to "keep a system and all of its components in working order" (Stephens, 2004, p.3), which also include repairs.

**Levitt (1997): "The Handbook of Maintenance Management"**

Maintenance management comprises the controlling of costs which are incurred "during the life cycle of an item" (European Committee for Standardization, 2010, p.6). Levitt (1997, pp.98-103) states and discusses five categories of such "life cycle costs", namely "ownership costs", "operating costs", "maintenance costs", "overhead costs" and "downtime costs".[2]

With regard to the identification of costs resulting from business consequences, the latter of these "core cost areas" (i.e., downtime costs) is of particular interest. Levitt (1997, pp.102-103) lists and explains six categories of downtime costs, examples of which are "revenue loss less recoverable costs like materials", "idle operator salary" and "tangible and intangible costs of customer dissatisfaction".

### 5.1.3. IT Service Management Literature

**IT Service Management** aims to "direct and control [...] activities and resources for the design, transition, delivery and improvement of services to fulfil [...] service requirements" (International Organization for Standardization, 2011, 3.30). In the context of IT Service Management, 'IT Service Continuity Management' identifies, assesses and controls risks which might affect IT services and, thus, supports Business Continuity Management (Taylor *et al.*, 2007a, p.301).

**Akkiraju *et al.* (2011) & Akkiraju *et al.* (2012): "Towards Effective Business Process Availability Management"**

Akkiraju *et al.* (2011) and Akkiraju *et al.* (2012) present an empirical approach to determine costs which are induced by planned maintenance measures causing business process unavailability. They use the information obtained in order to schedule maintenance activities and minimize the resulting expected total business impact. Through interviews the authors identified four main categories of business consequences resulting from outages. These are "financial losses", "business interruptions", "legal & regulatory impact" and "image & credibility impact" (Akkiraju *et al.*, 2012, pp.327-328).

As examples for the first category they mention, inter alia, "lost business opportunities", "lost discounts" and "late fees" Akkiraju *et al.* (2012, p.327). The second category of business consequences, business interruptions, captures backlogs of activities, such as disruptions of "loan processing" and "inter-company transfers" (ibid.). The third category comprises liabilities, such as "legal exposure" and "health and safety litigation exposure" (ibid.). Examples for the last category, which covers stakeholder reactions to outages, are impact on "brand reputation" and stock price (Akkiraju *et al.*, 2012, p.328).

**Balaouras *et al.* (2008): "Building the Business Case for Disaster Recovery Spending"**

Balaouras *et al.* (2008) argue that the quantification of costs induced by downtime and data loss is essential in order to build a business case for spending on disaster recovery and business continuity. They state that IT departments and business departments have to work together to identify risks (i.e., "not only disasters" (Balaouras *et al.*, 2008, p.2)), stipulate recovery objectives and select appropriate technologies and services.

[2] An overview on the field of life cycle costing is, for instance, provided by Dhillon (2010).

Therefore, the authors suggest to consider different categories of business consequences when calculating the financial impact of downtime and data loss. Examples for these categories are "impact to cash flow" and "penalties and loss of discounts" (Balaouras *et al.*, 2008, p.5).[3]

As examples of business consequences which fall into the category "impact to cash flow" the authors mention that a business department might be unable to "recognize revenue" or "process accounts receivable". The inability of a company to make use of early payment discounts is an example for the category "loss of discounts".

### 5.1.4. Quality Management Literature

**Quality Management** is responsible for the definition of quality policies and targets as well as for their implementation through activities such as "quality planning, quality control, quality assurance and quality improvement" (International Organization for Standardization, 2005, 3.3.4).

The measurement of 'quality costs' fosters the identification of quality problems and of opportunities for cost-reduction[4] and quality improvement (Dale & Plunkett, 1995, p.38). Furthermore, it supports managers in justifying the need for quality improvement initiatives and in monitoring their success (Dale & Plunkett, 1995, p.14). Quality costs are usually regarded as the costs which are incurred to avoid quality problems, the costs induced by evaluating the achieved quality level, and the costs resulting from imperfect quality (Yang, 2008).[5]

Accordingly, literature frequently distinguishes four categories of quality costs, namely "prevention cost", "appraisal cost", "internal failure cost" and "external failure cost" (British Standard Institution, 1990, p.1). Prevention costs result from activities aiming to examine, avoid or mitigate "the risk of nonconformity or defect" (British Standard Institution, 1990, p.2). The category 'appraisal cost' comprises costs incurred to evaluate if quality objectives are achieved (ibid.). Costs resulting from violation of quality objectives or defects which arise within a company belong to the category 'internal failure cost' (ibid.). Finally, external failure costs are the costs a company incurs due to customers experiencing violation of quality objectives or defects (ibid.).

Business costs monetarily quantify the impact of imperfect service quality on business operations but do neither comprise costs to prevent service incidents nor costs to evaluate the achievement of service level objectives. Therefore, internal and external failure costs are of particular interest with regard to the determination of business costs.

[3] The remaining categories are similar to types of business consequences that have been discussed in this chapter before. These are "revenue losses", "productivity losses", "compliance and/or reporting penalties", "impact to customers and strategic partners", "employee morale and employee confidence in IT" and "damages to reputation and goodwill" (Balaouras *et al.*, 2008, p.5).

[4] In this case, "cost-reduction" refers to the costs incurred due to imperfect quality.

[5] Yang (2008) provides a discussion about definitions of the terms 'quality costs' and 'cost of poor quality', which are basically used synonymously.

### Atkinson *et al.* (1994): "Standard Cost-of-Poor-Quality Element Listing"

Atkinson *et al.* (1994, pp.343-392) provide a comprehensive classification of quality costs, which is based on Campanella (1999, pp.187-204). They list and explain subcategories of internal and external failure costs, such as "operations failure costs" and "staff failure costs" (internal) as well as "liability costs", "penalty costs" and "lost sales" (external).

Quality cost elements from these subcategories might be induced by service incidents in specific business settings and, thus, correspond to types of business costs.

### Gryna (1999): "Quality and Costs"

Similarly, Gryna (1999, pp.8.4-8.8) differentiate between internal failure costs, external failure costs, appraisal costs and prevention costs and provide numerous examples of quality costs. They further subdivide the category internal failure costs into "failure to meet customer requirements and needs" (e.g., "scrap and rework" and "lost or missing information") and "cost of inefficient processes" (e.g., "variability of product characteristics" and "unplanned downtime of equipment"). Furthermore, with regard to external failure costs, they distinguish between "failure to meet customer requirements and needs" (e.g., "returned material" and "allowances") and "lost opportunities for sales revenue" (e.g., "new customers lost because of quality").

### Yang (2008): "Improving the Definition and Quantification of Quality Costs"

Yang (2008, p.175) suggests to consider "hidden costs" as a further main category of quality costs which complements the four "traditional" categories mentioned above. According to Yang (2008), this additional category comprises failure costs which are difficult to determine because they are either not recorded appropriately or not discovered at all. Moreover, the author provides an extensive classification of quality costs that is structured by categories of quality costs as well as by phases of a product life cycle (in each phase of which specific types of quality costs may be incurred).

### Résumé

Our literature review reveals that there are different ways to classify business consequences and corresponding costs. We find that the field of quality management offers the most comprehensive classifications. Furthermore, we state that in IT Service Management literature authors rather concentrate on causes and types of incidents.

Many classifications address a specific business context or focus on consequences which are regarded to be of particular importance. Therefore, we recommend analysts to not only consult classifications from their own field of application when determining business consequences. In other words, we suggest them to also consider classifications from other areas — particularly from the field of quality management — in order to take a different perspective and complement their results.

In this work, we assess business consequences in terms of business costs in order to be able to select cost-optimal service solutions. The following section argues why only cash-effective business consequences should be taken into consideration for this purpose.

## 5.2. Consideration of Cash-Effective Business Consequences in the Selection of Service Solutions

The selection of a service solution to support business operations from among different service offers — in consideration of the total service costs these entail — represents an investment decision. Baumol & Blinder (2012, p.377), for instance, define an investment as the "flow of resources into the production of new capital." The Merriam-Webster Online Dictionary (2014) describes it as a "process of exchanging income for an asset that is expected to produce earnings at a later time."

In practice, such investment opportunities are usually assessed using capital budgeting techniques[6] (cf., Baker & English, 2011). Therefore, we recommend IT departments to use capital budgeting approaches, such as the net present value decision rule (e.g., Berk & DeMarzo, 2011, pp.59), to evaluate provider offers.[7] In particular, we suggest the application of dynamic capital budgeting methods, which are more precise than static ones and allow the comparison of multi-period offers (Putnoki *et al.*, 2011, p.21). In general, these approaches operate on information about future cash flows (ibid.).[8]

Consequently, IT departments should concentrate on cash-effective business consequences when making service investment decisions. In other words, they only need to consider business consequences which either constitute decreases in cash receipts (cash inflows) or increases in cash payments (cash outflows). It is for this reason that we define 'business costs' to be cash-effective in this work.

Whether or not a certain business consequence is cash-effective has to be determined specifically for each business setting, however. Overtime work, for instance, does not necessarily imply additional cash outflows, unless employment contracts stipulate that overtime is paid. Similarly, productivity loss does not reduce cash inflows in any case, since demand might be satisfied through supply from stocks.

Analysts also need to define to what extent cascading effects of a specific business consequence have to be considered when deciding on the cash-effectiveness of this particular consequence. That is, they must specify, for instance, whether a productivity loss which induces supply from stock which, in turn, causes additional freight charges is considered to be cash-effective itself. Furthermore, they need to avoid double counting of business costs and to consider insurance coverage, which might induce additional cash flows (e.g., Hiles, 2011, p.205).

6 According to Fabozzi & Peterson-Drake (2010, p.546), capital budgeting is the "process of identifying and selecting investments in long-lived assets; that is, selecting assets expected to produce benefits over more than one year".

7 Other examples of dynamic capital budgeting approaches are the 'equivalent annual cost method' and the 'internal rate of return method'.

8 The International Accounting Standard 7 (IAS 7) "Statement of Cash Flows" (e.g., Zülch, 2012, pp.92) distinguishes three main categories of cash flows, namely "cash flows concerning operating activities", "cash flows concerning investing activities" and "cash flows concerning financing activities".

Despite these challenges, the determination of cash-effective business consequences is essential to arrive at a sound investment decision and select the cost-optimal service solution.

## 5.3. Summary

The identification of business consequences is a first step towards the determination of business costs. Therefore, in this chapter we initially provided an overview on classifications of business consequences and corresponding costs from different fields of research, which we identified via an extensive literature review.

These classifications may serve IT departments as a means to identify and categorize business consequences when analyzing a specific business setting. They provide guidance with regard to a systematic determination of business consequences in practice and mostly cite concrete examples of business consequences and corresponding costs.

Afterwards, we pointed out that the selection of a service solution represents an investment decision. We found that such investment opportunities are usually assessed through the application of dynamic capital budgeting approaches in practice, which generally operate on information about future cash flows.

Accordingly, we suggest IT departments to only consider cash-effective business consequences when making service investment decisions. That is, we recommend to focus on business consequences which either constitute decreases in cash receipts or increases in cash payments. It is for this reason that we define 'business costs' as cash-effective in this work.

In the following chapter, we present our simulation-based method for the determination of business cost functions, which also refers to the classifications discussed above.

# 6. Simulation-Based Determination of Business Cost Functions

IT departments have to understand the influence of service incidents on business operations in order to be able to weigh service quality against service price and to select cost-optimal service solutions (see Section 4.3).[1] Therefore, they need to work closely with business departments and support these in determining business cost functions.

In this chapter we present an approach which addresses this challenge and, thus, implements the first step of our method for Cost-Optimal Service Selection (see Section 4.4). It supports the assessment of the business costs induced by single service incidents at a certain time in case of well-defined business processes.

Our simulation-based method links the occurrence of service incidents at specific service incident levels (e.g., of an outage incident with a certain duration) and times of occurrence to changes in the values of process-related business parameters. That is, we provide an approach for IT and business departments to jointly determine the impact of service incidents on business operations through the adoption and usage of an established, formal technique for business process analysis. Using discrete-event simulation we are able to consider stochastic as well as dynamic aspects of business processes.[2]

As described in Section 4.4 our method builds on the BIA process by Wiedemann (2008, p.122).[3] Our approach consists of four steps (see Figure 6.1): First, a team of analysts from IT and business departments jointly identifies the cash-effective business consequences which are potentially induced by a disruption of business processes in

[1] The service price may be regarded as costs from a service customer point of view.
[2] Section 2.2.1 provides an overview on the field of simulation.
[3] We briefly outline the BIA process by Wiedemann (2008, p.122) in Section 4.3.2.

the business setting under consideration (step 1). Then, they define a measurement model (see Section 4.3) for each relevant business consequence identified (step 2). That is, they link business consequences to process-related business parameters (business process metrics). Afterwards, analysts simulate the impact of single service incidents on the values of those business process metrics (step 3).[4] Finally, based thereon, they estimate the business cost functions of the different service incident classes at different time periods (step 4).[5]

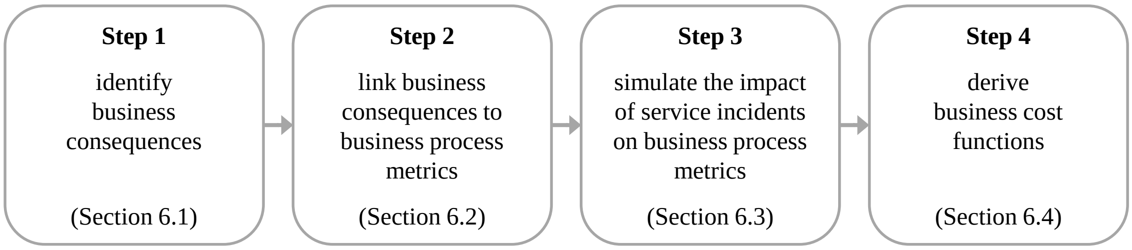


Figure 6.1.: Simulation-Based Method for the Determination of Business Cost Functions

In the following sections we elaborate on the four steps of our simulation-based method for the determination of business cost functions. In Chapter 8, we demonstrate its applicability using the example of an order-picking process (customer business process), which is supported by a warehouse management system (IT service).[6]

## 6.1. Identification of Cash-Effective Business Consequences

First, analysts from IT and business departments jointly determine business processes which are (directly as well as indirectly) supported by the IT service to be procured.[7] Then, they conduct expert interviews and analyze documents in order to identify effects which disruptions of these business processes may cause. That is, they detect cash-effective business consequences (see Section 5.2) which might be incurred due to disturbances of business operations in the particular business setting regarded.

4 In other words, they determine distinct values of business parameter functions (see Section 4.3.2) through business process simulation.

5 The steps 1 and 2 of our simulation-based method correspond to the steps 'identify dimensions of impact' and 'develop measurement models' of the BIA process by Wiedemann (2008, p.122). While step 3 of the BIA process suggests to assess the business impact of IT service outages through expert interviews and document analysis *alone* (qualitative approach), we use business process simulation to determine business cost functions (quantitative approach). Thus, our simulation-based method complements the BIA process. We also briefly refer to the work by Wiedemann (2008) in Section 6.6.

6 A former version of this chapter was published in the proceedings of the Fourth International Conference on Exploring Services Science (see Kieninger *et al.*, 2013b).

7 That is, "hierarchical relations" between business processes (e.g., Dumas *et al.*, 2013, p.37) have to be considered as well. Consequently, analysts have to decide which business processes have to be covered by the study.

The classifications discussed in the previous chapter may serve analysts as a guide to determine business consequences. They suggest numerous effects of business process disruptions which might be of relevance and, thus, facilitate their identification. Analysts could, for instance, use these classifications to structure expert interviews. As a result of this initial step, analysts obtain a list of cash-effective business consequences for each business process which depends on the IT service under consideration.

## 6.2. Mapping of Business Consequences to Business Parameters

Second, the team of analysts defines a measurement model (see Section 4.3) for each business consequence identified in the previous step. For this purpose, they determine (or define[8]) business parameters the values of which can be used to assess the financial impact of a business consequence. In other words, they link each business consequence to one or several business parameters in order to allow the calculation of the business costs it induces.

Furthermore, in a measurement model analysts specify how to assess the values of business parameters referred to. The following example illustrates such a method of measurement:

*Let us assume that the disruption of a production process results in penalty payments to a customer (business consequence). The corresponding measurement model states that the financial impact is calculated by multiplying the values of two business parameters, namely the 'number of products delivered late' (due to the disruption) and a 'penalty fee', which is defined per item. The method of measurement with regard to the first-mentioned business parameter states that the penalty fee is to be paid, if a product is delivered more than thirty minutes after the time assured.*

The selection and definition of business parameters hinges on the business consequences identified as well as on the business setting regarded. In the following, we concentrate on business parameters the values of which are influenced by disruptions of business processes. We denote such process-related business parameters as 'business process metrics' (e.g., the business parameter 'number of products delivered late' in the example above). We assume IT and business departments to determine the values of business parameters which are not dependent on business process results (e.g., the value of the business parameter 'product price' or the 'penalty fee' in the example above) through expert interviews and document analysis.

The result of this step is a set of measurement models each referring to a specific business consequence.

8 If adequate business parameters (e.g., specific business process metrics) have not been used to measure the business process so far, they need to be specified for this purpose.

## 6.3. Assessment of Service Incidents' Impact on Business Process Metrics

Third, the team of analysts determines distinct values of the business parameter functions (see Section 4.3.2) of the business process metrics determined (or defined) in the previous step. In order to assess the influence of service incidents on the values of these business process metrics we suggest to apply discrete-event business process simulation.

In the following, we elaborate on the development of a simulation model and the implementation of simulation experiments. The corresponding sub-steps of our method build on the general "framework for business process simulation" by Paul *et al.* (1998). The framework describes eight phases: (1) define the goal of the simulation experiments, (2) specify the business processes to be considered, (3) collect and analyze data on these business processes, (4) develop a business process simulation model, (5) test the business process simulation model, (6) conduct simulation experiments, (7) analyze simulation results and (8) recommend business process changes.

These tasks will usually not be carried out sequentially but in an iterative manner in order to achieve suitable results (Paul *et al.*, 1998). Similar phases of simulation studies are, for example, described by Kelton *et al.* (2004, pp.529-545) and Birta & Arbez (2013, pp.19-51).

In the context of our simulation experiments, analysts aim to determine the impact of single service incidents on the values of business process metrics. Therefore, they consider only those business processes in simulation experiments which are measured by the business process metrics determined (or defined).[9] Consequently, we do not need to address the first two phases of the framework by Paul *et al.* (1998) any more. Furthermore, we will not consider its last phase, since we do not aim to redesign business processes on the basis of simulation results.

Hence, we suggest analysts to address phases three to seven of the framework by Paul *et al.* (1998) in order to determine the business impact of service incidents on business process metrics. We relabel these phases as sub-steps 3.A to 3.E of our simulation based method (see Figure 6.2).

The first three sub-steps (3.A to 3.C) address the development of a suitable business process simulation model, while the remaining two sub-steps (3.D and 3.E) aim at the implementation of simulation experiments and the analysis of simulation outcomes. In the following paragraphs, we discuss each of these sub-steps in detail.

#### Step 3.A: Collect and Analyze Data on the Business Processes

First, analysts collect information about the "structural and behavioural features" (Birta & Arbez, 2013, p.38) of the business processes to be considered. That is, they

[9] That is, step 2 of our simulation-based method (see Section 6.2) already determines the business processes to be simulated.

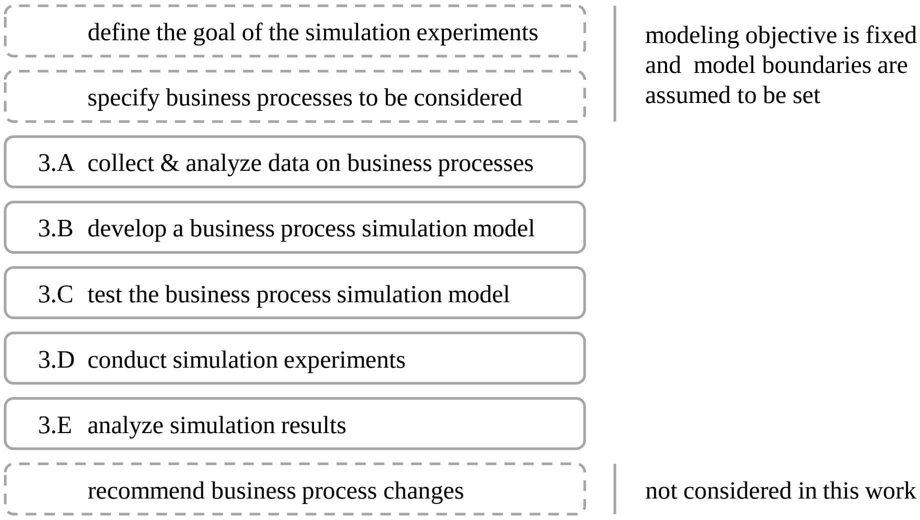


Figure 6.2.: Steps of Business Process Simulation (based on Paul *et al.*, 1998)

develop a formal "conceptual model" (ibid.), which may, according to Rosenkranz (2006, pp.97-116)[10], comprise information with regard to:

- activities and resources (e.g., dynamic behavior (such as changing processing times), capacity constraints or services used)
- structures of business processes (e.g., interdependencies between activities)
- events (e.g., service incidents or inter-arrival times of process instances) and
- routing (e.g., handling of process instances in case of incidents or high utilization)

Analysts should particularly investigate business process structure and behavior in case of disruptions, since they intend to measure effects of service incidents on business operations. This includes the identification of sub-processes which are specifically applied to circumvent problems ('contingency plans'[11]). Furthermore, analysts have to determine the way in which activities react to disruptions of the particular service under consideration. A process instance which is processed by an activity at the time a service incident occurs may (Gaver, 1962, pp.76)

- be resumed "from the point at which interruption took place" when the service incident is resolved ("preemptive-resume"),
- be completely processed without being impacted[12] ("preemptive-postpone") or

[10] A similar list of "basic ingredients" of business processes is presented by Dumas *et al.* (2013, pp.3-6 and pp.236-238).

[11] In contrast to 'workarounds' (e.g., Taylor *et al.*, 2007b, p.64), contingency plans describe ways of overcoming difficulties which are specified *before* those difficulties arise (ex ante).

[12] That is, the service incident only affects process instances the processing of which has not started yet.

- have to be completely reprocessed by this activity and potentially further upstream activities (“preemptive-repeat”).

Different methods to discover information about business processes (i.e., “evidence-based discovery”,[13] “interview-based discovery” and “workshop-based discovery”) are, for instance, discussed in Dumas *et al.* (2013, pp.161-165).

Finally, the group of analysts analyzes information about behavioral features of business processes gathered (e.g., processing times of a distinct activity and inter-arrival times of jobs). They apply statistical methods on sampled data in order to determine distribution functions which appropriately describe behavior. Analysts could, for instance, record a number of sample values of an activity’s processing time and, based thereon, assume this processing time to follow a distribution from a specific family of distributions (e.g., exponential, normal or Poisson distributions). Afterwards, they might apply maximum likelihood estimation in order to determine the distribution parameter value for which the sample is most likely. Finally, they could assess the distribution assumption’s goodness of fit through the application of a statistical test (e.g., a Chi-Square Test or a Kolmogorov-Smirnov Test).

An overview on (statistical) methods to determine behavioral features is provided, for instance, by Banks *et al.* (2010, pp.353-405) and Law (2011, pp.275-388).

### Step 3.B: Develop a Business Process Simulation Model

Based on the data collected in the previous step, analysts now develop a simulation model of the business processes using simulation software. They initially define a simple (high-level) model of the whole system (i.e., the business processes) and then expand and refine it until the required level of detail is reached.

In general, systems or parts of them can be formally described in multiple ways (Kelton *et al.*, 2004, p.288). Accordingly, there may be different approaches to model the occurrence of a service incident. The way in which a service incident is captured in a model depends, inter alia, on its particular service incident class and, thus, on the way it influences business process activities.

*For instance, the effects of an outage incident and a response time incident of the same IT service on a specific activity might have to be modeled differently. An outage incident could be modeled as failure of a resource (using the concept ‘resource’ provided by the simulation software) which represents the corresponding IT service. In contrast, a response time incident could be modeled as an increase in the processing time of an activity which stands for the IT service affected in a process model.*

Furthermore, analysts have to make sure that business process metrics which have been identified (or defined) in step 2 (see Section 6.2) are captured in the simulation model as well. The values of these business process measures have to be recorded when simulation experiments are conducted (i.e., they represent the simulation outcome).

[13] Evidence-based discovery, in turn, comprises “document analysis, observation, and automatic process discovery” (Dumas *et al.*, 2013, pp.161-162).

#### Step 3.C: Test the Simulation Model

Before the model is used in simulation experiments, analysts thoroughly test it "using as many model verification and validation techniques as feasible" (Hlupic & Robinson, 1998, p.1366) while ensuring economic efficiency. Thus, they ensure the credibility of simulation results (Birta & Arbez, 2013, p.41) which will be produced in the following step.

In order to find out if the model behaves as intended (verification, Kelton *et al.*, 2004, p.540), analysts set up test scenarios with predictable results (e.g., by using historical input data (Banks *et al.*, 2010, p.426)). Furthermore, they compare simulation results (e.g., values of business process metrics) to the outcome of the "real system" (validation, Kelton *et al.*, 2004, p.541).

Finally, analysts share the simulation model with individuals who are experts for the business processes it comprises.[14] These specialists might suggest corrections or further refinements. An overview on methods for model testing is provided, for instance, in Banks *et al.* (2010, pp.406-434) and Law (2011, pp.243-274).

#### Step 3.D: Conduct Simulation Experiments

Now, analysts first run the simulation model of business processes in absence of any service incident and, thus, determine the 'regular' values of business process metrics (i.e., their values under normal circumstances). In the following, we refer to this simulation scenario as the 'base case'.

Then, they simulate business operations in presence of a single service incident and record the corresponding values of business process metrics. They vary service incident class and service incident level as well as time of occurrence in different simulation experiments. In other words, analysts define a separate 'service incident scenario' — i.e., a simulation scenario which regards the occurrence of a single service incident — for each combination of service incident class, service incident level and time of occurrence to be considered. Therefore, for each service incident scenario, they create a specific configuration of the general simulation model, which has been developed and tested in the previous sub-steps of our method.[15]

Afterwards, analysts compare the resulting values of business process metrics of the different simulation experiments to those of the base case. Differences in the values of business process metrics are ascribed to the occurrence of the particular service incident.

In the following, we denote the particular set of simulation runs concerning a specific simulation scenario as a 'simulation experiment'. In the next section, we first describe issues with regard to the design of simulation experiments, which analysts should be

[14] Ideally, analysts share an animation of the simulation model with experts for the corresponding business processes (Kelton *et al.*, 2004, p.541). An animation illustrates the behavior of a simulation model and, thus, facilitates a comparison with the behavior of the real-world business process.

[15] In order to handle randomness with regard to the values of business process metrics obtained through simulation, each model configuration is run multiple times.

aware of. In particular, we elaborate on the appropriate definition of the simulation time frame and the number of simulation replications to conduct given a specific simulation scenario. Afterwards, we discuss the selection of meaningful simulation scenarios with respect to the determination of business cost functions.

***Sub-step 3.D.1: Set the Simulation Time Frame***

In general, there are two types of simulation studies, "terminating simulations"[16] and "steady-state simulations" (e.g., Kelton *et al.*, 2004, p.258; Birta & Arbez, 2013, p.72). Terminating simulations feature well-defined "starting and stopping conditions" (e.g., business hours), which are usually determined by the business setting modeled (Kelton *et al.*, 2004, p.258). In contrast, in case of steady-state simulations a stopping condition is not defined by the business setting (e.g., nonstop operation) and starting conditions are not of importance (ibid.).

In the context of our method, analysts aim to determine the influence of service incidents on business process metrics. Hence, they define the time interval over which business process behavior is simulated (i.e., the simulation time frame[17]) in such a manner that effects of service incidents are completely observed.

Consequently, in case of terminating simulation studies, analysts define the stopping condition in such a way that all effects on values of business process metrics are captured. With regard to the simulation of continuous business operations (i.e., in case of steady-state studies), analysts proceed as follows: First, they initiate a simulation run and await the system to achieve its steady state. Then, they trigger the occurrence of a service incident. At last, they stop the simulation run as soon as the system has recovered and reached its steady state again (Pang & Whitt, 2009).

We note that the simulation time frame needs to be the same in all simulation experiments, in order to ensure comparability between base case and service incident scenarios.[18]

***Sub-step 3.D.2: Set the Number of Replications***

The simulation model is supposed to emulate the behavior of business processes. Therefore, analysts have to precisely describe the behavior of its dynamic elements (e.g., processing times of activities and resources), which may follow certain probability distributions.

"Random input induces randomness in output" (Kelton *et al.*, 2004, p.38), however. That is, in case of stochastic behavior of model elements, business process metrics (output variables) turn into random variables as well. As a consequence, the values of a

[16] This type of simulation experiments is also referred to as "bounded horizon studies" (Birta & Arbez, 2013, p.72).

[17] The simulation interval is also denoted as "observation interval" (Birta & Arbez, 2013, p.21).

[18] With regard to the simulation of continuous business operations, a number of simulation runs is required in order to determine how long it takes in the worst case until the system reaches its steady state again (after the occurrence of a service incident). This worst case defines the simulation time frame.

business process metric which are recorded in different runs of a stochastic simulation experiment will usually differ from replication to replication as discussed.[19]

Therefore, analysts have to run each configuration of the general simulation model (which implements a specific simulation scenario) multiple times and handle randomness (with regard to the resulting values of business process metrics) through statistical analysis of simulation outcomes (Banks *et al.*, 2010, pp.439-440). They may, for instance, estimate the expected value of a business process metric — as well as the expected change thereof due to the occurrence of service incidents. In the following, we discuss how analysts may proceed in order to (approximately) determine the expected value of a single business process metric in a specific simulation scenario at reasonable cost.

Formally, each business process metric can be regarded as a random variable $Y$. Consequently, a recorded value of such a business process metric, which results from a single simulation run, may be considered as a 'realization' $Y_r$ thereof. The more often a configuration of a simulation model is run, the larger is the sample of realizations $(Y_1, Y_2, ..., Y_n)$ which the estimate of the expected value $E(Y)$ can be based upon.

Analysts may, for instance, calculate the sample mean $\bar{Y}$ (see Equation 6.1), which is an unbiased point estimate for the expected value of a random variable (Bol, 2003, p.71; Birta & Arbez, 2013, p.245) in case of independent[20] and identically distributed[21] realizations.[22]

$$\bar{Y} = \frac{1}{n} \sum_{r=1}^{n} Y_r \tag{6.1}$$

In general, the accuracy of the estimate increases with the number of sample values available. However, simulation experiments are costly not only in terms of time. Therefore, analysts may want to reduce the number of simulation runs $n$ and accept a lower degree of accuracy with regard to the estimate $\bar{Y}$. In such a case, they can (approximately) determine the minimum number $n_{min}$ of replications required to achieve a certain accuracy by taking the following steps, as suggested by Kelton *et al.* (2004, pp.261-262):

- First, analysts define how precise the estimate of the expected value has to be. That is, they specify the maximum difference between the estimated value $\bar{Y}$ and the actual expected value $E(Y)$ they are willing to accept. This deviation can be regarded as the half width $h$ of an interval with center $\bar{Y}$ which should

19 In contrast, every individual run of a deterministic simulation model (which does not consider random variables) results in the same "set of outputs" (Banks *et al.*, 2010, p.33). Section 2.2.1 provides a brief overview on different types of simulation models.

20 Since every replication uses specific random numbers, this condition is fulfilled.

21 This condition is satisfied since the same simulation model is used to generate simulation results (Kelton *et al.*, 2004, p.39).

22 We note that the sample mean is a random variable itself since it is calculated based on a random sample (Kelton *et al.*, 2004, p.610).

contain $E(Y)$.[23] Furthermore, they specify the confidence level $(1-\alpha)$ which defines the desired probability of $E(Y)$ to lie in the interval around $\bar{Y}$.

- Then, they record an initial number $n_{initial} > 30$ of realizations $(Y_1, Y_2, ..., Y_{n_{initial}})$ of the random variable regarded (i.e., the business process metric under consideration) through simulation runs.[24] Afterwards, analysts calculate the standard deviation $S_{initial}$ of the values recorded in the previous step.
- Finally, they determine the number of replications $n_{min}$ needed to achieve the predefined accuracy of estimate (see the first step), through solving the equation

$$n_{min} \approx z^2_{1-\frac{\alpha}{2}} \cdot \frac{S^2_{initial}}{h^2}, \tag{6.2}$$

where the term $z_{1-\frac{\alpha}{2}}$ represents the $1-\frac{\alpha}{2}$ quantile of the standard normal distribution. Appendix B elaborates on the derivation of Equation 6.2. Birta & Arbez (2013, p.246) present a similar method to calculate the minimum number of replications needed to achieve a sufficiently accurate estimate.

For each simulation scenario a specific number of replications is needed to achieve a certain accuracy. We suggest analysts to choose the maximum of these values as the target number of replications for experiments with all simulation scenarios. This ensures appropriate exactness of simulation results and facilitates the comparison of results across simulation scenarios.[25]

***Sub-step 3.D.3: Select Simulation Scenarios to be Analyzed***

As noted before, the basic configuration of our general simulation model corresponds to a simulation scenario in which no service incident occurs (i.e., to the base case). Therefore, analysts have to set up a separate configuration of the general simulation model for each service incident scenario to be examined. In other words, each service incident scenario needs to be implemented as a modified version of the general simulation model.

Service incident scenarios differ from one another in at least one of the following characteristics: the service incident class, the service incident level and the time of service incident occurrence regarded. These three characteristics of a particular service incident may have an influence on the values of business process metrics.

*If analysts, for instance, wanted to assess the influence of outage incidents (service incident class) within the time frame of a typical business day, they could define several model configurations, which differ in the duration of the outage incident (service incident level) and the time of the day at which the service incident occurs regarded.*

[23] $E(Y)$ might be lower or higher than the estimated value $\bar{Y}$ but has to lie in the interval $[\bar{Y}-h; \bar{Y}+h]$.

[24] For an explanation of why we need to assign a value greater than thirty to $n_{initial}$ see Appendix B.

[25] We discuss the comparison of simulation results across simulation scenarios in step 3.E.

Usually there is a large number of potential service incident scenarios. The more simulation experiments analysts conduct, the better they should be able to distinguish the business criticality of different service incident scenarios. This information is of particular interest for IT departments with respect to the negotiation of service offers with external providers.[26]

Since simulation experiments are costly, however, analysts may need to decide on a few scenarios to be studied in detail only. In order to reduce the number of options to choose from (i.e., the set of potential service incident scenarios), analysts might, for instance, consider a limited set of service incident levels. They might divide the value range of each attribute of a service incident class into a set of disjoint intervals and, afterwards, calculate the Cartesian product of the different sets of attribute intervals for each service incident class. The result is a discrete set of service incident levels for each service incident class as the following example illustrates:

*Two service attributes describe the service incident class 'reduced throughput incident', namely 'reduced throughput incident duration' (measured in minutes) and 'reduced throughput incident degree' (measured in Mbit/s). The value ranges of these attributes are (0;c] and (0;z). Analysts now divide the range of the duration interval into three parts — (0;a), [a;b) and [b;c] — and the range of the degree interval into four parts (0;w), [w;x), [x;y) and [y;z). Afterwards, they calculate the Cartesian product of these two sets of attribute intervals and, thus, obtain twelve service incident levels (3 x 4) to be examined further.*

Another way to reduce the number of potential service incident scenarios is to consider only distinct points in time at which service incidents might occur. Analysts might, for instance, split up the simulation time frame into a set of disjoint time periods. They could interview experts for the business processes regarded and ask them to differentiate time periods in which service incidents should usually have different business impacts. Afterwards, analysts could choose a point in time (i.e., a 'representative value') from each of these intervals. They could consider these points in time as the only times at which service incidents can occur when designing service incident scenarios. Finally, analysts select the set of service incident scenarios to be examined in simulation experiments and set up a separate configuration of the general simulation model for each of them.

***Sub-step 3.D.4: Simulate***

Now that analysts have defined the simulation time frame, the configurations of the general simulation model as well as the number of replications required to achieve accurate simulation results, they actually conduct the simulation experiments. The result of this step is a sample of realizations of each of the selected random variables for each simulation scenario. In other words, analysts obtain a separate set of values for every pair of business process metric and simulation scenario. As an illustration we continue the example given in step 2 (see Section 6.2):

[26] For instance, IT departments need to differentiate time periods with different business cost functions in a request for proposal and define penalty- and reward-rules accordingly in order to be able to select the cost-optimal service offer. We elaborate on this subject in Chapter 7.

*Analysts aim to determine the impact of service incidents on the 'number of products delivered late' (business process metric) through simulation experiments. Therefore, they have, inter alia, specified a service incident scenario which assumes an outage incident (service incident class) with a duration of five minutes (service incident level) to occur at 1:30 p.m. on a typical business day (time of occurrence). The simulation model considers normally-distributed processing times in different stages of production which cause variation with regard to the completion time of products. The result of the corresponding simulation experiment (which comprises five replications) is a set of values describing the number of products which have been delivered late in each simulation run (e.g., {22, 19, 21, 20, 18}).*

**Step 3.E: Analyze Simulation Results**

Based on the results of the simulation experiments conducted in the previous sub-step, analysts finally determine the impact of specific service incidents on business process metrics. In other words, they determine specific values of the business parameter function. Therefore, they compare — for every single business process metric — the simulation results of each service incident scenario with the simulation results of the base case.

We suggest analysts to calculate the 'mean difference' between the recorded values of a particular service incident scenario and the recorded values of the base case for each business process metric. Thus, they may (approximately) determine the expected impact of the corresponding service incident on a business process metric.[27]

In other words, we recommend to calculate the mean difference in order to estimate the expected difference using a set of variables (see Table 6.1). Let $E(\Delta Y_{m,i_j})$ be the expected change of a business process metric $m$ due to a specific service incident which is modeled in the service incident scenario $i_j$. This expected difference may be calculated as the expected value of the difference between the random variable of the business process metric $m$ in the service incident scenario $i_j$ (i.e., $Y_{m,i_j}$) and the random variable of the business process metric $m$ in the base case (i.e., $Y_{m,b}$; see Equation 6.3).

$$E(\Delta Y_{m,i_j}) = E(Y_{m,i_j}) - E(Y_{m,b}) = E(Y_{m,i_j} - Y_{m,b}), \forall(m; i_j) \tag{6.3}$$

In order to calculate the mean difference for a distinct pair of business process measure and service incident scenario analysts may proceed as follows:

First, analysts collate the recorded values regarding a specific business process metric $m$ of a service incident scenario $i_j$ and the base case $b$.

Then, they determine the differences between value pairs of corresponding simulation runs (replications) as illustrated in Table 6.2. For instance, they calculate the

[27] We note that IT departments who assess business impacts of service incidents based on mean values are considered as risk-neutral.

Table 6.1.: Description of Variables

| variable | description |
|---|---|
| $m$ | business process metric, $m \in \{1, 2, ..., u\}$ |
| $s$ | simulation scenario $s \in \{b\} \cup \{i_1, i_2, ..., i_v\}$, where $b$ denotes the base case and $i_j, j \in \{1, 2, ..., v\}$ represents a particular service incident scenario |
| $r$ | replication number with regard to a specific simulation scenario, $r \in \{1, 2, ..., w\}$ |
| $Y_m$ | random variable which represents the business process metric $m$ |
| $Y_{m,s}$ | random variable which represents the business process metric $m$ in simulation scenario $s$ |
| $Y_{m,s,r}$ | value of the random variable which represents the business process metric $m$ in replication $r$ of simulation scenario $s$ |

difference between the recorded value obtained in the third simulation run (i.e., in the third replication) of the service incident scenario $Y_{m,i_j,3}$ and the recorded value observed in the third simulation run of the base case $Y_{m,b,3}$. The resulting set of differences ('deltas') represents a sample of a random variable on its own.

We note that base case and service incident scenario have to be simulated with "common random numbers" (e.g., Banks *et al.*, 2010, pp.481-492) in order to allow the application of this approach (see Appendix C for details).

Table 6.2.: Calculation of the Mean Difference of Base Case Sample and Service Incident Sample

| sample | replication | | | | sample |
|---|---|---|---|---|---|
| | 1 | 2 | ... | w | mean |
| base case | $Y_{m,b,1}$ | $Y_{m,b,2}$ | ... | $Y_{m,b,w}$ | $\overline{Y}_{m,b}$ |
| service incident | $Y_{m,i_j,1}$ | $Y_{m,i_j,2}$ | ... | $Y_{m,i_j,w}$ | $\overline{Y}_{m,i_j}$ |
| difference | $\Delta Y_{m,i_j,1}$ | $\Delta Y_{m,i_j,2}$ | ... | $\Delta Y_{m,i_j,w}$ | $\overline{\Delta Y}_{m,i_j}$ |

Afterwards, analysts compute the sample mean $\overline{\Delta Y}_{m,i_j}$ of this set of deltas (using Equation 6.1) and, thus, (approximately) determine the expected change in the value of the business process metric $E(\Delta Y_{m,i_j})$, which can be attributed to the service incident.

Additionally, they may calculate the confidence interval of the sample mean of the set of deltas (see Appendix B). If this confidence interval includes the value 'zero',[28] there is no statistically significant difference between the scenarios at the confidence level specified (e.g., Kelton *et al.*, 2004, p.612; Banks *et al.*, 2010, p.483). That is, there is no strong statistical evidence for the service incident to affect the business process metric regarded.

[28] That is, the expected change in the value of the business metric might be zero.

We note that the mean difference is to be computed for each pair of business process metric and service incident scenario. A value pair of service incident level and expected change of a business process metric (with regard to a specific combination of service incident class and time of occurrence) represents a point on the corresponding business parameter function.

*For example, analysts find that a five minute outage incident (service incident level and class), which occurs at 1.30 p.m. (time), results in twenty more products which are delivered late (change of expected value with regard to a specific business process metric). This means that* $(5; 20)$ *is a point on the corresponding business parameter function.*[29]

The result of this third step of our simulation-based method is a set of points on different business parameter functions (see Section 4.3.2).

## 6.4. Determination of Business Cost Functions

Analysts now determine the business cost function for each potential combination of service incident class and time of occurrence. This again is a process of several sub-steps:

First, for each service incident scenario $i_j$, which has been compared to the base case in the previous step, analysts collate the expected changes in the values for all business process metrics $m$.

*For example, analysts find that a five minute outage incident at 1.30 p.m. (which is regarded in the service incident scenario* $i_1$*) results in twenty more products which are delivered late (*$E(\Delta Y_{1,i_1}) = 20$*) and ten more units of unusable semifinished products (*$E(\Delta Y_{2,i_1}) = 10$*).*[30] *Furthermore, they note that a ten minute outage incident at 1.30 p.m. (which is regarded in the service incident scenario* $i_2$*) causes fifty more products to be delivered late (*$E(\Delta Y_{1,i_2}) = 50$*) and thirty more semifinished products to become unusable (*$E(\Delta Y_{2,i_2}) = 30$*).*

We note that the expected changes in the values of business process metrics do not only depend on the particular class and level of a specific service incident (as the example might suggest) but also on its time of occurrence. In other words, the expected changes due to two individual service incidents which occur at different times might differ significantly although these service incidents are of the same service incident level and service incident class. Such differences could, for instance, hinge on the workload in business processes, which changes over time.

Second, for each expected change $E(\Delta Y_{m,i_j})$, analysts determine the expected business costs it induces ($c(E(\Delta Y_{m,i_j}))$) as defined in the corresponding measurement model.

*For instance, analysts calculate that the late delivery of twenty more products (in case of* $E(\Delta Y_{1,i_1}) = 20$*) induces business costs of €2,000, due to a penalty fee of*

29 The domain of this business parameter function is 'outage duration' and its co-domain is 'additional number of products delivered late'.

30 In this case, we assume semifinished products to become unusable if these are not processed in time.

*€100 per product delivered late. Furthermore, they calculate that ten more units of unusable semifinished products (for $E(\Delta Y_{2,i_1}) = 10$) result in business costs of €5,000, due to a loss of €500 per unit damaged. Analogously, they find that the late delivery of fifty more products (in case of $E(\Delta Y_{1,i_2}) = 50$) induces business costs of €5,000 (due to the penalty fee of €100 per product delivered late) and that thirty more units of unusable semifinished products (for $E(\Delta Y_{2,i_2}) = 30$) result in business costs of €15,000 (due to the loss of €500 per unit damaged).*[31]

Afterwards, for each service incident scenario[32] analysts add up the expected business costs — as calculated in the previous sub-step — across all business process metrics. Thus, they calculate the expected business costs of the particular service incident scenario (see Equation 6.4).

$$E(c(i_j)) = \sum_{m=1}^{u} c(E(\Delta Y_{m,i_j})) \tag{6.4}$$

*For example, they calculate that the business costs caused by the five minute outage incident at 1.30 p.m. (service incident scenario $i_1$) amount to €7,000 through adding $c(E(\Delta Y_{1,i_1}))$ =€2,000 and $c(E(\Delta Y_{2,i_1}))$ =€5,000. Analogously, they determine that the business costs caused by the ten minute outage incident at 1.30 p.m. (service incident scenario $i_2$) amount to €20,000 through adding $c(E(\Delta Y_{1,i_2}))$ =€5,000 and $c(E(\Delta Y_{2,i_2}))$ =€15,000.*

Finally, for each service incident class and time of occurrence,[33] analysts collate the pairs of service incident level and business costs induced. They use these discrete value pairs in order to estimate the corresponding business cost functions.

*For instance, analysts use the information obtained to determine the business cost function of outage incidents, which occur during the time period which is represented by the point in time '1.30 p.m.'. A five minute outage incident causes business costs in the amount of €7,000 and a ten minute outage incident results in business costs of €20,000. Based thereon, analysts might assume a business cost function through the points $(5; 7,000)$ and $(10; 20,000)$.*

As a result of this step, analysts obtain a set of points on each business cost function — i.e., for each pair of service incident class and time of occurrence considered. Based thereon, they estimate the particular business cost functions.[34]

[31] All values are assumed for purposes of illustration only.

[32] We recall that a service incident scenario is specified through a combination of service incident class, service incident level and time of service incident occurrence.

[33] We note that each time of occurrence (point in time) represents a specific time period, within which experts suppose the business impact induced by similar service incidents (i.e., by service incidents having a certain service incident level) to be the same.

[34] We recall that a business cost function refers to a specific period of time (time of service incident occurrence) and a single type of service incidents (service incident class) only. It describes the business costs induced by a single service incident with respect to its service incident level (see 4.1).

## 6.5. Discussion of the Approach

Our simulation-based method facilitates the estimation of business cost functions in case of well-defined business processes. There are, however, a number of limitations and challenges with regard to its application which we discuss in the following.

First, although simulation represents a powerful approach, it might not be the method of choice in any case.[35] For instance, there might be a lack of knowledge — with regard to the structure of business processes (e.g., if business processes are not defined or documented), the dynamic behavior of activities and resources (e.g., if processing times are unknown) or the dependencies between different business processes and services — which hampers the creation of an appropriate simulation model. In such cases, complementary approaches (see, for instance, Wiedemann (2008) and Akkiraju *et al.* (2012)), such as expert interviews and document analysis, might be more suitable. We have preferred simulation to analytical techniques for business process analysis due to its flexibility. In general, analytical techniques necessitate numerous assumptions and address particular questions only (van der Aalst, 2013, p.10). Moreover, "complex and realistic models [...] defy mathematical analysis" (Banks, 1996, p.245). Especially in case of business process models which contain features as, for instance, non-exponential service rates and "non-Poisson arrivals, routing based on priorities, [...] capacity constraints on subsystems" (Banks, 1996, p.245) or concurrent activities (Dumas *et al.*, 2013, p.235), simulation is a more popular approach (ibid.). Furthermore, many business process management tools allow for simulation, while analytical techniques are not prevalent in practice (van der Aalst, 2013, p.10).

Second, the development of an accurate simulation model might be costly and time-consuming. Consequently, it may be reasonable to evaluate the advantages of a simulation study over alternative approaches (e.g., in terms of accuracy of results) in a preliminary study.

Third, we have supposed analysts to determine business costs induced by single service incidents based on expected changes of business process metrics. That is, we have concentrated on the case of risk-neutral IT departments, who decide considering expected values only (e.g., Bamberg *et al.*, 2012, pp.81-82). However, IT departments with other risk preferences could also evaluate a sample of values in a different way. They could, for instance, consider the sample standard deviation as well.

Fourth, in our simulation-based method, we suggest to assess the impact of single service incidents on business operations. Thus, we implicitly assume business consequences which result from different service incidents not to affect one another. If such mutual independence of business consequences of 'different origin' must not be assumed in a certain business setting, we can simulate the impact of combinations of service incidents (which, for instance, occur at the same time or one shortly after the other) on business process metrics as well. We could, for example, define a simulation scenario in which an outage incident with a duration of ten minutes follows one hour

[35] Banks *et al.* (2010, pp.22-24) present a list of ten rules which state when simulation is not an appropriate approach. Furthermore, they discuss advantages and disadvantages of simulation.

after a reduced throughput incident with a duration of thirty minutes and a degree of 900 Mbit/s.

## 6.6. Related Work

This section presents related work which focusses on the assessment of the adverse business impact of imperfect service quality without particular consideration of service quality measures.[36] The articles discussed in the following represent the selection of most closely related publications from the results of an extensive literature review.

Conducting our literature review, we followed the approaches suggested by vom Brocke *et al.* (2009) and Webster & Watson (2002). We used Google Scholar as primary source and conducted a keyword-based full-text search without temporal restrictions. Search terms applied were 'business impact', 'business loss', 'business cost', 'business process', 'process simulation', 'discrete-event simulation', 'service interruption', 'service disruption', 'service incident', 'service outage', 'downtime cost', 'downtime loss', 'interruption cost' and 'operational risk', as well as combinations thereof. In abstract- and, if appropriate, full-text-analyses, we identified peer-reviewed journal and conference papers as well as doctoral theses. We complemented the results obtained through a systematic forward and backward search using the same selection criteria.

**Akkiraju *et al.* (2011) & Akkiraju *et al.* (2012): "Towards Effective Business Process Availability Management"**

Akkiraju *et al.* (2011) and Akkiraju *et al.* (2012) present an empirical approach to determine business costs which are induced by planned maintenance measures causing business process unavailability. They use the information obtained in order to schedule maintenance activities and minimize the resulting expected total business impact.

The authors consider that the business impact of a planned outage on a certain business process depends on outage duration, time of outage occurrence and the country the particular process is operated in. For each business process they collect data about the influence of outages on "business performance metrics" (Akkiraju *et al.*, 2012, p.321). Data on "qualitative metrics" is gathered through surveys among process owners in the different countries using Likert scales (Akkiraju *et al.*, 2012, p.329). Data on "quantitative metrics" is obtained from data analysis (e.g., databases and transaction logs) or collected through surveys as well (Akkiraju *et al.*, 2012, p.329). The authors note that quantitative data is usually not readily available but dispersed across various systems — which makes data collection a difficult task.

As opposed to Akkiraju et al., we do not only conduct surveys and analyze existing data but also use business process simulation in order to measure the expected adverse impact of service incidents on the values of business process metrics. Furthermore, we exclusively consider the business impact of 'unplanned' service incidents.

[36] In Section 4.5 we have presented approaches which concentrate on the selection of optimal service quality objectives, which we attribute to the field of Service Level Management.

### Cheng *et al.* (2005): "Modeling Operational Risk in Business Processes"

Cheng *et al.* (2005) develop an approach to model and assess operational risks in business processes and to identify cost-efficient countermeasures to mitigate these. Therefore, the authors describe the frequency distribution and the severity distribution of "events" through two independent distribution functions.[37] The severity of an event is regarded as its duration.

Events are assumed to affect distinct resources which, in turn, enable certain tasks of business processes. Thus, the influence of resource malfunction (i.e., of the "human, physical and logical infrastructure" (Cheng *et al.*, 2005, p.23)) on business processes is modeled. Furthermore, the amount of "operational loss" due to not operating a certain task for a specific time period is described through a random variable. That is, the model allows for the case of non-constant marginal business cost functions as well.

Cheng *et al.* (2005) regard an internal business setting. Consequently, they assume the optimizing party to have complete information about business processes and resources as well as about distributions of operational loss, event frequencies and event severities.

In contrast to the work at hand, the authors do neither consider service quality measures nor their suitability for the assessment of business impact from a service customer point of view. While we consider the total business costs of different service offers in order to select the cost-optimal service solution, Cheng *et al.* (2005) analyze the operational risk of an already existing 'solution'.

Additionally, this thesis suggests a method for the determination of business cost functions using discrete-event simulation (see Chapter 6), whereas Cheng *et al.* (2005) assume the existence of records about operational loss with respect to various types of events and durations of business disruption. Consequently, the approach by Cheng *et al.* (2005) requires business processes to have been analyzed for some time in the past.

### Pang & Whitt (2009): "Service Interruptions in Large-Scale Service Systems"

Following an analytical approach, Pang & Whitt (2009) examine the impact of (single) service interruptions on the performance of "many-server service systems" (e.g., call centers, Pang & Whitt, 2009, p.1510)). More precisely, they study Markovian queueing models of these "large-scale service systems" in order to find out if the adverse impact of service outages increases with the number of "servers" (i.e., the amount of units which provide service) the system consists of. The authors aim to use this knowledge in order to weigh efficiency gains due to economies of scale (i.e., due to the design of even larger service systems) against the corresponding adverse business impact due to service interruptions.

[37] This is common practice in operational risk management (e.g., Bolancé *et al.*, 2012, p.6; Kenett & Raanan, 2011, p.30).

In their article, Pang & Whitt (2009) consider different performance measures, such as "recovery time" (the time the service system requires to return to its steady state), "maximum queue length" (the maximum number of entities waiting to be served) and "average waiting time" during service outage (the time an entity has to wait on average until it is served if an interruption occurs). Furthermore, they analyze different cases with regard to the behavior of entities during service outage. For instance, they regard a case in which new entities are not allowed to enter the service system during a service interruption and waiting entities leave it.

The approach by Pang & Whitt (2009) allows for a general understanding of the influence of service outages on many-server service systems. However, the article does not provide a general approach to assess the influence of service incidents on business operations.

**Salmela (2007): "Analysing Business Losses Caused by Information Systems Risk — A Business Process Analysis Approach**

Based on a literature review, Salmela (2007) compares different methods to identify business costs which may be induced by "realised [information] system risk[s]", namely "labour cost analysis", "lost profit analysis", "information asset value analysis", "business process analysis" and "stock market analysis".

Furthermore, the author suggests the application of business process analysis in order to identify business costs which result from system availability problems. Salmela (2007) first identifies key business processes and describes their main steps in process models. Afterwards, for each of these steps he determines potential business consequences which might result from information system outages. Finally, the author assesses each business consequence in terms of "business loss" considering its probability and severity. He applies the approach in two case studies and, thus, demonstrates its applicability in practice.

In contrast to the simulation-based method presented in this thesis, Salmela (2007) follows a qualitative approach, which is based on interviews and workshops. He focusses on the severity and probability of business consequences but does not discuss in detail that the business impact of a service incident depends on its particular service incident level.

**Setzer *et al.* (2010): "Change Scheduling based on Business Impact Analysis of Change-Related Risk"**

Setzer *et al.* (2010) present an approach to support service providers in scheduling service changes.[38] Their model aims to minimize expected provider costs due to change-related service outages, which disrupt customer business processes and, thus, oblige a provider to pay penalties.

The work by Setzer *et al.* (2010) comprises an "analytical model" to estimate the business impact of service outages — which differ in duration and occur at different

[38] ITIL defines a change as "the addition, modification, or removal of anything that could have an effect on IT services" (Taylor *et al.*, 2007b, p.227; Hunnebeck *et al.*, 2013, p.444).

times — from a provider point of view. Therefore, the model leverages information about customer business operations, which is stipulated in service level agreements, namely process transaction-based penalty rules.[39]

Penalty rules define the amount of money to be paid in dependence of the number of transactions which are disrupted in a certain way and to a certain degree (e.g., delayed by five minutes). The model allows a formal description of process and service dependencies as well as IT service hierarchies. Finally, change scheduling models are evaluated using discrete-event simulation.

In contrast, our simulation-based method provides guidance for service customers (i.e., IT and business departments) on how to proceed in order to assess business costs induced by single service incidents. Additionally, we support the identification of business consequences, which are mapped to business process-related parameters (measurement model).

Furthermore, in our method, service customers simulate comprehensive, formal business process models in order to assess the effects of service incidents on business process metrics and, thus, quantify resulting business costs. As opposed to Setzer *et al.* (2010), we use this information to actually define penalty- and reward-rules (see Section 7.1) for service level agreements. Additionally, our business process simulation-based approach allows for the consideration of mitigating effects of contingency plans on the adverse impact of service incidents.

**Suh & Han (2003): "The IS Risk Analysis Based on a Business Model"**

The authors present a method to value IS assets "from the viewpoint of operational continuity" (Suh & Han, 2003, p.149), which is based on the Analytical Hierarchy Process. Therefore, they identify the "relative importance" of assets for business functions and estimate the value created by each business function by considering its contribution to the organization's daily average income.

Afterwards, the authors assess the impact of outage incidents which affect an asset assuming the income that is usually created by the corresponding business functions to be lost. In other words, they link asset unavailability to business impact, which depends on an outage's duration, the organization's average income per time unit and an asset's relative importance. The information needed to assess outage incidents is obtained in expert interviews.

In contrast to the work at hand, Suh & Han (2003) do not regard the influence of incidents on business process metrics. Furthermore, they estimate business impact considering an organization's daily average income only. That is, they do not take into account further potential business consequences and, thus, might obtain imprecise results.

[39] In their model, Setzer *et al.* (2010) use penalties as a proxy for the actual financial impact which customer business departments incur. Since penalties are usually stipulated depending on actual business costs, their model allows to compare service outages with regard to 'relative' business impact.

**Wiedemann (2008): "IT-Notfallvorsorge im betrieblichen Risikomanagement — Entwicklung eines Gestaltungsmodells unter Berücksichtigung ökonomischer Aspekte am Beispiel einer TK-Unternehmung"**

Similar to Salmela (2007), Wiedemann (2008) presents a method to monetarily assess the business impact of IT service outages, which focuses on business processes and is based on expert interviews as well as document analysis. More precisely, he applies business impact analysis in the field of Business Continuity Management dealing with rare service incidents, which have a severe impact on business operations (see also Section 4.3.2, which describes the approach in more detail).

In contrast to this thesis, Wiedemann (2008) does not provide a quantitative approach to measure the influence of single service incidents on business operations. We support the determination of distinct values of business parameter functions through discrete-event business process simulation and, thus, complement his approach.

**Zambon *et al.* (2007): "Model-Based Mitigation of Availability Risks"**

Zambon *et al.* (2007) present a method to quantitatively assess and mitigate availability risks due to IT infrastructure failure. In particular, they suggest an approach to model service delivery environments and describe the propagation of outage incidents across the IT architecture.

Thus, Zambon *et al.* (2007) determine the influence of technical failure on business processes, which are supported by certain service delivery environments. Based on historical data on frequencies and durations of service component failures as well as information about "cost[s] associated to the downtime" of business processes (Zambon *et al.*, 2007, p.77) they calculate the expected overall costs of downtime. Furthermore, the authors elaborate on the selection of countermeasures to reduce these downtime costs. Therefore, they leverage information about the influence of each countermeasure on incidents' frequency or duration and the costs induced by this particular countermeasure. The goal is to identify the "best set of countermeasures" (Zambon *et al.*, 2007, p.80) which minimizes the sum of overall downtime costs and costs of countermeasures.

In contrast to this thesis (and similar to Cheng *et al.* (2005)), Zambon *et al.* (2007) address an internal business setting only and do not consider the definition of service quality measures, which could be stipulated in service level agreements. Moreover, their approach does not support the determination of business cost functions.

**Résumé**

In this section we have discussed related work from different fields of research, such as business continuity management, business process management and operational risk management. Whereas some authors suggest to conduct expert interviews and to analyze data and documents in order to determine the adverse business impact of imperfect service quality, others propose to primarily apply analytical and statistical approaches.

Thus, our literature review shows that there is a variety of different but complementary approaches to assess the business impact of service incidents. However, it also illustrates that, so far, there has not been any comprehensive, simulation-based method to guide service customers (i.e., IT and business departments) in assessing the business costs which result from single service incidents. Thus, the literature review also confirms the novelty and contribution of our work.

## 6.7. Summary

In this chapter we developed a simulation-based method to guide IT and business departments in jointly determining business cost functions in case of well-defined business processes. Our method, which builds on the BIA process by Wiedemann (2008), implements the first step of COSS. It consists of four steps:

First, analysts from IT and business departments identify cash-effective business consequences which are potentially induced by a disruption of business processes. Therefore, they consider the classifications which we introduced in the previous chapter. Second, the team of analysts links each relevant business consequence to one or several business process metrics.[40] Third, analysts simulate the impact of single service incidents on the values of those business process metrics. In other words, they determine specific points on business parameter functions. Finally, based thereon, analysts estimate the particular business cost functions of the different service incident classes at different time periods.

Furthermore, we discussed limitations and challenges with regard to the application of our method and reviewed related work, which focusses on the assessment of the adverse business impact of imperfect service quality. In the following chapter, we present an approach to support IT departments in negotiating service offers with multiple providers.

[40] More precisely, they define a measurement model for each business consequence identified.

# 7. Cost-Optimal Selection of Services through Procurement Auctions

The knowledge about both, service incident patterns and corresponding business cost functions, is a prerequisite for the assessment of the total business costs a service solution induces. This information, in turn, is needed to determine total service costs and, thus, to identify a cost-optimal service solution. Consequently, IT departments have to ensure that providers disclose their knowledge about service incident behavior when offering a service solution and do not have an incentive to define service offers strategically.

Therefore, in this chapter we present two approaches which fulfill this condition and support IT departments in negotiating service offers with multiple risk-neutral providers. More precisely, we suggest procurement auctions[1] as a mechanism to enable IT departments to solve their optimization problem to select cost-optimal service solutions among different service offers.

Our approaches address the last phase of a typical contract negotiation process in IT outsourcing,[2] in which a buyer negotiates contract details with selected risk-neutral providers (e.g., Bräutigam, 2009, p.797). We assume an IT department and providers to exclusively discuss service incident patterns and service price at this stage.

The following sections analyze how IT departments should define penalty and reward functions in order to make providers disclose their private information with regard to service incident behavior. Afterwards, we present our approaches to solve the IT department's optimization problem: a multi-attribute and a single-attribute

[1] In Section 2.2.2 we provide a brief overview on the field of auction theory.

[2] The preceding phases of this typical contract negotiation process are 'request for information' (RfI) and 'request for proposal' (RfP) (e.g., Bräutigam, 2009, p.797).

procurement auction. Finally, after a short discussion of these approaches, we refer to related work and briefly summarize our results.[3]

## 7.1. Definition of Penalty and Reward Functions

Every single service incident induces business costs on the service customer side. Consequently, we suggest that every deviation from service incident patterns promised in a service contract (i.e., deviations from the adverse business impact agreed upon) should either result in a penalty or a bonus payment.

That is, we recommend that a provider has to compensate the customer if (at the end of a reference period) the number of service incidents at a specific service incident level has exceeded the amount of service incidents which is defined in the corresponding service incident pattern for this particular service incident level. Accordingly, we propose that, at the end of a reference period, a provider receives a bonus payment for each service incident at a certain service incident level which has been avoided compared to the amount which is defined in the corresponding service incident pattern for this specific service incident level. Based on these considerations, we formally introduce the concepts 'penalty function' and 'reward function' as follows:

> **Penalty Function:** A penalty function specifies the amount of money a service provider has to pay a service customer due to the occurrence of a service incident — depending on its service incident level — if this service incident is not permitted by the corresponding service incident pattern. We define that a penalty function refers to a specific period of time and a single class of service incidents only.
>
> **Reward Function:** A reward function specifies the amount of money the service customer has to pay a service provider due to the avoidance of a service incident — depending on its service incident level — which would have been permitted by the corresponding service incident pattern but has not occurred. We define that a reward function refers to a specific period of time and a single type of service incidents only.

The following simple example illustrates both of these concepts:

*Let us assume a service incident pattern which states the number (absolute frequencies[4]) of outage incidents while distinguishing between three levels of outage duration (specified by the time intervals* $(0;10]$, $(10;20]$ *and* $(20;30]$ *minutes). Let* $f_{(0;10]} = 5$, $f_{(10;20]} = 7$ *and* $f_{(20;30]} = 1$ *be these absolute frequencies. Furthermore, let* $\pi_{(0;10]}$, $\pi_{(10;20]}$ *and* $\pi_{(20;30]}$ *be the penalty rates the service provider has to pay the customer in case the number of outage incidents with a specific duration exceeds the number*

3 This chapter is based on Kieninger *et al.* (2012b) and Kieninger *et al.* (2013a).

4 For reasons of clarity, we use absolute frequencies of service incidents at distinct service incident levels in this example. Assuming outage durations of 5, 15 and 25 minutes to be representative values regarding the different service incident levels, the service incident pattern can also be specified through the relative frequencies $f_{rel,(0;10]} = 12\%$, $f_{rel,(10;20]} = 51\%$ and $f_{rel,(20;30]} = 37\%$ in combination with the reference value $r = 205$ minutes.

*stipulated in the service incident pattern at the end of a reference period (penalty function). In addition, let* $\rho_{(0;10]}$, $\rho_{(10;20]}$ *and* $\rho_{(20;30]}$ *be the reward rates the customer agreed to pay the provider if, at the end of a reference period, the amount of outage incidents with a certain duration is lower than the number stipulated in the service incident pattern (reward function).*

*Now assume* $f_{(0;10]} = 3$, $f_{(10;20]} = 8$ *and* $f_{(20;30]} = 1$ *to be the absolute frequencies realized in a specific time period in the service delivery phase. Then, the provider receives a bonus payment in the amount of* $2 \cdot \rho_{(0;10]}$ *due to the avoidance of two short outage incidents which were permitted. Moreover, the provider has to pay a penalty in the amount of* $\pi_{(10;20]}$ *due to an additional outage incident of medium duration.*

Different penalty and reward functions might influence the decision of a service provider when preparing a service offer. A provider could, for instance, increase the price of a service solution with increasing penalty rates (regarding distinct service incident levels) if it is aware of quality issues which might result in future penalty payments. On the other hand, a provider could also reduce the price of a service solution with increasing reward rates (regarding distinct service incident levels) if it expects bonus payments and aims to underbid competitors' prices.

Therefore, in the following, we examine how an IT department should define penalty and reward functions in order to be able to calculate total service costs even if service incident behavior is not perfectly stable. In other words, we analyze how an IT department can make providers disclose as much of their knowledge about service incident behavior as possible. Using a detailed but simple example,[5] we show by *falsification* (see Section 7.1.1 and Section 7.1.2) that an IT department will not be able to select the cost-optimal service offer in any case if incident-related penalty and reward rates are either both lower or higher than the business costs a service incident actually induces.

*Let us assume an IT department to ask two external service providers, A and B, to submit their offers regarding a new IT service required. In this case, a service offer consists of two elements, a specification of the absolute frequency of outage incidents with a duration of more than 30 minutes* $f_{(30;\infty)}$ *and a statement about the service price* $p$. *The IT department is aware that each outage incident which lasts longer than 30 minutes actually induces business costs* $b_{actual}$ *of €10,000. In the following we refer to such service incidents as 'long outage incidents'.*

*Both service providers have private information about the service incident behavior of the service solutions they intend to offer. They know the probabilities of their service solutions* $s_A$ *and* $s_B$ *to be unavailable 5, 6 or 7 times for more than 30 minutes within a month (see Table 7.1). Service provider B, for instance, expects its service solution* $s_B$ *to fail 6 times with probability* $P(f_{(30;\infty)} = 6)$ *of 0.2.*[6]

[5] For reasons of clarity, we focus on a single service incident pattern which considers one service incident level only.

[6] We note that this private provider information represents a distribution which refers to a single service incident level of a specific service incident class. Generally, service providers could describe such 'probabilistic' service incident behavior by specifying distributions of service

Table 7.1.: Probabilistic Incident Behavior of Service Solutions $s_A$ and $s_B$

| service solution | $P(f_{(30;\infty)} = 5)$ | $P(f_{(30;\infty)} = 6)$ | $P(f_{(30;\infty)} = 7)$ |
|---|---|---|---|
| $s_A$ | 0.1 | 0.9 | $\approx 0.0$ |
| $s_B$ | 0.1 | 0.7 | 0.2 |

*The IT department defines the penalty and reward rates as follows: If, at the end of a reference period, long outage incidents have occurred which are not covered by the amount stated in the service contract the provider will have to pay a penalty $\pi_{(30;\infty)}$ of €10,000 for each of these outage incidents. On the other hand, the service provider selected will receive a bonus payment $\rho_{(30;\infty)}$ of €10,000 for each long outage incident which has not occurred during the reference period but which would have been permitted according to the number stated in the service contract. Since $\pi_{(30;\infty)}$ and $\rho_{(30;\infty)}$ have the same value as $b_{actual}$, they reflect the adverse business impact which is actually induced by an additional service incident or avoided due to the non-occurrence of a service incident.*

In the following two sections we continue our example and illustrate the effects of incident-related penalty and reward rates which do not correspond to actual business costs.

### 7.1.1. Penalty and Reward Rates Below Business Costs

Let us first consider the case in which the IT department specifies the values of $\pi_{(30;\infty)}$ and $\rho_{(30;\infty)}$ to be lower than $b_{actual}$:

*Based on the information given above, the providers prepare their service offers. First, they determine the costs of delivering the service required using their particular service solutions $s_A$ and $s_B$ and add a profit mark-up. Let the resulting values be €50,000 and €48,100 (net service prices). Then, A and B decide on the amount of outage incidents to state in their service offers. In this example, we assume both providers to define $f_{(30;\infty)} = 6$ since this is the most likely number of incidents to occur. Afterwards, having defined $f_{(30;\infty)}$, they calculate the amounts of penalty and bonus payments they expect due to deviations from the number of long outage incidents selected.*

In the following, we denote the difference between expected penalty payments and expected bonus payments due to deviations from service incident patterns stated as the 'surcharge' . We assume service providers to be risk-neutral and to decide based on expected values (e.g., Bamberg *et al.*, 2012, pp.81-82; Laux *et al.*, 2014, p.96) with regard to future penalty and bonus payments only. Furthermore, we assume service providers to add the surcharge determined to the net service price[7] in order to obtain the (gross) service price.

incident patterns. That is, for each service incident class to be considered they would define a set of tuples of service incident pattern and corresponding probability of occurrence.

7 We denote the sum of service provision costs and profit mark-up as the net service price.

*Provider A does not expect penalty payments since* $P(f_{(30;\infty)} = 7) \approx 0.0$ *(see Table 7.1) but anticipates bonus payments in the amount of* $P(f_{(30;\infty)} = 5) \cdot \rho_{(30;\infty)} = 0.1\cdot$ *€10,000 = €1,000. Consequently, A assumes a surcharge of €-1,000. On the other hand, B expects penalty payments in the amount of* $P(f_{(30;\infty)} = 7) \cdot \pi_{(30;\infty)} = 0.2\cdot$ *€10,000 = €2,000 and bonus payments in the amount of* $P(f_{(30;\infty)} = 5) \cdot \rho_{(30;\infty)} = 0.1\cdot$ *€10,000 = €1,000. Therefore, B assumes a surcharge of €1,000.*

*Finally, both providers add the surcharge determined to the net service price and obtain the service price. In this case, A reduces the net service price due to expected bonus payments (i.e., a negative surcharge) in order to win the contract. Consequently, A offers* $s_A$ *at a price of €49,000 and B specifies a price of €49,100 for* $s_B$*. According to our business setting (see Section 3.2), we describe the resulting service offers* $s_A$ *and* $s_B$ *through tuples defining amount of long outage incidents and service price as* $s_A = (6,$*€49,000) and* $s_B = (6,$*€49,100).*

*Table 7.2 summarizes the preparation of service offers regarding the service solutions* $s_A$ *and* $s_B$*. It also considers cases in which* $\pi_{(30;\infty)}$ *and* $\rho_{(30;\infty)}$ *take values lower than* $b_{actual}$*.*[8] *We notice that service prices differ from case to case since surcharges depend on* $\pi_{(30;\infty)}$ *and* $\rho_{(30;\infty)}$*, which are predefined by the service customer.*

Table 7.2.: Preparation of Service Offers in Case of Penalty and Reward Rates Below Business Costs (Provider Perspective)

| $\pi_{(30;\infty)}=$ $\rho_{(30;\infty)}$ [Euro] | service solution | net service price [Euro] | amount of outage incidents $f_{(30;\infty)}$ stated | surcharge [Euro] | service price [Euro] |
|---|---|---|---|---|---|
| 10,000 | $s_A$ | 50,000 | 6 | -1,000 | 49,000 |
| | $s_B$ | 48,100 | 6 | 1,000 | 49,100 |
| 9,000 | $s_A$ | 50,000 | 6 | -900 | 49,100 |
| | $s_B$ | 48,100 | 6 | 900 | 49,000 |
| 8,000 | $s_A$ | 50,000 | 6 | -800 | 49,200 |
| | $s_B$ | 48,100 | 6 | 800 | 48,900 |

*Having received the service offers* $s_A$ *and* $s_B$*, the IT department analyzes these. First, it calculates the total business costs which would result from the application of the service solutions* $s_A$ *and* $s_B$*. That is, for each service offer it multiplies the number of long outage incidents* $f_{(30;\infty)}$ *stated with the business costs* $b_{actual}$ *that such a service incident causes. Afterwards, it adds the service price and, thus, determines the total service costs this service offer entails. Finally, the IT department selects the cost-optimal service offer.*

*The IT department expects total business costs of* $b_{actual} \cdot f_{(30;\infty)} = 6 \cdot$ *€10,000 = €60,000 through the application of either service solution. Table 7.3 gives an overview on the IT department's considerations when analyzing the service offers* $s_A$ *and* $s_B$*. It considers cases in which* $\pi_{(30;\infty)}$ *and* $\rho_{(30;\infty)}$ *take different values lower than* $b_{actual}$*.*

[8] Within a case we assume $\pi_{(30;\infty)}$ and $\rho_{(30;\infty)}$ to have the same value.

Table 7.3.: Selection of Service Offers in Case of Penalty and Reward Rates Below Business Costs (Customer Perspective)

| $\pi_{(30;\infty)}$= $\rho_{(30;\infty)}$ [Euro] | service solution | amount of outage incidents $f_{(30;\infty)}$ stated | amount of business costs [Euro] | service price [Euro] | total service costs [Euro] | selected |
|---|---|---|---|---|---|---|
| 10,000 | $s_A$ | 6 | 60,000 | 49,000 | 109,000 | $s_A$ |
| | $s_B$ | 6 | 60,000 | 49,100 | 109,100 | |
| 9,000 | $s_A$ | 6 | 60,000 | 49,100 | 109,100 | |
| | $s_B$ | 6 | 60,000 | 49,000 | 109,000 | $s_B$ |
| 8,000 | $s_A$ | 6 | 60,000 | 49,200 | 109,200 | |
| | $s_B$ | 6 | 60,000 | 48,900 | 108,900 | $s_B$ |

We find that only in the first case (in which $\pi_{(30;\infty)} = \rho_{(30;\infty)} = b_{actual}$) the IT department actually selects the cost-optimal service solution $s_A$, which actually minimizes total service costs (see Table 7.3). In all other cases (in which $\pi_{(30;\infty)} = \rho_{(30;\infty)} < b_{actual}$) the IT department considers $s_B$ to be cost-optimal.

### 7.1.2. Penalty and Reward Rates Above Business Costs

In the second part of this example, we consider cases in which the IT department specifies the values of $\pi_{(30;\infty)}$ and $\rho_{(30;\infty)}$ to be greater than $b_{actual}$.

*We now change the sum of service provision costs and profit mark-up of service solution $s_B$ to €47,900. Table 7.4 corresponds to Table 7.2 and summarizes the preparation of service offers regarding these cases.*

Table 7.4.: Preparation of Service Offers in Case of Penalty and Reward Rates Above Business Costs (Provider Perspective)

| $\pi_{(30;\infty)}$= $\rho_{(30;\infty)}$ [Euro] | service solution | net service price [Euro] | amount of outage incidents $f_{(30;\infty)}$ stated | surcharge [Euro] | service price [Euro] |
|---|---|---|---|---|---|
| 10,000 | $s_A$ | 50,000 | 6 | -1,000 | 49,000 |
| | $s_B$ | 47,900 | 6 | 1,000 | 48,900 |
| 11,000 | $s_A$ | 50,000 | 6 | -1,100 | 48,900 |
| | $s_B$ | 47,900 | 6 | 1,100 | 49,000 |
| 12,000 | $s_A$ | 50,000 | 6 | -1,200 | 48,800 |
| | $s_B$ | 47,900 | 6 | 1,200 | 49,100 |

*We recall that service prices differ from case to case since surcharges depend on $\pi_{(30;\infty)}$ and $\rho_{(30;\infty)}$, which are predefined by the service customer. Again, the IT department assumes total business costs of $b_{actual} \cdot f_{(30;\infty)} = 6 \cdot$ €10,000 = €60,000 to be incurred* *due to the application of either service solution. Table 7.5 provides an overview on the IT department's information and decision while considering cases in which* $\pi_{(30;\infty)}$ *and* $\rho_{(30;\infty)}$ *take different values higher than* $b_{actual}$.

Table 7.5.: Selection of Service Offers in Case of Penalty and Reward Rates Above Business Costs (Customer Perspective)

| $\pi_{(30;\infty)}=$ $\rho_{(30;\infty)}$ [Euro] | service solution | amount of outage incidents $f_{(30;\infty)}$ stated | amount of business costs [Euro] | service price [Euro] | total service costs [Euro] | selected |
|---|---|---|---|---|---|---|
| 10,000 | $s_A$ | 6 | 60,000 | 49,000 | 109,000 | |
| | $s_B$ | 6 | 60,000 | 48,900 | 108,900 | $s_B$ |
| 11,000 | $s_A$ | 6 | 60,000 | 48,900 | 108,900 | $s_A$ |
| | $s_B$ | 6 | 60,000 | 49,000 | 109,000 | |
| 12,000 | $s_A$ | 6 | 60,000 | 48,800 | 108,800 | $s_A$ |
| | $s_B$ | 6 | 60,000 | 49,100 | 109,100 | |

*Once more, we recognize that the IT department is only able to select the cost-optimal service solution in the first case (in which* $\pi_{(30;\infty)} = \rho_{(30;\infty)} = b_{actual}$*). If penalty and reward rates are higher than the business costs a long outage incident induces the IT department will erroneously choose service solution* $s_A$.

**Résumé**

By *falsification* we could show that an IT department will not be able to select the cost-optimal service solution among different offers in any case if penalty and reward functions are not identical with business cost functions. Therefore, we suggest to use business cost functions as penalty and reward functions.

If penalty and reward functions correspond to business cost functions it does not make any difference from a customer point of view whether the service performs as promised, whether the customer receives penalty payments which completely compensate the adverse business impact incurred or whether the customer pays rewards in the amount of business costs avoided. The IT department is still able to calculate service solutions' total service costs in case of fluctuations in service quality.

In the next section we show that this rule also avoids that a service provider is tempted to define service offers strategically.

## 7.2. Strategic Behavior of Service Providers

Information asymmetry is the "fact that different people know different things" (e.g., Stiglitz, 2003, p.15). In the business setting addressed in this thesis, only the service customer can obtain precise information about business cost functions regarding specific periods of time and different service incident classes. On the other hand, service providers have detailed information about the service incident behavior of service solutions they offer and the corresponding service provision costs they incur.

In Chapter 4 we found that an IT department requires precise information about service incident patterns in order to be able to determine total service costs and, thus, to identify the cost-optimal service offer. Therefore, service providers must not have any incentive to state service incident behavior strategically.

Furthermore, in the previous section we learned that penalty and reward functions need to be identical with business cost functions to enable an IT department to select cost-optimal service solutions in *any* case. In the following, we bring forward a further argument to define penalty and reward rules which correspond to business cost functions: We show that the '1:1 rule' may also be a prerequisite for making providers specify service incident patterns or service prices in service offers at all.

### 7.2.1. Penalty and Reward Rates Below Business Costs

In this section we will see that a rational provider will not define a service incident pattern in a service offer if the provider is convinced that the values of penalty and reward functions are below those of the corresponding business cost function (with regard to a certain service incident class).[9] Instead, a provider will pretend that its service solution does not cause service incidents and agree to compensate the service customer for every service incident to occur in the amount of money defined in the corresponding penalty function. Furthermore, it will state a service price, which incorporates expected future penalty payments.[10]

This is due to the fact that IT departments value service incident patterns using business cost functions in order to determine total business costs. Service providers, however, use penalty and reward functions to value service incident behavior and to calculate the amounts of penalty and bonus payments they expect when defining service offers.

Therefore, a service solution seems to be more expensive to the IT department if service incident behavior is specified through service incident patterns and not factored into the service price in terms of penalty payments expected. That is, a service provider can pretend that a service solution is cheaper if it does not state service incident patterns — in case the values of penalty and reward functions are lower than those of business cost functions. In other words, service providers have an incentive to state service incident behavior incorrectly. The following simple example illustrates these considerations:[11]

*Let us assume a service provider A to offer a service solution s to the IT department. In this case, the offer consists of two elements, a specification of the absolute frequency of outage incidents with a duration of more than 30 minutes* $f_{(30;\infty)}$ *and a statement about the service price p.*

9 A provider might, for instance, know the business process to be supported by its service solution and, thus, be able to estimate actual business cost functions.

10 In the previous section, we implicitly assumed service providers to consider penalty and reward functions to correspond to business cost functions.

11 For reasons of clarity, we focus on a single service incident pattern which considers one service incident level only.

*The IT department defines the penalty and reward rates as follows: At the end of a reference period the provider will have to pay a penalty* $\pi_{(30;\infty)}$ *of €1,000 for each long outage incident which has occurred and which is not covered by the amount stated in the service contract. On the other hand, the service provider will receive a bonus payment* $\rho_{(30;\infty)}$ *of €1,000 for each long outage incident which has not occurred during the reference period but which would have been permitted according to the number stated in the service contract.*

*From the analysis of historical data the service provider knows that s is unavailable 5 times for more than 30 minutes within a month (*$f_{(30;\infty)} = 5$*). It expects the sum of service provision costs and profit mark-up to amount to €50,000 (net service price). Therefore, the provider considers to offer the service solution defining a long outage frequency* $f_{(30;\infty)}$ *of 5 and a price p of €50,000.*

*The provider is convinced, however, that the actual business costs which the customer incurs in case of a long outage incident must be at least ten times greater than the penalty rate predefined. Consequently, it assumes that the service customer would incur business costs in the amount of at least €50,000 if a service solution is unavailable five times. That is, the provider knows that the service customer would assume total service costs of more than €100,000.*

*Therefore, the provider considers to offer the same service solution specifying a long outage frequency* $f_{(30;\infty)}$ *of 4 only. Since, in this case, it would have to pay a penalty of €1,000 (due to the occurrence of one more long outage incident) it adds this amount to the sum of service provision costs and profit mark-up. The resulting service offer would state the tuple (*$f_{(30;\infty)} = 4; p =$*€51,000). The provider is aware that the service customer valuing this offer would at least expect business costs of €40,000 and, thus, total service costs of more than €91,000.*

*Further thinking about the service offer the provider finds that it should not state a long outage frequency at all but a service price only. It understands that factoring expected penalty payments into the service price makes a service offer look cheaper to a service customer in terms of total service costs. Table 7.7 summarizes these thoughts.*

The example shows that understating the number of service incidents the service provider is able to conceal business costs from the IT department — in case the values of penalty and reward functions are lower than those of business cost functions. Thus, a service provider can pretend that a service solution is cheaper in terms of total service costs than it actually is. Even if service providers agreed to state service incident patterns they would have an incentive to understate the number of service incidents.

Let us now examine how a service provider will behave if the values of penalty and reward functions are greater than the values of business cost functions.

### 7.2.2. Penalty and Reward Rates Above Business Costs

If a service provider is aware that the values of penalty and reward functions are above those of the corresponding business cost function (with regard to a certain

Table 7.6.: Preparation of Service Offers in Case of Penalty and Reward Rates Below Business Costs (Provider Perspective Considering Strategic Behavior)

| number of long outages stated | expected +penalty / –bonus [Euro] | net service price [Euro] | service price stated [Euro] | total service costs (assumption) [Euro] |
|---|---|---|---|---|
| 5 | 0 | 50,000 | 50,000 | 100,000 |
| 4 | 1,000 | 50,000 | 51,000 | 91,000 |
| 3 | 2,000 | 50,000 | 52,000 | 82,000 |
| ... | ... | ... | ... | ... |
| 0 | 5,000 | 50,000 | 55,000 | 55,000 |

service incident class), it will not specify a service price (or even a negative one) in its service offer. Instead, it will state a higher number of service incidents than it actually expects. Thus, it can make sure that it will receive bonus payments in the future (due to service incidents 'avoided') and is able to reduce the service price in the amount of expected rewards at the time the contract is concluded.

In other words, for each additional service incident stated the provider is able to allow a price discount which is greater than the business costs that the service customer will assume due to this additional service incident. Again, this is due to the fact that IT departments value service incident patterns using business cost functions in order to determine total business costs. Service providers, however, use penalty and reward functions to value service incident behavior and to calculate the amounts of penalty and bonus payments they expect when defining service offers.

Therefore, a service solution seems to be more expensive to the IT department if service incident behavior is factored into the service price and not specified through service incident patterns — in case the values of penalty and reward functions are greater than those of business cost functions. That is, a service provider can pretend that a service solution is cheaper in terms of total service costs if it does not state a service price (or even a negative one). We continue our example from above to demonstrate this effect:

*Now, let* $\pi_{(30;\infty)} = \rho_{(30;\infty)} =$ *€12,500 be the penalty and reward rates predefined by the service customer. Furthermore, let* $b_{actual}$ *=€10,000 be the business costs a long outage incident causes and let the sum of service provision costs and profit mark-up of service solution s be €50,000.*

*This time, the provider is convinced that the actual business costs which the customer incurs in case of a long outage incident must be at least one fifth lower than the reward rate predefined. Consequently, it assumes that the service customer would incur business costs of not more than €50,000 if a service solution is unavailable five times. That is, the provider knows that the service customer would assume total service costs of less than €100,000.*

*Therefore, the provider considers to offer the same service solution specifying a long outage frequency* $f_{(30;\infty)}$ *of 6. Since, in this case, it would receive a bonus payment of €12,500 (due to the avoidance of a long outage incident permitted) it could give a price discount in order to win the contract. That is, it could subtract this amount from the sum of service provision costs and profit mark-up. The resulting service offer would state the tuple (*$f_{(30;\infty)} = 6; p =$*€37,500). The provider is aware that the service customer valuing this offer would expect business costs of at most €60,000 and, thus, total service costs of not more than €97,500.*

*The provider realizes that it should further increase the number of long outage incidents and lower the service price. It finds that factoring expected bonus payments into the service price makes a service offer look cheaper to a service customer in terms of total service costs. Table 7.7 summarizes these considerations.*

Table 7.7.: Preparation of Service Offers in Case of Penalty and Reward Rates Above Business Costs (Provider Perspective Considering Strategic Behavior)

| number of long outages stated | expected +penalty / –bonus [Euro] | net service price [Euro] | service price stated [Euro] | total service costs (assumption) [Euro] |
|---|---|---|---|---|
| 5 | 0 | 50,000 | 50,000 | 100,000 |
| 6 | –12,500 | 50,000 | 37,500 | 97,500 |
| 7 | –25,000 | 50,000 | 25,000 | 95,000 |
| 8 | –37,500 | 50,000 | 12,500 | 92,500 |
| 9 | –50,000 | 50,000 | 0 | 90,000 |

We notice that the service provider is able to conceal total service costs from the IT department by overstating the number of service incidents if the values of penalty and reward functions are greater than those of business cost functions. Even if service providers agreed to state a (positive) service price they would have an incentive to overstate the number of service incidents.

Next, we discuss how a service provider will behave if penalty and reward functions are identical with business cost functions.

### 7.2.3. Penalty and Reward Rates Equal to Business Costs

In this section we will see that a provider does not have any incentive to state service incident patterns strategically if penalty and reward functions correspond to business cost functions. Furthermore, we will find that a risk-neutral provider is not able to conceal business costs from the IT department in this case.

This is due to the fact that, in case of deviations from stipulated service incident patterns, a provider faces the same monetary impact as its customer. If an additional service incident occurs the provider will have to compensate the customer in the amount of business costs caused. On the other hand, the provider will receive a bonus payment in the amount of business costs avoided if a permitted service incident does not occur.

Consequently, a provider would have to factor a penalty into the service price if it concealed a service incident it expects.[12] This penalty would increase the service price in the amount of business costs which such a service incident causes.

Likewise, a provider would reduce the service price if it pretended that a service incident will occur which it does not actually expect.[13] This price reduction would correspond to the amount of bonus payments due to the additional service incident stated. Once more, we continue our example from above in order to illustrate these effects:

*Let $\pi_{(30;\infty)} = \rho_{(30;\infty)} =$ €10,000 be the penalty and reward rates predefined by the service customer. Furthermore, let these rates correspond to the business costs a long outage incident causes ($b_{actual} =$€10,000) and let the sum of service provision costs and profit mark-up of service solution s be €50,000.*

*This time, the provider knows that the actual business costs which the customer incurs in case of a long outage incident are identical with the penalty and reward rates predefined. Consequently, it is aware that the service customer would incur business costs of €50,000 if a service solution is unavailable five times. That is, the provider knows that the service customer would assume total service costs of €100,000.*

*Now, the provider considers to offer the same service solution specifying a long outage frequency $f_{(30;\infty)}$ of 4 only. Since, in this case, it would have to pay a penalty of €10,000 (due to the occurrence of one more long outage incident) it adds this amount to the sum of service provision costs and profit mark-up. The resulting service offer would state the tuple ($f_{(30;\infty)} = 4; p =$€60,000). The provider is aware that the service customer valuing this offer would expect business costs of €40,000 and, again, total service costs €100,000.*

*Furthermore, the provider considers to offer the same service solution stating a long outage frequency $f_{(30;\infty)}$ of 6. In this case, it would give a price discount (in order to win the contract) since it would receive a bonus payment of €10,000 (due to the avoidance of a long outage incident permitted). That is, it would subtract this amount from the sum of service provision costs and profit mark-up. The resulting service offer would quote the tuple ($f_{(30;\infty)} = 6; p =$€40,000). The provider finds that the service customer assessing this offer would expect business costs of €60,000 and, once more, total service costs of €100,000.*

*The provider realizes that factoring expected penalty or bonus payments into the service price does not make a service offer look cheaper to a service customer in terms of total service costs. Table 7.8 summarizes these thoughts.*

[12] That is, the provider does not state the service incident in the service incident pattern deliberately.

[13] As stated above, we assume providers to reduce the service price in the amount of bonus payments they expect in order to underbid their competitors.

Table 7.8.: Preparation of Service Offers in Case of Penalty and Reward Rates Correspond to Business Costs (Provider Perspective)

| number of long outages stated | expected +penalty / –bonus [Euro] | net service price [Euro] | service price stated [Euro] | total service costs (assumption) [Euro] |
|---|---|---|---|---|
| 4 | 10,000 | 50,000 | 60,000 | 100,000 |
| 5 | 0 | 50,000 | 50,000 | 100,000 |
| 6 | –10,000 | 50,000 | 40,000 | 100,000 |

The example shows that understating or overstating the number of service incidents a service provider cannot pretend that a service solution is cheaper in terms of total service costs than it actually is. Therefore, a provider does not have any incentive to state service incident patterns strategically in this case.

In the following, we analyze how a service provider will specify a service offer if the IT department does not define reward functions but penalty functions only.

### 7.2.4. The Effect of Applying Penalty Functions Only

When deciding on service incident patterns, providers have to predict future service incident behavior (see Section 4.6.1). Particularly, they need to be aware of deviations from stipulated service incident patterns since these may induce penalty and bonus payments.

If IT departments do not predefine reward functions, but penalty functions only, a service provider is in an uncomfortable situation: It will not benefit from its service solution to outperform service incident patterns stated. On the other hand, it will have to pay penalties if its service solution underperforms these service incident patterns.

Since there is always a possibility that a service solution outperforms the service incident behavior predicted, service providers will not state service incident patterns at all. Instead they will predict future service incident behavior and incorporate expected penalty payments into the service price. Consequently, service providers will 'save' penalty payments if a service solution performs better than expected. Otherwise, if unexpected additional service incidents occur they will have to pay penalties anyway. That is, there is no disadvantage compared to the case in which the provider specifies a service incident pattern and the service performs worse than expected.

**Résumé**

This section demonstrated that service providers will have an incentive to define service offers strategically if they know that the values of penalty and reward functions do not correspond to those of business cost functions. That is, they might try to pretend that their service solutions are cheaper in terms of total service costs. Furthermore, we could show that providers do not have any incentive to state service incident patterns strategically if penalty and reward functions are identical with business cost functions. Moreover, we found that service providers might not specify service incident patterns at all if they cannot benefit from outperforming these.

For these reasons, we suggest the IT department to define penalty as well as reward functions which both correspond to business cost functions. Thus, service providers will not have an incentive to state service incident patterns incorrectly and they cannot conceal total service costs from the IT department.

Based on the insights gained, we present two approaches to support IT departments in negotiating service offers with multiple risk-neutral providers in the subsequent section. More precisely, we suggest procurement auctions as a mechanism to enable IT departments to solve their optimization problem to select cost-optimal service solutions among different service offers.

## 7.3. Procurement Auctions to Negotiate with Multiple Service Providers

Contract negotiations between a service customer and potential external providers in the field of IT outsourcing are usually structured as follows (Bräutigam, 2009, p.797): First, the customer looks for providers offering the service in question and sends them a request for information (RfI). Then, analyzing the information contained in responding providers' answers, the customer further specifies service requirements regarding technical, economic and legal characteristics (e.g., specification of number of users, traffic, mission-criticality, service partnership models) in a 'request for proposals' (RfP). Afterwards, this document is sent to selected providers. Finally, based on the proposals received, the customer defines the short list of providers it starts into negotiations with.

In this work, we use service procurement auctions to conduct these final negotiations. Auctions are a means to elicit private information from bidders (Krishna, 2010, p.6). In our case, we aim to enable the IT department to obtain knowledge about providers' reservation prices for service solutions with particular service incident behavior. Thus, we support the comparison of different service offers in terms of total service costs.

We suppose the IT department to already have selected a few service providers to take part in a procurement auction. Furthermore, we make three assumptions about the content of the RfP: First, the RfP defines all general properties of the service required — except for service quality objectives and service price — according to business needs (see Section 3.2). The contract duration and the length of monitoring periods are specified in the RfP as well.

Second, the RfP specifies service incident classes and corresponding service incident levels in a way that any service incident falls into a single service incident class and has one specific service incident level only. Thus, the IT department ensures that double counting of service incidents is avoided. Risk-neutral service providers sum up all penalty payments (positive) and bonus payments (negative) they expect and add these to the net service price in order to calculate the service price.

Third, the RfP states penalty and reward functions for each service incident class which are identical with business cost functions. In other words, we assume the IT department to reveal the adverse business impact of service incidents to potential providers in order to be able to identify the cost-optimal service offer (see Section 7.1). The application of procurement auctions incentivizes providers to offer service solutions at market-oriented prices, since they "compete for the right to sell" (Krishna, 2010, p.1). Consequently, the IT department will not suffer a disadvantage due to the 'disclosure' of business cost functions.

In the following, we first present a procurement auction approach in which service providers are invited to offer service solutions by stating a tuple of service incident patterns and a service price. Afterwards, we address a scenario in which service providers are asked to state a service price only and have to assure to compensate the service customer for every service incident that occurs if their service offer is selected.

For reasons of clarity, we consider a single reference period only. We note, however, that a service offer might also refer to several reference periods and state service prices and service incident patterns which change over time. Capital budgeting approaches (e.g., Baker & English, 2011) represent a way to compare such multi-period service offers.

### 7.3.1. A Cost-Optimal Multi-Attribute Mechanism

In this section, we assume providers are asked to state their bids regarding a specific service which is requested by the IT department. A bid consists of a tuple of service incident patterns[14] and a service price. Since penalty as well as reward functions are defined in the RfP, providers will not have an incentive to state service incident patterns strategically.

In the following, we define a multi-attribute, first-price, sealed-bid procurement auction (see Section 2.2.2) to structure and conduct the final negotiation of service offers between the customer and potential service providers.[15] That is, we develop an "optimal mechanism" (Krishna, 2010, p.67) in order to minimize the customer's total service costs.

[14] We recall that service providers need to specify a service incident pattern for each combination of service incident class and time period predefined by the service customer.

[15] We preferred the first-price, sealed-bid auction to other 'standard' types of auctions since it maintains confidentiality of providers' service offers and is more widely applied for procurement in practice than second-price, sealed-bid auctions (e.g., Eichstädt, 2008, p.121). According to the revenue equivalence principle, "the [customer's] expected revenue of a first-price auction is the same as the expected revenue in a second-price auction", given the assumptions of the 'symmetric independent private values' (SIPV) model are met (Krishna, 2010, p.19). In case the individual values of bidders are distributed according to different distribution functions, however, "no general ranking of the revenues [of first-price and second-price auctions] can be obtained" (Krishna, 2010, p.46). An explanation of the SIPV model is provided, for instance, by Berninghaus *et al.* (2010, p.242).

In a first-price sealed-bid procurement auction, every service provider is allowed to state a single bid only (one round of bids). Furthermore, the content of bids will not be disclosed to other providers — neither during nor after the procurement auction. At the end of the procurement auction, the particular service offer which induces the lowest total service costs wins the auction.

With their offers service providers assure to deliver the service at the service price stated and according to the conditions specified in the RfP. The procurement auction is structured into two phases. First, all parties prepare the procurement auction. Then, the procurement auction is conducted. The auction process consists of the following steps:

**Preparing the Procurement Auction**

1. Preparation of the RfP (customer): In particular, the customer formally defines the service incident classes as well as the corresponding service incident levels to be considered for the service requested. Furthermore, for each service incident class it specifies the penalty and reward function which are identical with the business cost function (see Chapter 6).

2. Preparation for bidding (providers): Each provider analyzes the service incident behavior of the service solution it intends to offer in order to identify its characteristic service incident patterns (see Section 4.6). Furthermore, each provider determines the service provision costs of its service solution and decides on the profit it wants to make.

**Conducting the Procurement Auction**

3. Invitation of providers (IT department): The IT department invites selected providers to participate in the procurement auction. Therefore, it sends the RfP to these trusted potential suppliers.

4. Definition of bids (providers): Based on the knowledge about its service solution's incident behavior and the penalty and reward functions defined in the RfP, each risk-neutral provider calculates the penalty and bonus payments it has to expect when stating the characteristic service incident patterns. Afterwards, providers add expected penalty payments to and subtract expected bonus payments from the net service price in order to determine the service price (see Section 7.1).

5. Bidding (providers): Each service provider sends the IT department a sealed bid stating the tuple of characteristic service incident patterns and the service price of its service solution.

6. Winner determination (IT department): The IT department selects the particular service solution among all service offers which results in the lowest total service costs. Therefore, for each offer it calculates the total business costs which are induced by the set of characteristic service incident patterns and adds these to the corresponding service prices.

On the one hand, the provider winning the auction will have to compensate the customer in the amount of additional business costs induced if the actual business costs are higher than those the characteristic service incident patterns would have caused. On the other hand, the customer will have to reward the provider in the amount of business costs avoided if actual business costs incurred are lower than those which the stipulated service incident patterns would have induced. Consequently, the total service costs the customer incurs are determined by the service offer selected (fixed value).

In Section 7.2.4 we found that providers will not state service incident patterns in their service offers if IT departments do not define reward functions, but penalty functions only. Let us now examine this case, in which service providers bid nothing but a service price.

### 7.3.2. A Cost-Optimal Single-Attribute Mechanism

If service incident behavior regarding a specific service incident class varies significantly over time it might be difficult to describe it through changing service incident patterns. That is, it could happen that for a certain service solution no characteristic, stable service incident patterns are observable.[16]

Since providers might not want to state service incident patterns in such a case, in this section, we address a scenario in which service providers are asked to only state a service price for the service solution they offer. With their price bids they assure to compensate the service customer for every service incident to occur in the amount defined in the corresponding penalty functions, if their offer is finally selected.[17]

In the following, we define a single-attribute, first-price, sealed-bid procurement auction (see Section 2.2.2) to structure and conduct the final negotiation of service offers.[18] That is, every service provider is allowed to state a single bid only (one round of bids). Furthermore, the content of bids will not be disclosed to other providers — neither during nor after the procurement auction. Finally, the particular service offer which states the lowest service price wins the auction.

With their offers service providers assure to deliver the service at the service price stated and according to the conditions specified in the RfP. The procurement auction is structured into two phases. First, all parties prepare the procurement auction. Then the procurement auction is conducted. We now briefly describe the single steps of the auction process. It differs from the multi-attribute approach since service incident patterns and bonus payments are not considered.

16 We also denote this as the case of 'non-stable' service incident behavior.

17 Reward functions are not considered since providers are not supposed to state service incident patterns in this case.

18 Again, we preferred the first-price, sealed-bid auction to other 'standard' types of auctions for the above mentioned reasons.

**Preparing the Procurement Auction**

1. Preparation of the RfP (customer): In particular, the customer formally defines the service incident classes as well as the corresponding service incident levels to be considered for the service requested. Furthermore, for each service incident class it specifies the penalty function which is identical with the corresponding business cost function (see Chapter 6).
2. Preparation of bids (providers): Each provider analyzes the service incident behavior of the service solution it intends to offer. Furthermore, each provider determines the service provision costs of its service solution and decides on the profit it wants to make.

**Conducting the Procurement Auction**

3. Invitation of providers (IT department): The IT department invites selected providers to participate in the procurement auction. Therefore, it sends the RfP to these trusted potential suppliers.
4. Definition of bids (providers): Based on the knowledge about its service solution's incident behavior and the penalty functions defined in the RfP each provider calculates the penalty payments it expects. Afterwards, providers add expected penalty payments to the net service price and, thus, determine the service price (see Section 7.1).
5. Bidding (providers): Each service provider sends the IT department a sealed bid stating the service price of its service solution.
6. Winner determination (IT department): The IT department accepts the particular bid among all service offers which states the lowest service price.

The service price which is stated in the service offer selected determines the total service costs the customer incurs (fixed value). If a service incident occurs, the service provider chosen will have to compensate the customer in the amount of business costs caused.

## 7.4. Discussion of the Approaches

In the previous sections we suggested the use of procurement auctions to support the last phase of a typical IT outsourcing contract negotiation process. In case of frequent negotiations (e.g., due to changing business requirements) procurement auctions represent an efficient way to consider service offers by multiple risk-neutral providers at the same time. In contrast, bilateral 'bargaining' with the same number of providers might turn out to be more time-consuming. We note, however, that service requirements need to be specified more precisely if a procurement auction is conducted since it does not allow for service properties to be discussed. Advantages of auctions and bilateral negotiations are examined, for instance, by Bajari *et al.* (2008) and Asker & Cantillon (2010).

The procurement auction approaches introduced enable IT departments to solve their optimization problem to select cost-optimal service solutions among different service offers — in case of constant and non-constant marginal business cost functions. They avoid that risk-neutral providers have an incentive to state service incident patterns strategically. Additionally, they ensure that customers do not have a disadvantage if a service incident pattern assured is not realized. The total service costs the customer incurs are determined by the service offer selected (fixed value).

Furthermore, both approaches allow providers to predict the total business costs their service solutions will cause and, thus, to better tailor service solutions to customers' business requirements. Since we do not urge providers to further specify their service solutions, but focus on the adverse impact of service incident behavior providers may freely balance service components (which form the service delivery environment) they apply to support clients' business operations.

We argue that an IT department requires information about business costs in order to make an economically well-founded decision about the service offer to select. Therefore, we assume the IT department to disclose business cost functions to selected service providers by stating these as penalty and reward functions. As we have seen in Section 7.1 even minor distortions of business cost functions might significantly influence service investment decisions with regard to total service costs. We recall that the application of procurement auctions incentivizes providers to offer service solutions at market-oriented prices (Krishna, 2010, p.1) and, thus, avoids that the IT department suffers a commercial disadvantage (Taylor *et al.*, 2007a, p.97) due to the revelation of business cost functions.

We are convinced that providers will agree to penalty rates which correspond to business costs — as long as these are not 'out of balance' with service prices — since they can incorporate all expected penalties into the service price. Furthermore, providers can insure themselves against penalty payments due to service incidents which are caused by disasters. If penalty rates are not in a fair relation to the service price, however, providers might not accept this '1:1 rule' due to the high financial risk it represents.

Moreover, in this work, we suppose providers which take part in a procurement auction to be risk-neutral. Thus, we assume them to add expected penalties to and subtract expected bonus payments from the sum of service provision costs and profit mark-up (net service price) when determining a service price. This assumption should at least be realistic for large providers which are able to spread and diversify risks across numerous, heterogeneous customers. In contrast, risk-averse or risk-seeking service providers would value potential deviations from service incident patterns in a different way and, thus, consider different surcharges when defining service prices. Furthermore, they might value service solutions' service incident behavior differently and, thus, state different service incident patterns. This would hamper the selection of cost-optimal service offers for service customers.

Following our approaches, there is no financial risk with regard to service incident behavior for the service customer. The total service costs the customer incurs are determined by the service offer selected (fixed value, see Sections 7.3.1 and 7.3.2).[19]

[19] Our multi-attribute mechanism defines that the provider winning the auction will have to

The service customer, however, takes the risk that business cost functions, which are applied as penalty and reward functions, reflect the business impact the service customer will actually incur due to service incidents in the future. If the estimate of business cost functions turns out to be imprecise after the service contract has been signed (ex post) the service offer selected might not have been the cost-optimal choice. In other words, through the application of our auction-based mechanisms the service customer selects the particular service offer which is to be considered as the cost-optimal one at the time the contract is signed (ex ante).

We are aware that — in spite of compensation payments by providers — ongoing service incidents might significantly influence customer reputation and even ruin a company. This issue is particularly obvious in this work due to our focus on customers' business costs. Although we assume such adverse business consequences to be adequately reflected in business cost functions, we note that customers should have the opportunity to terminate service contracts (e.g., Sturm *et al.*, 2000, p.68) and switch to other providers — as usual in service relationships.

Consequently, providers have an incentive to stick to service incident patterns promised. If a provider has a bad reputation — e.g., because service incident patterns realized in the past deviated significantly from those defined in service level agreements — it might not be invited to procurement auctions any more.

## 7.5. Related Work

In this chapter, we adapted a 'standard' auction approach to our business setting and considered penalty and reward functions which are 'commonly' specified, essential parts of SLAs (e.g., Hiles, 2002, p.82; Benyon & Johnston, 2006, p.8). As a consequence, we could reference a huge body of literature related to this part. In the following, we exemplarily highlight a few articles which concentrate on service procurement auctions as well as on the definition of penalty and reward functions. Furthermore, we point out similarities of our approaches with work from the field of operational risk management.

### 7.5.1. Service Procurement Auctions

Service procurement auctions foster competition among potential providers (e.g., Asker & Cantillon, 2010) and, thus, support cost-optimal service selection from a customer point of view. In Chapter 2.2.2 we provided a general overview on the field of auction theory.

compensate the customer in the amount of additional business costs induced if the actual business costs are higher than those the characteristic service incident patterns would have caused. Furthermore, our multi-attribute mechanism states that the IT department will have to reward the provider in the amount of business costs avoided if actual business costs incurred are lower than those which the stipulated service incident patterns would have induced. On the other hand, our single-attribute mechanism defines that the service provider selected will have to compensate the customer in the amount of business costs caused if a service incident occurs.

**Bichler (2000): "An Experimental Analysis of Multi-Attribute Auctions"**

Bichler (2000) presents the results of a laboratory experiment in which different forms of single-sided, sealed-bid procurement auctions are compared. The experiment addresses a business setting in which a single buyer aims to acquire a financial derivative that is offered by multiple sellers and that is specified through multiple attributes. Therefore, the buyer discloses its scoring function (which is applied to value offers in multi-attribute procurement auctions) to all potential sellers. Similarly, we suppose the service customer to disclose its business cost functions to potential service providers in our approaches.

The experiment is conducted using an electronic brokerage system. Bichler (2000) assumes additive scoring functions, which implies "mutual preferential independence between the attributes" (Bichler, 2000, p.256). Likewise, we suppose service incidents from different service incident classes not to influence one another and regard each service incident pattern as an individual attribute in our procurement auctions. Bichler (2000) finds that first-score auctions perform significantly better than second-score or English auctions from a buyer point of view — in single-attribute as well as in multi-attribute settings. Furthermore, his experiment results suggest that buyers achieve higher utility scores in multi-attribute auctions in comparison with single-attribute auctions.

**Blau *et al.* (2009): "How to Coordinate Value Generation in Service Networks — A Mechanism Design Approach"**

Blau *et al.* (2009) present a multi-attribute, sealed-bid, second-price procurement auction to identify the socially-efficient composition of services in service value networks.[20] That is, considering every potential composition of services from different providers which fulfills business requirements the authors aim to maximize the common welfare of all parties forming the service value network (i.e., the customer and potential providers). Blau *et al.* (2009) consider the aggregation of service quality attributes across different service modules. Furthermore, they apply contractual penalties in order to discourage providers from stating service quality attributes and service costs untruthfully.

In contrast to Blau *et al.* (2009), we address the selection of the particular service solution (i.e., the "composite service" (e.g., Blau, 2009, p.24)) among different provider offers which minimizes total service costs from a customer point of view. Additionally, we also consider reward functions which specify the amount of money a provider receives in case the service quality achieved is higher than the service quality assured.

### 7.5.2. Definition of Penalties and Rewards

The definition of penalties and rewards is an important aspect of SLA design (Sauvé *et al.*, 2005). Penalties should deter service providers from "faulty behavior" (Rana

[20] Service value networks are "goal-oriented business networks" in which providers offer "complementary as well as substitutive standardized service modules" which may be combined to complex services (Blau, 2009, p.51).

*et al.*, 2008, p.134) whereas rewards may be defined to incentivize service providers to perform better than promised or to state service quality objectives non-strategically. An example for a stepwise penalty function which reflects different service incident levels is provided by Buco *et al.* (2004, p.170). The concept of "reward clauses" is discussed, for instance, by Marques *et al.* (2009, p.82).

**Rana *et al.* (2008): "Monitoring and Reputation Mechanisms for Service Level Agreements"**

Rana *et al.* (2008) differentiate between two main types of penalty approaches, namely "reputation based mechanisms" and "monetary fines" (Rana *et al.*, 2008, p.126). In the first case, a provider's ability to comply with service quality objectives is quantified through a numerical value. Reputation values may be considered by clients when selecting a service provider. They decline if an SLA is violated. The authors note that in practice usually monetary fines are stipulated in SLAs.

They mention that SLAs can also define obligations for service clients — e.g., to react on provider data requests within a certain time. If a client does not fulfill such obligations, a provider might be relieved from its duty to pay corresponding penalties. Furthermore, Rana *et al.* (2008) point out that monitoring is of particular importance with regard to the determination of SLA violations.

**Sturm *et al.* (2000): "Foundations of Service Level Management"**

Sturm *et al.* (2000, pp.68-71) discuss the definition of penalties from a practical point of view. They state that the aim of penalties usually is to serve as an incentive for providers to achieve service quality assured but not to compensate the customer for non-performance. They argue that providers generally have not defined a standard process for penalty payments. According to the authors, provider employees therefore should try everything to avoid such exceptions since they require the approval of decision makers and, thus, increase the visibility of issues within their company.

Sturm *et al.* (2000) consider only those types of penalties as effective which result in discomfort on the provider side. They mention that, on the one hand, smart providers might not accept large penalty payments. On the other hand they admit that unscrupulous providers might avoid expensive efforts to achieve service quality objectives and deliberately pay penalties instead.

## 7.5.3. Operational Risk Management

Operational risk management aims at different objectives as, for instance, the reduction of losses, the improvement of "awareness, objectivity, transparency and accountability of risk" and the improvement of the "efficiency and effectiveness of controls and processes" (Tattam, 2011, p.16). We briefly referred to the field of operational risk management in Section 3.3.2.

In operational risk management risk events are usually characterized by their severity and their frequency (e.g., Cruz, 2002, p.103), which can both be modeled using parametric probability distribution functions. Common approaches leverage these distribution functions in order to calculate an "aggregated loss distribution" (e.g., by running a Monte Carlo simulation) and, based thereon, predict operational loss with a certain degree of confidence (e.g., Kenett & Raanan, 2011, p.30).[21] The most widely applied example is the 'value at risk' approach which uses percentiles of the aggregated loss distribution as risk measures (Bolancé *et al.*, 2012, p.8).

Generally, quantitative approaches from the field of operational risk management assume frequency and severity distribution functions of risk events to be known. In contrast, we address an external setting (see Section 3.1.3) and consider information asymmetry in this work. Furthermore, we do not aim to measure the operational risk induced by a specific service solution already in use but to select the cost-optimal service among different service offers.[22]

## 7.6. Summary

In this chapter we presented two approaches to support IT departments in negotiating service offers with multiple providers. More precisely, we suggested procurement auctions as a mechanism to enable IT departments to solve their optimization problem to select cost-optimal service solutions among different service offers. We defined both procurement auctions in a way which ensures that providers disclose their private knowledge about service incident behavior when offering a service solution and do not have an incentive to define service offers strategically.

We developed our procurement auction approaches based on findings which we discussed in the first two sections of this chapter: Initially, we showed that an IT department will not be able to select the cost-optimal service solution among different offers in any case if penalty and reward functions are not identical with business cost functions. Afterwards, we demonstrated that service providers will have an incentive to define service offers strategically if they know that the values of penalty and reward functions do not correspond to those of business cost functions. That is, they might try to pretend that their service solutions are cheaper in terms of total service costs. Furthermore, we could show that this incentive does not exist if penalty and reward functions are identical with business cost functions. Moreover, we found that service providers might not specify service incident patterns at all if they cannot benefit from outperforming these.

For these reasons, we suggested the IT department to define penalty and reward functions in a way that their values are identical with those of business cost functions.

Afterwards, we introduced our procurement auction approaches. In the first approach, service providers are invited to offer service solutions by stating a tuple of service incident patterns and a service price. Afterwards, we addressed a scenario in which service providers are asked to state a service price only. In that case they have to assure to compensate the service customer for every service incident that occurs if their service offer is chosen.

[21] Therefore, the random variables used to model severity and frequency of a specific type of risk event are often assumed to be stochastically independent.

[22] In Section 3.3.2 we identify further differences between SLE and operational risk management.

Finally, we discussed limitations and challenges with regard to the application of our procurement auction approaches and briefly referred to selected related work focussing on service procurement auctions, the definition of penalty and reward functions and operational risk management. In the following chapter, we present an exemplary instantiation of our method for Cost-Optimal Service Selection.

# Part III.

# Evaluation of the Method for Cost-Optimal Service Selection

# 8. Exemplary Instantiation of the Approach — A Case Study

Following our method for Cost-Optimal Service Selection an IT department first determines the business cost functions concerning a certain service which is required to support business operations. In other words, it collaborates with business departments in order to measure the influence of single service incidents with regard to business consequences using our business process simulation-based method (see Chapter 6). Afterwards, the IT department applies one of our procurement auction approaches as a mechanism to negotiate service offers with multiple providers (see Chapter 7). Thus, the IT department is able to solve its optimization problem to select the cost-optimal service solution among different provider offers.

This chapter demonstrates the applicability of COSS using the example of a typical order picking process (business process) at a warehouse, which is to be supported by a warehouse management system (business service).[1] First, we introduce the theoretical business setting which is addressed in the subsequent sections. Then, we apply our simulation-based method to this scenario and determine the business cost functions with regard to service outage incidents. Afterwards, we illustrate how a procurement auction can be conducted in this scenario, assuming two multi-attribute service offers by two different service providers. Finally, we summarize the findings obtained through the exemplary instantiation of COSS.

For reasons of simplicity, we consider the influence of service incident behavior only on a single business process (order picking), which is heavily dependent on the warehouse management system. Furthermore, we concentrate on a single service incident class (service outage incidents).

[1] In Kieninger *et al.* (2013b) we presented the application of our simulation-based method for the determination of business cost functions on a simplified order picking process.

## 8.1. Introduction of the Case

We take the perspective of the IT department within a mail order company. This IT department aims to support business operations through a warehouse management system, which is to be hosted, managed, and offered as a business service by an external service provider.[2]

Warehouse management systems (WMS) are complex software systems which coordinate the “flow of people, machines, and product[s]” through a warehouse (Bartholdi & Hackman, 2014, p.33). They facilitate the management of “inventory, storage locations, and the workforce” and support various warehouse processes (ibid.).

Order picking belongs to the group of outbound processes. It is the process of removing specific amounts of different articles from stock in order to fulfill a customer order (ten Hompel & Schmidt, 2010, p.34). The picking process is followed by activities such as checking, packing, and shipping (e.g., ten Hompel & Schmidt, 2010, p.51-53; Bartholdi & Hackman, 2014, p.23). Customer orders may be regarded as ‘shopping lists’ which initiate this process.

The picking system which we analyze in this chapter represents an abstraction of real, more extensive picking systems.[3] It is subdivided into three picking areas with four picking stations each (see Figure 8.1). The different areas and picking stations are linked by conveyor belts. Subsequently, we denote the conveyor belts which connect the picking stations within a certain picking area as secondary circuits. Analogously, we refer to the conveyor belts which link the different picking areas as well as the entry and the exit point of the picking system as the primary circuit.

When customer orders enter the picking system, for each order a separate box is manually placed on the conveyor belt of the primary circuit. A box needs to visit one or several picking stations depending on the articles ordered by the particular customer.

At any time the WMS controls which picking stations a box still needs to visit and which of the articles ordered it has collected already. At each picking station a “pick-to-light system” (e.g., Bartholdi & Hackman, 2014, p.43) indicates an operator which quantity of articles from a certain rack and shelf has to be collected and placed into a certain box.

That is, the WMS (business service) enables the company’s order picking process and, thus, measurably contributes to the business value created. A temporary unavailability of the WMS results in an abrupt halt of all conveyors, which will stop immediately if the information required to pick articles and to control the flow of boxes through the picking system is not available.

[2] Hosted application management is a specific form of a ‘business service’ (see Section 2.1.1).

[3] The picking system was designed by Dr. Sven Spieckermann who is chairman of a service provider that specializes in the simulation of company processes (SimPlan AG). He uses the scenario in lectures on ‘Discrete-event Simulation in Production and Logistics’ as a realistic example of picking systems. We embed the picking system in our business setting and analyze the impact of the WMS’s incident behavior on business operations. In contrast, Dr. Spieckermann concentrates on the optimization of the order picking process.

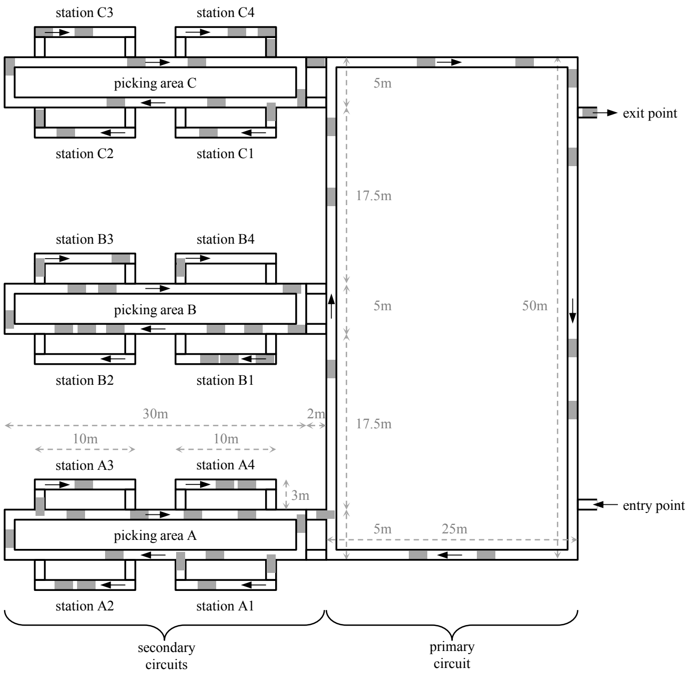


Figure 8.1.: Business Setting: A Picking System with Three Picking Areas

In our theoretical case study, we assume that the IT department aims to apply COSS in order to purchase the cost-optimal service solution from among different provider offers with regard to the needed WMS business service. For this purpose, the IT department and the logistics department requesting the business service have already defined all general properties of the required hosted application management service — except for service quality objectives and service price — according to business needs. The order picking process to be supported has also been analyzed in detail.

We suppose that the IT and logistics departments intend to jointly apply our simulation-based method (see Chapter 6) on this business setting in order to determine business cost functions with regard to service outage incidents. Furthermore, we assume that the IT department intends to negotiate with external service providers using our multi-attribute auction approach for the procurement of cost-optimal service solutions (see Chapter 7).

## 8.2. Determination of Business Cost Functions

The previous section introduced the business process used for the case study. This section illustrates how the IT and logistics departments proceed to determine discrete business cost functions for service outage incidents using our simulation-based four-step method (see Chapter 6). Therefore, a team of analysts estimates changes in the expected values of two business process metrics due to the occurrence of service incidents.

### Step 1: Identification of Cash-Effective Business Consequences

The IT and the logistics department identify two cash-effective business consequences which might be induced by service outage incidents concerning the WMS business service: First, in the evening storekeepers have to work at the picking stations until all orders of the day have been processed. Consequently, service outage incidents may result in additional paid 'overtime work' and, thus, increase cash payments.

Second, the mail order company will grant its customers shipping fee waivers if orders which have been handed in before 1 p.m. are not sent out the same day. In order to be shipped the same day, orders have to be processed before 4 p.m. Due to this condition, outage incidents, which delay order picking, may result in additional shipping fee waivers, which induce decreases in cash receipts.

### Step 2: Linking Business Consequences to Business Parameters

The IT and logistics departments define overtime work as the number of minutes after 6 p.m. which is needed until all orders are picked. Therefore, they record the time at which the last box leaves the order picking system (business process metric). Orders are picked by a group of twelve storekeepers — i.e., one operator per picking station. Each minute of overtime induces cash-effective business costs of €0.25 per storekeeper.

In order to determine business costs induced by shipping fee waivers the IT and logistics departments determine the number of orders which have been handed in before 1 p.m. but leave the picking system after 4 p.m. (business process metric). The shipping fee of a single order amounts to €4.00. The values of process-related metrics are recorded in each simulation experiment.

### Step 3: Simulating the Impact of Service Incidents on Business Process Metrics

Building on the previous steps, the team of analysts from the IT and the logistics departments develops a simulation model of the picking system in order to determine the impact of single outage incidents on the values of the process-related business metrics specified above.

First, they collect and analyze further data on the business process (step 3.A): They determine that the time it takes a storekeeper to pick an order at a picking station is uniformly-distributed with a minimum value of 0.1 and a maximum value of 0.3 minutes. Furthermore, they measure the dimensions of the primary and the secondary circuits (as depicted in Figure 8.1) as well as the speed of conveyors (100 meters per minute). Moreover, they analyze the probabilities of a box having to stop at a specific number of picking stations (see Table 8.1).

The maximum number of containers allowed in each of the secondary circuits is limited to 60 boxes. In other words, a container will only be able to leave the primary circuit and to enter into a picking area if this secondary circuit is not crowded. If a box cannot enter a picking area which it needs to visit, it continues to circle on the conveyor belt of the primary circuit.

Table 8.1.: Probability Distribution of the Number of Picking Stations to be Visited by an Order

| Number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Probability | 14% | 13% | 10% | 10% | 9% | 8% | 7% | 6% | 6% | 6% | 6% | 5% |

Similarly, picking stations have a capacity of three containers each. While one order is processed, two more boxes can be buffered. Boxes which cannot access a picking station they need to visit because the buffer is (temporarily) full have to go past this picking station and continue to circle on the conveyor belt of the corresponding picking area (secondary circuit). As soon as all articles of an order have been placed in the corresponding box this container is sent to the exit point of the system.

The picking system features two characteristics with regard to the flow of containers: First, a box does not have to visit the different picking stations (which it needs to stop at) in a specific order (i.e., picking stations are independent of one another). Second, containers are not dynamically routed to specific picking stations or picking areas. In other words, a box tries to enter a picking area or picking station (which it needs to visit in order to collect a certain article) whenever it passes a corresponding switch. [4]

Order picking begins at 10 a.m. every day. A set of orders is released into the picking system at the top of each hour as a batch. Orders which arrive before order picking starts enter the process at 10 a.m. The orders which arrive after 10 a.m., but before 11 a.m., are handed in at 11 a.m., et cetera. At 5 p.m. the last set of orders enters the picking system. Table 8.2 depicts typical batch sizes depending on the time of the day which we will use in our analysis of the order picking process. Altogether, 4,800 orders are processed per day.

Table 8.2.: Size of Order Batches Released for Picking in the Course of a Day

| Time | 10 a.m. | 11 a.m. | 12 noon | 1 p.m. | 2 p.m. | 3 p.m. | 4 p.m. | 5 p.m. |
|---|---|---|---|---|---|---|---|---|
| Size | 384 | 1,152 | 1,056 | 288 | 432 | 864 | 528 | 96 |
| Percent | 8% | 24% | 22% | 6% | 9% | 18% | 11% | 2% |

If all orders which are in the picking system are handled before the next batch is released we will regard the remaining time as idle time. Recall that the mail order company promises its customers same-day shipping of orders which are placed before 1 p.m. In order to keep this promise such orders have to be picked before 4 p.m. ('cutoff time'). Boxes which leave the picking system after 4 p.m., however, will not be sent before the next day.

Second, the team of analysts develops a discrete-event simulation model of the business process (step 3.B; see Figure 8.2) in AnyLogic (e.g., AnyLogic, 2015; Grigoryev, 2012).

[4] Dynamic routing of containers would require a more complex control system and would, thus, be much more expensive.

In particular, they ensure that the number of overtime minutes and the number of shipping fee waivers granted are recorded.

A service incident is modeled as a pair of events — incident start event and incident end event — which occur at predefined points in time. The incident start event stops all conveyors in the picking system (i.e., the speed of all conveyors is set to zero) since the WMS, which controls the routing of boxes through the picking system, becomes unavailable. Furthermore, the incident start event interrupts order picking at all picking stations (the capacity of picking stations is set to zero) since the information required to pick articles is not available any more.[5] The incident end event, in turn, resets the speed of all conveyors as well as the capacity of all picking stations in the picking system to their original values. The time between incident start event and incident end event represents the duration of a service outage incident (i.e., its service incident level).

Before the model is used in simulation experiments, it is thoroughly tested (step 3.C). The team of analysts verifies the simulation model by setting up test scenarios the results of which are predictable.[6] Using deterministic input they test the following characteristics:

- Speed and capacity of conveyors (e.g., through the creation of single orders to examine their transfer time, or through the creation of large batches in order to determine the maximum number of boxes which can be conveyed at once)
- Routing of boxes (e.g., through the definition of orders which require the corresponding box to visit specific picking stations and by analyzing if each picking station is visited)
- Recording of business process metrics (e.g., through the specification of orders which enter and leave the picking system at certain points in time in order to ensure that the number of shipping fee waivers counted is correct)
- Reaction to service outage incidents (e.g. through the occurrence of a service outage incident with a certain duration in order to check if the routing of boxes and the processing of orders at the picking stations is influenced as expected)

Afterwards, the team of analysts sets the simulation time frame, specifies the service incident scenarios to be analyzed and determines the number of replications required to achieve accurate estimates with regard to the expected values of business process parameters (step 3.D).

The analysts decide to model the order picking process as a terminating system since orders are usually processed between 10 a.m. and 6 p.m. only (not considering overtime). They define the simulation period to start in the morning when there are no customer orders in the picking system (initial condition) and to end when the

[5] More precisely, the capacity of a picking station changes to zero as soon as the order which is currently being processed is finished.

[6] Since we regard a hypothetical picking system we cannot compare simulation results to outcomes of a real system.

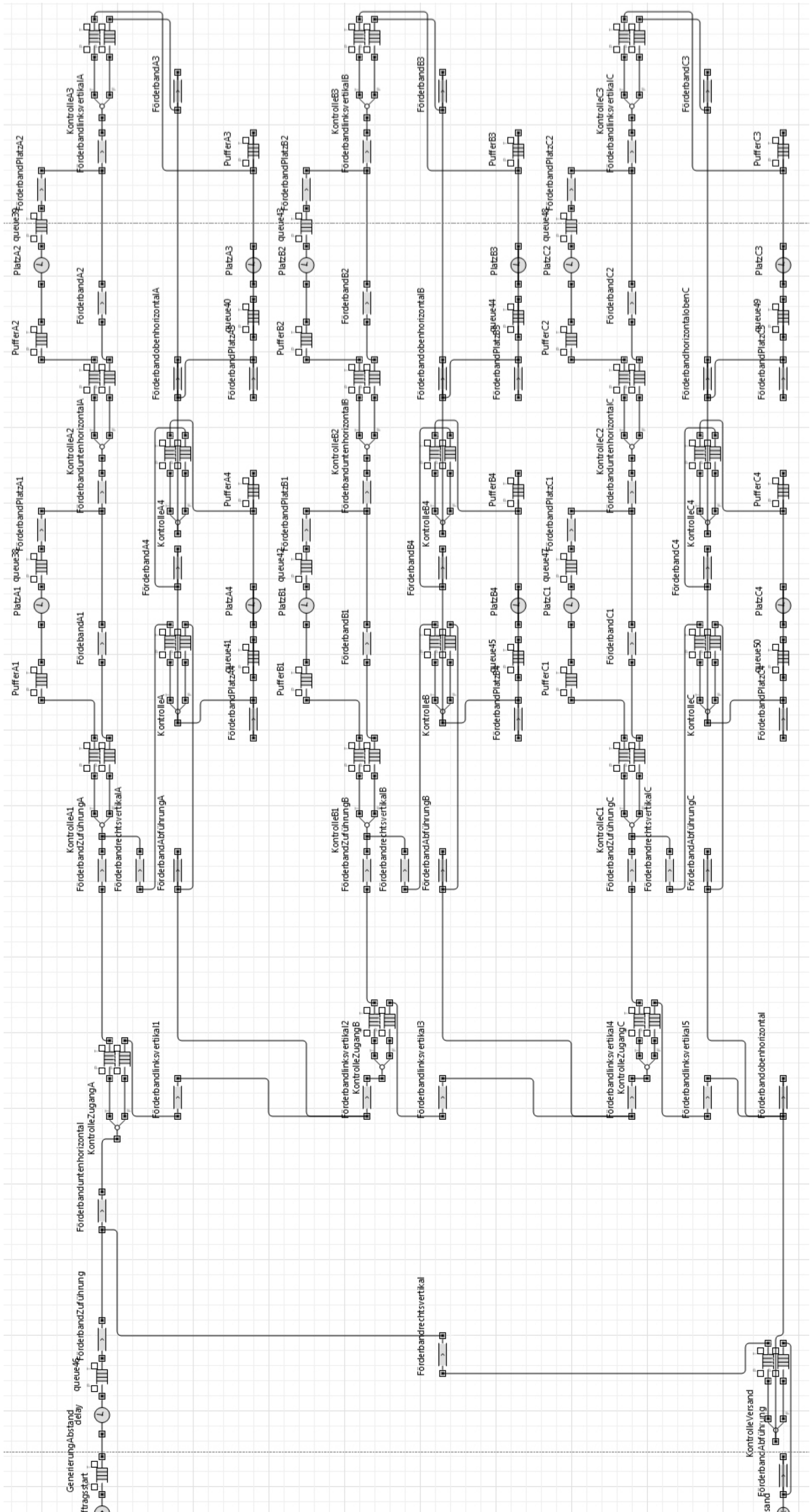

Figure 8.2.: Simulation Model of the Order Picking Process in AnyLogic

last order leaves the picking system (terminating condition). Precisely, they agree to simulate the order picking process from 10 a.m. to 11 p.m. (780 minutes). This simulation time frame suffices to cover the impact of service outage incidents on the expected values of the business process parameters under consideration: Since storekeepers have to work until all orders of a day are processed, service outage incidents do not impact the next day's operations. In other words, delays caused by outage incidents will be made up through overtime work.[7]

Now the team of analysts defines a number of service incident scenarios in order to determine the impact of outage incidents (service incident class) on overtime work and the number of shipping fee waivers. Service incident scenarios differ from one another in the outage duration (service incident level) or in the time of outage occurrence considered.[8] Consequently, a service incident scenario is represented by a tuple of values with regard to these two characteristics.

In order to reduce the number of service incident scenarios to be regarded, analysts concentrate on specific outage durations and times of occurrence (see Table 8.3). Since outages which last longer than 80 minutes are supposed to be extremely rare, they agree to only examine outage durations between 10 and 80 minutes. In order to obtain a comprehensive overview on the impact of outage incidents with certain service incident levels during the day they specify 18 times of occurrence[9] to be considered in the simulation experiments. They keep in mind that the first batch of orders is released into the picking system at 10:00 a.m., and that the last set of orders enters the picking system at 5 p.m. Altogether, they define 145 simulation scenarios (i.e., 144 service incident scenarios (based on the 18 times of occurrence combined with eight outage duration levels) and the base case, in which no outage incident occurs).

Table 8.3.: Service Incident Scenarios Considered

| Time of Occurrence | Outage Duration [Minutes] | | | | |
|---|---|---|---|---|---|
| | 10 | 20 | ... | 70 | 80 |
| 10:00 a.m. | (10:00; 10) | (10:00; 20) | ... | (10:00; 70) | (10:00; 80) |
| 10:30 a.m. | (10:30; 10) | (10:30; 20) | ... | (10:30; 70) | (10:30; 80) |
| 11:00 a.m. | (11:00; 10) | (11:00; 20) | ... | (11:00; 70) | (11:00; 80) |
| ... | ... | ... | ... | ... | ... |
| 05:30 p.m. | (05:30; 10) | (05:30; 20) | ... | (05:30; 70) | (05:30; 80) |
| 06:00 p.m. | (06:00; 10) | (06:00; 20) | ... | (06:00; 70) | (06:00; 80) |
| 06:30 p.m. | (06:30; 10) | (06:30; 20) | ... | (06:30; 70) | (06:30; 80) |

7 The analysts may make this assumption since they will also limit the maximum duration of outage incidents to be considered in the following (see below).

8 We recall that only one outage incident occurs in each service incident scenario.

9 That is, they analyze the impact of outage incidents every half an hour between 10:00 a.m. and 06:30 p.m. Times of occurrence after 06:30 p.m. are not regarded since the team of analysts is aware that outage incidents which occur later only result in little overtime work.

Afterwards, the team of analysts specifies the maximum difference between estimated and actual expected values of additional overtime and number of additional shipping fee waivers (due to outage incidents) they are willing to accept. That is, they decide how precise the estimate of the expected value has to be.[10] They define that the estimated value 'additional overtime work' may be at most 16.67 minutes lower or higher than the actual expected value.[11] Furthermore, they state that the estimated number of additional shipping fee waivers may deviate from the actual expected number by at most 12.5.[12]

Then, the team of analysts records an initial number of 31 realizations of the business process metrics under consideration for each simulation scenario through simulation runs. For each simulation scenario and business process metric they determine the number of replications needed to achieve the predefined accuracy of estimate at a confidence level of 95% using Equation 6.2.[13]

The analysts choose the highest number of replications required as the target number of simulation runs for all simulation scenarios. Consequently, they conduct 50 simulation runs for each simulation scenario and record the numbers of shipping fee waivers needed and the minutes of overtime required to pick all orders.[14]

Finally, the team of analysts examines the obtained simulation results (step 3.E). That is, for each service incident scenario they determine the business impact of the particular outage incident on both the considered business process metrics. In other words, they determine the corresponding business parameter functions (see Section 4.3). Table 8.4[15] and Table 8.5 show the influence of outage incidents with specific durations at different times of occurrence on the number of additional shipping fee waivers[16] and the additional minutes of overtime required, respectively.

### Step 4: Determination of Business Cost Functions

Based on these findings, the team of analysts finally estimates the incurred business costs due to different outage incidents. We recall that each additional minute of overtime induces business costs of €3.00 and that the shipping fee of a single order

[10] The terms 'estimated value' and 'actual expected value' are introduced in the paragraph 'Set the Number of Replications' of step 3.D of our simulation-based method (see Section 6.3.

[11] 16.67 minutes of overtime work correspond to €50 of business costs since a single minute of overtime induces business costs of €3.

[12] 12.5 shipping fee waivers induce €50 of business costs since a single waiver results in business costs of €4.

[13] Section 1 of Appendix D provides the corresponding simulation results for all service incident scenarios in Table D.1 (regarding the number of additional shipping fee waivers) and Table D.2 (regarding additional overtime).

[14] Section 2 of Appendix D includes the detailed simulation results in Table D.3 (regarding the number of additional shipping fee waivers) and Table D.4 (regarding additional overtime).

[15] For presentation purposes, Table 8.4 only considers times of occurrence until 04:00 p.m. The complete results are depicted in Table D.5 of Appendix D.

[16] Since the workload on the picking system is high, a small number of shipping fee waivers (between 0.0 and 1.1 on average) needs to be granted even if there is no outage incident before 4:00 p.m. (see Table D.5).

Table 8.4.: Number of Additional Shipping Fee Waivers (Sample Mean)

| Time of Occurrence | Outage Duration [Minutes] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
| 10:00 a.m. | 0.7 | 1.0 | 0.8 | 3.1 | 4.6 | 14.6 | 30.0 | 63.8 |
| 10:30 a.m. | 0.8 | 1.1 | 1.6 | 2.7 | 6.2 | 16.3 | 26.7 | 65.4 |
| 11:00 a.m. | 1.2 | 2.8 | 7.2 | 15.2 | 29.7 | 66.9 | 150.7 | 239.2 |
| 11:30 a.m. | 1.8 | 2.6 | 5.4 | 12.5 | 31.3 | 72.5 | 140.5 | 229.8 |
| 12:00 noon | 1.2 | 2.6 | 7.4 | 16.0 | 35.9 | 69.6 | 136.1 | 232.2 |
| 12:30 p.m. | 1.5 | 3.5 | 7.9 | 16.9 | 40.9 | 69.7 | 129.9 | 228.0 |
| 01:00 p.m. | 1.2 | 3.3 | 7.6 | 14.6 | 33.3 | 67.1 | 141.5 | 229.7 |
| 01:30 p.m. | 2.0 | 2.6 | 6.9 | 15.1 | 34.3 | 73.5 | 137.3 | 243.4 |
| 02:00 p.m. | 0.7 | 2.4 | 7.1 | 18.1 | 36.8 | 67.2 | 154.6 | 228.3 |
| 02:30 p.m. | 1.1 | 3.5 | 8.0 | 12.1 | 35.5 | 70.5 | 149.2 | 244.2 |
| 03:00 p.m. | 1.4 | 3.2 | 7.1 | 14.9 | 26.5 | 71.4 | 72.3 | 68.1 |
| 03:30 p.m. | 0.8 | 1.6 | 8.1 | 6.4 | 7.0 | 6.3 | 7.2 | 5.7 |
| 04:00 p.m. | 0.6 | 0.6 | 0.5 | 0.7 | 0.6 | 0.4 | 0.5 | 0.9 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |

Table 8.5.: Additional Minutes of Overtime (Sample Mean)

| Time of Occurrence | Outage Duration [Minutes] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
| 10:00 a.m. | 0.8 | 1.2 | 2.8 | 13.0 | 20.7 | 32.3 | 43.2 | 52.9 |
| 10:30 a.m. | 1.0 | 1.1 | 3.4 | 11.7 | 22.3 | 32.7 | 42.0 | 52.5 |
| 11:00 a.m. | 7.0 | 16.1 | 26.5 | 38.0 | 46.6 | 57.4 | 68.7 | 76.8 |
| 11:30 a.m. | 6.6 | 17.6 | 26.5 | 36.0 | 46.5 | 55.7 | 67.9 | 77.0 |
| 12:00 noon | 7.7 | 15.6 | 27.7 | 36.2 | 46.6 | 57.1 | 66.6 | 77.3 |
| 12:30 p.m. | 6.8 | 16.7 | 27.2 | 37.2 | 48.2 | 57.7 | 66.5 | 76.6 |
| 01:00 p.m. | 7.4 | 16.7 | 27.0 | 37.2 | 48.0 | 57.5 | 67.8 | 76.7 |
| 01:30 p.m. | 7.7 | 16.5 | 27.1 | 36.8 | 46.6 | 56.7 | 67.4 | 76.4 |
| 02:00 p.m. | 6.6 | 16.4 | 27.3 | 37.6 | 47.2 | 55.8 | 67.2 | 75.9 |
| 02:30 p.m. | 7.5 | 17.5 | 28.3 | 36.9 | 47.2 | 57.0 | 67.7 | 77.9 |
| 03:00 p.m. | 6.9 | 17.2 | 27.0 | 37.1 | 46.1 | 57.0 | 68.0 | 77.1 |
| 03:30 p.m. | 6.2 | 16.1 | 26.6 | 35.5 | 46.4 | 56.2 | 67.4 | 76.6 |
| 04:00 p.m. | 8.1 | 16.5 | 27.5 | 37.2 | 47.0 | 57.6 | 66.9 | 77.7 |
| 04:30 p.m. | 8.5 | 17.8 | 28.0 | 37.6 | 48.2 | 58.0 | 67.3 | 78.0 |
| 05:00 p.m. | 7.4 | 17.0 | 27.6 | 37.1 | 47.2 | 56.3 | 66.8 | 78.1 |
| 05:30 p.m. | 0.0 | 9.6 | 19.9 | 29.4 | 39.1 | 49.1 | 59.2 | 68.9 |
| 06:00 p.m. | 4.4 | 9.4 | 9.1 | 9.4 | 18.9 | 19.6 | 26.4 | 28.2 |
| 06:30 p.m. | 1.0 | 0.8 | 1.5 | 0.7 | 0.6 | 0.8 | 1.0 | 0.7 |

amounts to €4.00.[17]

The team of analysts observes that the business costs caused by outage incidents differ significantly depending on outage duration and time of incident occurrence. They distinguish five time slots with regard to the impact of outage incidents on business costs (see Figure 8.3). Business costs develop similarly between 10:00 a.m. and 10:30 a.m., from 11:00 a.m. to 02:30 p.m., between 03:30 p.m. and 05:30 p.m. and from 06:00 p.m. to 06:30 p.m. Furthermore, business costs develop differently at 03:00 p.m. In Figure 8.3 discrete points on business cost functions which fall into a certain time slot are marked through squares, triangles, circles, diamonds or asterisks.

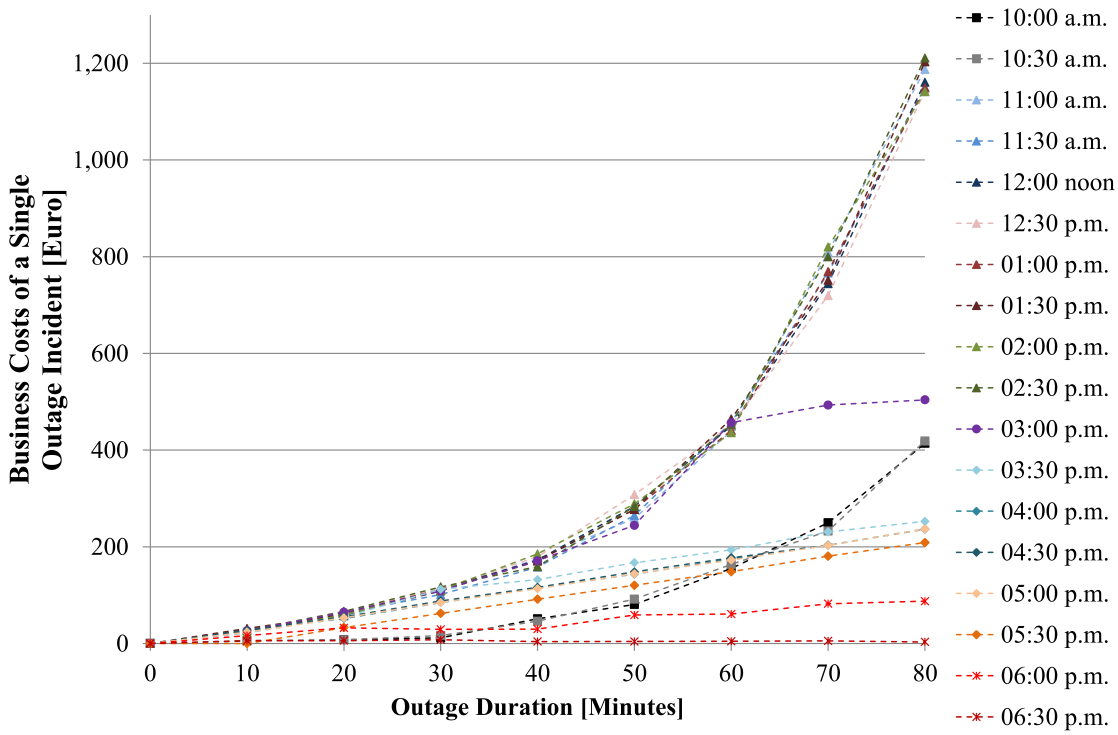


Figure 8.3.: Business Costs of Single Service Outage Incidents with Specific Times of Occurrence (on Average)

The analysts deduce that business costs develop disproportionally (non-constant marginal business cost functions) in all but two of the five mentioned time slots. More precisely, business costs increase linearly only after 03:30 p.m. Therefore, the team of analysts from the IT and the logistics departments argues that service incident patterns have to be used as service quality measures in order to be able to assess the adverse impact of outages of the warehouse management system on business operations.

[17] In Section 3 and Section 4 of Appendix D we provide detailed Figures and Tables distinguishing between business costs induced by additional shipping fee waivers and by additional minutes of overtime.

In the next section, we illustrate how the IT department uses these findings to prepare and conduct negotiations with potential external service providers, which could host and manage the business service ‘warehouse management system’.

## 8.3. Negotiation with Service Providers

The IT department decides to apply our multi-attribute auction approach in order to procure the cost-optimal service solution (see Chapter 7). It has already identified providers offering the business service in question and sent them a ‘request for information’. Considering these providers’ answers the IT department specifies the request for proposals.

**Preparation of the RfP**

In the RfP, the IT department defines all general properties of the required service (according to business needs) except for outage-related service incident behavior and service price. Furthermore, it specifies the length of a reference period to be one month and the service time as ‘10:00 a.m. to 08:00 p.m. on business days’.

We recall that the team of analysts identified five time slots during a business day in which the business costs due to outage incidents develop similarly. Therefore, the IT department decides to only define a business cost function with regard to each of these distinct time slots in the RfP and, thus, to reduce the number of business cost functions to be considered in service offers.

Additionally, the IT department decides to reduce the number of considered service incident levels. Since the business costs due to outage incidents with a duration of 10, 20 and 30 minutes do not differ considerably the IT department calculates the average business costs resulting from such outage incidents for each of the five time slots. It specifies the results as the business costs for outage incidents which last up to 30 minutes. Accordingly, it specifies the business costs due to outage incidents with a duration between 30 and 50 minutes. With regard to outage durations in the intervals $(50; 60]$, $(60; 70]$ and $(70; 80]$ they use the average business costs which are induced by outage incidents lasting 50 and 60 minutes, 60 and 70 minutes and 70 and 80 minutes, respectively. Table 8.6 illustrates the resulting discrete business cost functions which are distinguished by the time of incident occurrence.

Moreover, the IT department states penalty and reward functions for service outage incidents to be identical with these discrete business cost functions. Since service prices and service incident patterns are not assumed to change over time, the IT department decides to consider only a single reference period when comparing different service offers.

**Invitation of Providers**

After the determination of the business cost functions, the IT department invites selected providers to participate in the procurement auction. Therefore, it sends the RfP — including the information from Table 8.6 — to these potential suppliers.

Table 8.6.: Business Costs due to an Outage Incident in Dependence of its Time of Occurrence and Duration (in Euro)

| Time Slot | Outage Duration [Minutes] | | | | |
|---|---|---|---|---|---|
| | (0;30] | (30;50] | (50;60] | (60;70] | (70;80] |
| [10:00 a.m.; 10:30 a.m.] | 8 | 66 | 155 | 250 | 414 |
| (10:30 a.m.; 02:30 p.m.] | 65 | 217 | 440 | 809 | 1187 |
| (02:30 p.m.; 03:00 p.m.] | 67 | 208 | 457 | 493 | 504 |
| (03:00 p.m.; 05:30 p.m.] | 57 | 132 | 177 | 203 | 236 |
| (05:30 p.m.; 08:00 p.m.] | 26 | 44 | 61 | 82 | 87 |

The IT department has selected these providers based on their answers to an initial 'request for information'.

**Winner Determination**

The IT department receives two sealed bids, $s_A$ and $s_B$, from two different providers (A and B)[18], each stating a characteristic service incident pattern (with regard to service outage incidents) and a service price (see Table 8.7).[19] Both providers declare that the stated service incident pattern refers to a time period of 200 hours of service time. More precisely, their bids refer to the time period '10:00 a.m. to 08:00 p.m.' on business days of one month (assuming 20 business days per month).

Table 8.7.: Service Incident Pattern and Service Price per Offered Service Solution

| Service Solution | Outage Duration [Minutes] | | | | | Service Price [Euro] |
|---|---|---|---|---|---|---|
| | (0;30] | (30;50] | (50;60] | (60;70] | (70;80] | |
| $s_A$ | 4 | 2 | 1 | 1 | 1 | 1,400 |
| $s_B$ | 5 | 3 | 1 | 1 | 0 | 1,600 |

Both providers do not further split up this time period since the number of orders in the picking system ('business demand' of the service) is supposed to have no influence on their service solutions' incident behavior. In other words, they do not consider thinner 'time slices' in their service offers because they assume the probability of an outage incident to occur at a certain time during a business day to be independent from the workload on the picking system and equally-distributed over service time.

Thus, the probability of an outage incident — which is stated in a service incident pattern — to occur in a specific time slot depends on the share of this time slot within service time. Table 8.8 shows the corresponding probabilities with regard to

[18] For reasons of simplicity, we only use two offers to illustrate the applicability of our approach.
[19] Again, for reasons of clarity, we use absolute frequencies of service incidents at distinct service incident levels in this example (see also our example in Section 7.1).

the five time slots. For instance, only ten of 200 hours of monthly service time fall into the time slot '10:00 a.m. to 10:30 a.m.' (i.e., 0.5 hours per business day within a month).

Table 8.8.: Probability of an Outage Incident to Occur within a specific Time Slot

| Time Slot | Hours per Month | Probability |
|---|---|---|
| [10:00 a.m.; 10:30 a.m.] | $0.5 \cdot 20 = 10$ | 0.05 |
| (10:30 a.m.; 02:30 p.m.] | $4.0 \cdot 20 = 80$ | 0.40 |
| (02:30 p.m.; 03:00 p.m.] | $0.5 \cdot 20 = 10$ | 0.05 |
| (03:00 p.m.; 05:30 p.m.] | $2.5 \cdot 20 = 50$ | 0.25 |
| (05:30 p.m.; 08:00 p.m.] | $2.5 \cdot 20 = 50$ | 0.25 |

Therefore, the risk-neutral IT department calculates the expected business costs due to a service outage incident (considering the business costs from Table 8.6 and the probabilities from Table 8.8) for each of the five service incident levels.[20] Table 8.9 shows the results of this calculation.

Table 8.9.: Expected Business Costs due to an Outage Incident Depending on its Duration

| | Outage Duration [Minutes] | | | | |
|---|---|---|---|---|---|
| | (0;30] | (30;50] | (50;60] | (60;70] | (70;80] |
| Expected Business Costs [Euro] | 50 | 144 | 266 | 432 | 602 |

Afterwards, using these expected business cost values, the IT department determines the aggregated business costs (i.e., the business costs induced by a service incident pattern) for each of the service offers $s_A$ and $s_B$. Therefore, it multiplies the number of outage incidents stated in a service incident pattern with the corresponding expected business cost value for each service incident level and, then, sums up the results.[21] Finally, the IT department calculates the total service costs for each of the offered service solutions by adding aggregated business costs and service price. Table 8.10 depicts aggregated business costs, service price and total service costs for both service offers.

Since service offer $s_B$ results in lower total service costs (per month) compared to service offer $s_A$, provider B wins the procurement auction. Hence, the IT department and provider B conclude a service contract based on the conditions stated in the RfP and the service offer $s_B$, which specifies the service incident pattern to be achieved

[20] For instance, the expected business costs due to an outage incident with a duration in the interval (60; 70] is calculated as follows: 0.05·€250+0.40·€809+0.05·€493+0.25·€203+0.25·€82=€432.

[21] For instance, the aggregated business costs of service offer $s_A$ are calculated as follows: 4·€50+2·€144+1·€266+1·€432+1·€602≈€1,790.

Table 8.10.: Aggregated Business Costs, Service Price and Total Service Costs of the Offered Service Solutions

| Service Solution | Aggregated Business Costs [Euro] | Service Price [Euro] | Total Service Costs [Euro] |
|---|---|---|---|
| $s_A$ | 1,790 | 1,400 | 3,190 |
| $s_B$ | 1,383 | 1,600 | 2,983 |

by the provider and the service price to be paid by the customer. The penalty and reward functions which were specified in the RfP (see Table 8.6) are an integral part of this contract.

## 8.4. Summary

In this chapter we demonstrated the applicability of our method for Cost-Optimal Service Selection using the example of a typical order picking process at a warehouse. We took the perspective of an IT department within a mail order company which needs to support business operations through a warehouse management system. We assumed that this warehouse management system is to be hosted, managed, and offered as a business service by an external service provider.

First, we described how the IT department applies our discrete-event simulation-based method (step 1; see Chapter 6) in order to determine business cost functions with regard to service outage incidents. We considered two cash-effective business consequences, namely overtime work and decrease in cash receipts due to shipping delays, and described the definition of service incident scenarios. The simulation results show that the business costs caused by outage incidents differ significantly depending on outage duration and time of incident occurrence. Furthermore, we recognized that business costs develop disproportionally (non-constant marginal business cost functions) in many cases.

Afterwards, we illustrated how the IT department prepares and conducts a multi-attribute procurement auction (step 2; see Chapter 7) with two external service providers regarding the required business service. In particular, we showed how the complexity of a request for proposal can be decreased by conflating time periods, in which business costs due to outage incidents develop similarly, and by reducing the number of service incident levels to be distinguished.

With the following chapter we conclude this thesis. In particular, we summarize its contributions and suggest directions for future research in the field of Service Level Engineering.

# 9. Conclusion

The overall objective of this work was to develop an approach which is suitable for a customer IT department to select the cost-optimal service solution among different provider offers. In particular, we concentrated on four facets of this problem: The description of the decision problem in economic terms, the definition of appropriate service quality measures, the monetary assessment of the business impact induced by specific service quality levels, and the negotiation of service offers with service providers.

We developed the method for Cost-Optimal Service Selection (COSS), which addresses these challenges by explicitly considering the trade-off between service price and service quality with regard to services' incident behavior.

In the next section, we summarize the contributions and findings of this thesis. Afterwards, we derive implications for decision makers of IT departments, business departments and IT providers. Finally, we address limitations of this work and point out questions to be addressed in future research in the field of Service Level Engineering.

## 9.1. Contribution

In order to approach the overall research problem, we defined four concrete research questions in Chapter 1. In the following, we summarize the answers to these questions which are provided by this work and, thereby, illustrate its contribution.

**Research Question 1: How can the decision problem of an IT department be described in economic terms?**

First, we introduced Service Level Engineering as a field of research, which aims to determine business-relevant service quality measures and, based thereon, select cost-optimal service solutions (see Chapter 3). SLE strives to solve a quantitative optimization problem and to choose the particular service offer for which the sum of service delivery costs and monetarily quantified, adverse business impact — resulting from imperfect service quality — reaches its minimum.

Next, in Chapter 4, we described the IT department's decision problem to select the cost-optimal IT service among different provider offers as an SLE optimization problem. We analyzed service quality indicators commonly used in industry with regard to their suitability to describe the business impact that a service causes. We found that the IT department will usually not be able to precisely determine the adverse business impact induced by imperfect service quality if aggregating, incident-based service quality measures are applied. This is due to the fact that these service quality indicators 'merge' information about the incident behavior of services. Consequently, we stated that it will normally be impossible for the IT department to solve its optimization problem in case such service quality measures are used. Our second research question addresses this issue.

**Research Question 2: Which characteristics of a service and of the business environment which it is to support does an IT department have to consider in order to identify the cost-optimal service solution with regard to service incident behavior?**

We argue that imperfect service quality has to be described by service quality measures, which allow the assessment of services' adverse business impact (see Chapter 4). Only thus, the IT department is able to monetarily assess a service's adverse business impact in terms of 'total business costs', to weigh service price against service quality and, finally, to select the cost-optimal service solution.

We found that IT departments have to consider the impact of single service incidents on business operations in order to be able to determine services' adverse business impact. Therefore, we elaborated on the characteristics of service incidents, defined a number of concepts, and formally described the relationship between service incidents and the monetarily quantified, adverse business impact (business costs) these induce (see Section 4.3).

In order to address the shortcomings of current service quality measures, we suggested an advanced form of incident-based, non-aggregating service quality measures: service incident patterns, which represent services' characteristic combination of service incident attribute values and, thus, indicate the business impact a service causes. Furthermore, we recommended service customers to determine business cost functions, which describe the cash-effective adverse business impact a single service incident induces depending on its attribute values and time of occurrence. This challenge is tackled by our third research question.

**Research Question 3: How can IT departments quantify the impact of service incidents on business operations in case of well-defined business processes?**

We addressed this question through the development of a simulation-based method to guide IT and business departments in jointly determining business cost functions in case of well-defined business processes (see Chapter 6). Our method, which builds on the BIA process by Wiedemann (2008), implements the first step of COSS. Our simulation-based method consists of four steps:

First, analysts from IT and business departments identify cash-effective business consequences which are potentially induced by a disruption of business processes. Therefore, they consider classifications of business consequences and corresponding costs from different fields of research, which we found via an extensive literature review and referred to in Chapter 5. Second, the team of analysts links each relevant business consequence to one or several business process-related metrics. That is, they link the occurrence of a service incident with specific properties at a certain time to changes in the values of business process-related metrics. Third, analysts simulate the impact of single service incidents on the values of those business process metrics. Finally, based thereon, they estimate the particular business cost functions for the different types of service incidents to be considered at different time periods.

With our simulation-based method we address a scenario in which one or more business processes depend on a specific service. Using discrete-event simulation, we determine the impact of service incidents on business operations through the adoption and usage of an established, formal technique for business process analysis. Thus, we are able to consider stochastic as well as dynamic aspects of business processes.

Considering research question 2, we suggested service incident patterns as service quality measures that indicate the business impact which a service causes. We also pointed out that service incident patterns need to be stated truthfully by providers to allow the determination of cost-optimal service offers. That is, IT departments need to ensure that providers are not tempted to state service incident patterns in a strategic manner. Research question 4 seeks a strategy which makes sure that providers do not have an incentive to untruthfully state incident patterns in their service offers.

**Research Question 4: Which strategy should an IT department follow in a negotiation with risk-neutral external service providers when service incident patterns are used as quality measures?**

We developed two auction-based, cost-optimal mechanisms to support IT departments in negotiating service offers with multiple providers. More precisely, we suggested procurement auctions as a mechanism to enable IT departments to solve their optimization problem to select cost-optimal service solutions among different service offers. We defined both procurement auctions in a way which ensures that providers disclose their private knowledge about service incident behavior when offering a service solution and do not have an incentive to strategically define service offers.

Initially, we showed that an IT department will not be able to select the cost-optimal service solution among different offers in any case if penalty and reward functions are not identical with business cost functions. Afterwards, we demonstrated that service providers will have an incentive to define service offers strategically if they know that the values of penalty and reward functions do not correspond to those of business cost functions. That is, they might try to pretend that their service solutions are cheaper in terms of total service costs. Furthermore, we could show that this incentive does not exist if penalty and reward functions are identical with business cost functions. Moreover, we found that service providers might not specify service incident patterns at all if they cannot benefit from outperforming these. For these reasons, we suggested the IT department to define penalty and reward functions in a way that their values are identical with those of business cost functions.

Afterwards, we introduced our mechanisms — a multi-attribute and a single-attribute, first-price procurement auction. In the first approach, service providers are invited to offer service solutions by stating a tuple of service incident patterns and a service price. In the second approach, we addressed a scenario in which service providers are asked to state a service price only. In that case they have to assure to compensate the service customer for every service incident that occurs if their service offer is chosen.

Based on the answers to the four research questions, we developed our method for Cost-Optimal Service Selection to solve the IT department's SLE optimization problem. On the one hand, COSS supports the determination of total service costs given a service offer and, on the other hand, provides guidance for IT and business departments on how to jointly assess the adverse impact of imperfect service quality on business operations. Furthermore, COSS offers a method which facilitates and structures negotiations between the IT department and external providers meanwhile enabling the selection of the cost-optimal service solution.

In the following, we provide a number of recommendations for IT departments, business departments and service providers, which we derive from the results obtained in this work.

## 9.2. Managerial Implications

Only by considering the trade-off between services' adverse business impact and service price, IT departments achieve cost-optimal support of business operations. Therefore, in this thesis, we state that IT departments have to monetarily quantify the business impact which is caused by imperfect service quality. IT departments have to work closely with business departments and to support these in determining and assessing the business consequences which may result from service incidents.

This requires that IT and business departments are aware of the dependencies between services and business processes. Consequently, we argue that IT departments have to formally document these interrelations and have, at least, a basic understanding of the business processes they support. A formal definition and description of business processes as well as of their interdependencies (which may, for instance, be modeled in the context of Business Continuity Management activities) facilitate this task. Furthermore, we suggest regular meetings with business departments in order to ensure that IT departments are informed about changes of business processes and resulting differences in service requirements.

In the course of this work, we also pointed out that IT departments must be aware of services' incident behavior in order to be able to select cost-optimal service offers. Therefore, we argue that IT departments ought to apply service incident patterns as an advanced form of service quality measures in service level agreements. In a first step, an IT department could define simple, discrete service incident patterns which consider a few service incident levels only. For instance, it might define a service incident pattern, which only distinguishes between three levels of outage duration.[1]

Moreover, we state that the usage of service incident patterns also enables service providers to considerably better understand and meet customers' business requirements. Consequently, in this work, we also suggest providers to apply service incident patterns in order to achieve a significant advantage over competitors, which are not able to precisely specify service incident behavior. We suggest that service providers either calculate or simulate end-to-end service incident behavior bottom-up (i.e., based on information about service components' incident behavior) or determine it through end-to-end service monitoring.[2] In both cases, they need to be aware of the particular service components forming a service delivery environment in order to allow for the prediction of its characteristic service incident patterns. Therefore, we argue that providers need to standardize service components and service solutions.[3]

## 9.3. Limitations

Having developed our method for Cost-Optimal Service Selection, there are a number of limitations and challenges that need to be mentioned. First, we assumed that providers are able to state service incident patterns in their service offers. However, as noted above, this may be still too advanced for today's practices. Providers need to further standardize service components and service solutions as well as gather and analyze detailed data about service incidents in order to allow for accurate predictions of service incident patterns.[4]

Furthermore, we require the customer company to determine the business costs which it incurs due to the occurrence of service incidents. We are aware that it might be difficult, costly and time-consuming to achieve a precise estimation of business cost functions in practice. Nevertheless, we state that an IT department requires information about business costs in order to make an economically well-founded decision about the service offer to select. We argue that a good estimate of business costs might already significantly reduce total service costs. Our simulation-based method addresses this issue.[5]

[1] That is, the value range of the service incident attribute 'outage duration' is divided into a set of three disjoint intervals.

[2] Section 4.6 illustrates how a provider might proceed in order to determine characteristic service incident patterns of a service solution.

[3] Huntley *et al.* (2014) also argue that providers need to standardize and industrialize offerings in order to be able to react to changing business needs.

[4] Section 4.6 discusses challenges for service providers which emerge from the application of service incident patterns as service quality measures in more detail.

[5] Challenges with regard to the determination of business costs as well as limitations of our simulation-based approach are discussed in Section 6.5 in more detail.

In addition, we assumed that providers will agree to penalty rates which correspond to business costs — as long as these are not 'out of balance' with service prices. This is a reasonable assumption since providers can incorporate all expected penalties into the service price. Furthermore, they can insure themselves against penalty payments due to service incidents which are caused by disasters. If penalty rates are not in a 'fair relation' to the service price, however, providers might not accept this '1:1 rule' due to the high financial risk it represents.[6]

Moreover, for the development of COSS we assumed providers to be risk-neutral. Consequently, we supposed that they add expected penalties to and subtract expected bonus payments from the net service price. This assumption should at least be realistic for large providers which are able to spread and diversify risks across numerous, heterogeneous customers. In contrast, risk-averse or risk-seeking service providers would value potential deviations from service incident patterns in a different way and, thus, consider different surcharges when defining service prices. Furthermore, they might value service solutions' service incident behavior differently and therefore state different service incident patterns. This would hamper the selection of cost-optimal service offers for service customers.

Finally, we only evaluated and demonstrated the applicability of our method for Cost-Optimal Service Selection using a theoretical business setting. Although we addressed a realistic scenario to evaluate COSS,[7] the approach needs to be further tested by tackling scenarios directly taken from practice. The following section suggests directions for further research in the field of Service Level Engineering.

## 9.4. Recommended Further Research

Further research may concentrate on both steps of COSS (i.e., the determination of business cost functions and the negotiation with service providers) as well as on its foundation (i.e., the Service Level Engineering approach and the idea to apply service incident patterns). Moreover, it may address theoretical or methodological as well as empirical research questions.

### Determination of Business Cost Functions

Our simulation-based approach for the determination of business cost functions (see Chapter 6) supports IT and business departments to monetarily assess the business impact of service incidents in case of well-defined business processes.

- As stated before, the development of an accurate simulation model of a complex network of interdependent business processes may be time-consuming. Therefore, future work could elaborate ways of combining our simulation-based method with less precise but also less costly approaches, such as expert interviews and document analysis (e.g., Wiedemann, 2008; Akkiraju *et al.*, 2012).

[6] Further challenges with regard to the negotiation of service offers with external service providers as well as limitations of our auction-based mechanisms are discussed in Section 7.4.

[7] The picking system which we analyzed in Chapter 8 represents an abstraction of real, more extensive picking systems.

- Furthermore, the application of our simulation-based method to different scenarios in practice will raise further research questions, which can be addressed in future work. Case studies could, for instance, focus on other types of service incidents, such as 'reduced response time incidents' and 'reduced throughput incidents' which are specified by two attributes each (i.e., their duration and their degree of reduction). Moreover, case studies could analyze and compare the suitability of different classifications of business consequences (see Chapter 5) for the identification of potential adverse impacts of service incidents.

### Negotiation with Service Providers

Our cost-optimal procurement auction approaches (see Chapter 7) support IT departments in negotiating service offers with multiple providers and selecting the cost-optimal service solution among different service offers.

- We preferred procurement auctions to bilateral 'bargaining' since procurement auctions represent an efficient way to consider service offers by multiple providers at the same time.[8] We noted, however, that service requirements need to be specified more precisely if a procurement auction is conducted since it does not allow for service properties to be discussed.[9] If a business settings requires that various service properties are refined during the last phase of contract negotiation, different approaches might achieve better results. Therefore, with regard to such cases, we suggest the conduct of laboratory experiments (e.g., Bichler, 2000) in order to determine and compare the suitability of different types of procurement auction approaches and bilateral negotiations as well as of combinations of these.

- Moreover, models considering risk-sharing between service customer and provider could be developed and analyzed in future work (see e.g., Satzger & Kieninger, 2011). An IT department and a service provider might, for instance, specify a maximum penalty to be charged if a specific attribute value of a service incident exceeds a certain threshold.[10]

### Service Level Engineering and Service Incident Patterns

Finally, we recommend further research with respect to the determination of services' characteristic incident patterns, which should be of particular interest for service providers.

---

8 In contrast, bilateral bargaining with the same number of providers might turn out to be more time-consuming.

9 In this work, we assumed that all general properties of the service required — except for service quality objectives and service price — are specified according to business needs.

10 An IT department could, for example, define that a service provider has to pay a penalty of €1.000.000 (maximum value) for all service incidents with a duration of more than 60 minutes (threshold) only. On the one hand, this would lower the service price, since providers would not incorporate higher penalty payments into the service price when defining service offers. On the other hand, however, the IT department would take a higher financial risk and would not be able to select the cost-optimal service solution in any case.

- First, we suggest to examine, which attributes of the service components that form a specific service delivery environment (e.g., the hardware and the operating system used) have influence on service incident behavior (see e.g., Bogojeska *et al.*, 2013 and Bogojeska *et al.*, 2014). Based on this knowledge, it will be possible to predict and control the future service incident behavior of service solutions which comprise similar service components.
- Furthermore, we propose to analyze the influence of the workload on business processes on the incident behavior of a service which supports these particular business processes. Similarly, correlations between the utilization of a service delivery environment's hardware resources (e.g., memory utilization) and its service incident behavior could be examined (ibid.). The implementation of such analyses also requires the identification or development of appropriate methods.

To summarize, following a Service Level Engineering approach, in this work we developed COSS, a method to support IT departments in selecting cost-optimal service solutions among different provider offers. In order to solve this optimization problem, COSS leverages discrete-event simulation to monetarily assess the adverse business impact of service incidents on business operations and applies auction theory to facilitate and structure the contract negotiation process between the IT department and potential service providers. COSS is based on the concept of 'service incident patterns', an advanced form of service quality measures suggested by this thesis, which allows for a precise assessment of the business impact that results from imperfect service quality.

We are convinced that this work provides valuable insights for IT departments and service providers with respect to the challenge of selecting or offering services in a new way — by exploiting the results of business impact analyses and the determination of services' incident behavior. At the same time, the recommended directions for further research also constitute challenges which need to be addressed in order to further increase the effectiveness of Service Level Engineering in practice.

# Bibliography


Addy, Rob. 2007. *Effective IT Service Management: To ITIL and Beyond!* Berlin et al.: Springer.

Akkiraju, Rama, Bhattacharjya, Debarun, Plaskon, Joseph, & Jennings, Derek. 2011. Towards Effective Business Process Availability Management. *Pages 242–251 of: Proceedings of the 1st Global Conference of the Service Research Innovation Initiative (SRII).* San Jose, California, USA: IEEE.

Akkiraju, Rama, Bhattacharjya, Debarun, & Gupta, Sammukh. 2012. Towards Effective Business Process Availability Management. *Journal of Service Science Research*, **4**(2), 319–351.

Ambrose, Christopher, & Huntley, Helen. 2012. *Outsourcing Trends, 2011-2012: Get More From Vendor Management.* Tech. rept. G00226754. Gartner.

AnyLogic. 2015. *Multimethod Simulation Software.* http://www.anylogic.com/, accessed on 02/16/2015.

Asker, John, & Cantillon, Estelle. 2010. Procurement When Price and Quality Matter. *The RAND Journal of Economics*, **41**(1), 1–34.

Atkinson, Hawley, Hamburg, John, & Ittner, Christopher (eds). 1994. Standard Cost-of-Poor-Quality Element Listing. *Chap. Appendix B, pages 343–392 of:* Atkinson, Hawley, Hamburg, John, & Ittner, Christopher (eds), *Linking Quality to Profits - Quality-Based Cost Management.* Milwaukee: ASQC Quality Press.

Bajari, Patrick, McMillan, Robert, & Tadelis, Steven. 2008. Auctions Versus Negotiations in Procurement: An Empirical Analysis. *Journal of Law, Economics, and Organization*, **25**(2), 372–399.

Baker, Harold, & English, Philip (eds). 2011. *Capital Budgeting Valuation: Financial Analysis for Today's Investment Projects.* Hoboken: Wiley.

Balaouras, Stephanie, Schreck, Galen, & Batiancila, Rachel. 2008. *Building the Business Case for Disaster Recovery Spending.* Tech. rept. Forrester Research.

Bamberg, Günter, Coenenberg, Adolf, & Krapp, Michael. 2012. *Betriebswirtschaftliche Entscheidungslehre.* 15th edn. Munich: Vahlen.

Banks, Jerry. 1996. *Discrete Event System Simulation*. 2nd edn. Upper Saddle River: Prentice Hall.

Banks, Jerry, Carson II, John, Nelson, Barry, & Nicol, David. 2010. *Discrete-Event System Simulation*. 5th, int. edn. Upper Saddle River et al.: Pearson.

Barnes, Peter. 2011. Business Impact Analysis. *Chap. 12.1, pages 166–182 of:* Hiles, Andrew (ed), *The Definitive Handbook of Business Continuity Management*, 3rd edn. Chichester: Wiley.

Bartholdi, John, & Hackman, Steven. 2014. *Warehouse & Distribution Science*. Release 0.96. Atlanta: Georgia Institute of Technology.

Bartsch, Christian. 2010. *Modellierung und Simulation von IT-Dienstleistungsprozessen*. Ph.D. thesis, Karlsruhe Institute of Technology.

Basel Committee on Banking Supervision. 2006. *Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework*. compr. edn. http://www.bis.org/publ/bcbs128.pdf, accessed on 11/29/2014.

Baumol, William, & Blinder, Alan. 2012. *Economics: Principles and Policy*. 12th edn. Mason: South-Western.

Benyon, Robert, & Johnston, Robert. 2006. *Service Agreements: A Management Guide*. Zaltbommel: van Haren Publishing.

Berger, Thomas. 2007. *Konzeption und Management von Service-Level-Agreements für IT-Dienstleistungen*. Saarbrücken: Verlag Dr. Müller.

Berk, Jonathan, & DeMarzo, Peter. 2011. *Corporate Finance*. 2nd edn. Boston et al.: Pearson.

Berlet, Bruno, & Riese, Adam. 1860. *Die Coß von Adam Riese*. Annaberg: Hasper.

Berninghaus, Siegfried, Ehrhart, Karl-Martin, & Güth, Werner. 2010. Auktionstheorie. *Chap. 5, pages 231–278 of:* Berninghaus, Siegfried, Ehrhart, Karl-Martin, & Güth, Werner (eds), *Strategische Spiele: Eine Einführung in die Spieltheorie*, 3rd edn. Berlin & Heidelberg: Springer.

Bichler, Martin. 2000. An Experimental Analysis of Multi-Attribute Auctions. *Decision Support Systems*, **29**(3), 249–268.

Birta, Louis, & Arbez, Gilbert. 2013. *Modelling and Simulation: Exploring Dynamic System Behaviour*. 2nd edn. Simulation Foundations, Methods and Applications. London et al.: Springer.

Blau, Benjamin. 2009. *Coordination in Service Value Networks: A Mechanism Design Approach*. Ph.D. thesis, Universität Karlsruhe, Germany.

Blau, Benjamin, Block, Carsten, & Stößer, Jochen. 2008. How to Trade Electronic Services?: Current Status and Open Questions. *Pages 1–9 of: Proceedings of the Joint Conference of the INFORMS Section on Group Decision and Negotiation, the EURO Working Group on Decision and Negotiation Support, and the EURO Working Group on Decision Support Systems.* Coimbra, Portugal: INFORMS.

Blau, Benjamin, Dinther, Clemens, Conte, Tobias, Xu, Yongchun, & Weinhardt, Christof. 2009. How to Coordinate Value Generation in Service Networks. *Business & Information Systems Engineering*, **1**(5), 343–356.

Bogojeska, Jasmina, Lanyi, David, Giurgiu, Ioana, Stark, George, & Wiesmann, Dorothea. 2013. Classifying Server Behavior and Predicting Impact of Modernization Actions. *Pages 59–66 of: Proceedings of the 9th International Conference on Network and Service Management.* Zurich, Switzerland: IEEE.

Bogojeska, Jasmina, Giurgiu, Ioana, Lanyi, David, Stark, George, & Wiesmann, Dorothea. 2014. Impact of HW and OS Type and Currency on Server Availability Derived from Problem Ticket Analysis. *Pages 1–9 of: Proceedings of the Network Operations and Management Symposium.* Krakow, Poland: IEEE.

Bol, Georg. 2003. *Induktive Statistik.* 3rd edn. Munich & Vienna: Oldenbourg.

Bolancé, Catalina, Guillén, Montserrat, Gustafsson, Jim, & Nielson, Jens. 2012. *Quantitative Operational Risk Models.* online edn. Boca Raton et al.: CRC Press.

Bräutigam, Peter. 2009. Rechtliche Vorbereitung, Vertragsverhandlung und Vertragsmanagement. *Chap. 12, pages 753–816 of:* Bräutigam, Peter (ed), *IT-Outsourcing: Eine Darstellung aus rechtlicher, technischer, wirtschaftlicher und vertraglicher Sicht*, 2nd edn. Berlin: Schmidt.

Breitgand, David, Henis, Ealan, Shehory, Onn, & Lake, John. 2007. Derivation of Response Time Service Level Objectives for Business Services. *Pages 29–38 of: Proceedings of the 2nd International Workshop on Business-Driven IT Management.* Munich,Germany: IEEE.

British Standard Institution. 1990. Guide to the Economics of Quality: Part 2: Prevention, Appraisal and Failure Model. *BS 6143-2:1990.*

Brooks, Peter. 2006. *Metrics for IT Service Management.* 1st edn. Zaltbommel: van Haren Publishing.

Buco, Melissa, Chang, Rong, Luan, Laura, Ward, Christopher, Wolf, Joel, & Yu, Philip. 2004. Utility Computing SLA Management Based upon Business Objectives. *IBM Systems Journal*, **43**(1), 159–178.

Bullinger, Hans-Jörg, Fähnrich, Klaus-Peter, & Meiren, Thomas. 2003. Service Engineering: Methodical Development of New Service Products. *International Journal of Production Economics*, **85**(3), 275–287.

Burr, Wolfgang. 2003. Service-Level-Agreements: Arten, Funktionen und strategische Bedeutung. *Chap. 2, pages 33–46 of:* Bernhard, Martin, Mann, Hartmut, Lewandowski, Winfried, & Schrey, Joachim (eds), *Praxishandbuch Service-Level-Management: Die IT als Dienstleistung organisieren*, 1st edn. Düsseldorf: Symposion.

Business Continuity Institute. 2011. *Dictionary of Business Continuity Management Terms*. http://www.thebci.org/glossary.pdf, accessed on 07/03/2015.

Campanella, Jack (ed). 1999. *Principles of Quality Costs: Principles, Implementation and Use*. 3rd edn. Milwaukee: American Society for Quality.

Cardoso, Jorge, Schmitz, Björn, & Kieninger, Axel. 2016. Service Research. *Chap. 10, pages 343–377 of:* Cardoso, Jorge, Fromm, Hansjörg, Nickel, Stefan, Satzger, Gerhard, Studer, Rudi, & Weinhardt, Christof (eds), *Fundamentals of Service Systems*. Heidelberg et al.: Springer.

Cassady, Ralph. 1967. *Auctions and Auctioneering*. Berkeley et al.: University of California Press.

Che, Yeon-Koo. 1993. Design Competition through Multidimensional Auctions. *RAND Journal of Economics*, **24**(4), 668–680.

Cheng, Feng, Gamarnik, David, Jengte, Nitin, Min, Wanli, & Ramachandran, Bala. 2005. *Modeling Operational Risk in Business Processes*. Tech. rept. IBM Research, Yorktown Heights, New York, USA.

Chung, Lawrence. 1991. Representation and Utilization of Non-Functional Requirements for Information System Design. *Pages 5–30 of:* Andersen, Rudolf, Bubenko, Janis, & Sølvberg, Arne (eds), *Proceedings of the 3rd International Conference on Advanced Information Systems Engineering*. Trondheim, Norway: Springer.

Cornish, Malcolm. 2011. Business Continuity Management Methodology. *Chap. 9, pages 121–135 of:* Hiles, Andrew (ed), *The Definitive Handbook of Business Continuity Management*, 3rd edn. Chichester: Wiley.

Cruz, Marcelo. 2002. *Modeling, Measuring and Hedging Operational Risk*. Chichester: Wiley.

Dale, Barrie, & Plunkett, James. 1995. *Quality Costing*. 2nd edn. London et al.: Chapman & Hall.

Darmawan, Budi, Gummadavelli, Ravinder, Kuppusamy, Sathyabama, Tan, Celena, Rintoul, David, Anglin, Howard, Chuan, Huang, Subhedar, Aleem, Firtiyan, Andri, Nambiar, Pradeep, Lall, Prem, Dhall, Rohit, & Pires, Ronaldo. 2008. *IBM Tivoli Composite Application Manager Family: Installation, Configuration, and Basic Usage*. Tech. rept. IBM.

Dhillon, Balbir. 2010. *Life Cycle Costing for Engineers*. Boca Raton et al.: CRC Press.

Dibbern, Jens, Goles, Tim, Hirschheim, Rudy, & Jayatilaka, Bandula. 2004. Information Systems Outsourcing: A Survey and Analysis of the Literature. *The Data Base for Advances in Information Systems*, **35**(4), 6–102.

Dowd, Victor. 2003. Measurement of Operational Risk: The Basel Approach. *Chap. 3, pages 31–48 of:* Alexander, Carol (ed), *Operational Risk - Regulation, Analysis and Management*. London et al.: Financial Times Prentice Hall.

Duden Online. 2014. *Duden Online*. http://www.duden.de/suchen/dudenonline/social_costs, accessed on 09/12/2014.

Dumas, Marlon, La Rosa, Marcello, Mendling, Jan, & Reijers, Hajo. 2013. *Fundamentals of Business Process Management*. Berlin & Heidelberg: Springer.

Eichstädt, Tilman. 2008. *Einsatz von Auktionen im Beschaffungsmanagement: Erfahrungen aus der Einkaufspraxis und die Verbreitung auktionstheoretischer Konzepte*. Wiesbaden: Gabler.

Energy Economics. 2014. *Energy Economics: Aims and Scope*. http://www.journals.elsevier.com/energy-economics, accessed on 07/24/2014.

Engelhardt, Werner, Kleinaltenkamp, Michael, & Reckenfelderbäumer, Martin. 1993. Leistungsbündel als Absatzobjekte: Ein Ansatz zur Überwindung der Dichotomie von Sach- und Dienstleistungen. *Zeitschrift für betriebswirtschaftliche Forschung*, **45**(5).

European Committee for Standardization. 2010. Maintenance: Maintenance Terminology. *DIN EN 13306:2010-12*.

Fabozzi, Frank, & Peterson-Drake, Pamela. 2010. *The Basics of Finance: An Introduction to Financial Markets, Business Finance, and Portfolio Management*. online edn. Hoboken: Wiley.

Fox, James, Brammall, Jadd, & Yarlagadda, Prasad. 2008. Determination of the Financial Impact of Machine Downtime on the Post Large Letters Sorting Process. *Journal of Achievements in Materials and Manufacturing Engineering*, **31**(2), 732–738.

Franke, Ulrik. 2012. Optimal IT Service Availability: Shorter Outages, or Fewer? *IEEE Transactions on Network and Service Management*, **9**(1), 22–33.

Fromm, Hansjörg, Habryn, François, & Satzger, Gerhard. 2012. Service Analytics: Leveraging Data Across Enterprise Boundaries for Competitive Advantage. *Pages 139–149 of:* Bäumer, Ulrich, Kreutter, Peter, & Messner, Wolfgang (eds), *Globalization of Professional Services: Innovative Strategies, Successful Processes, Inspired Talent Management, and First-Hand Experiences*. Berlin, Heidelberg: Springer.

Gadrey, Jean. 2000. The Characterization of Goods and Services: An Alternative Approach. *Review of Income and Wealth*, **46**(3), 369–387.

Gaver, Donald. 1962. A Waiting Line with Interrupted Service, Including Priorities. *Journal of the Royal Statistical Society*, **24**(1), 73–90.

German Institute for Standardization. 1998. Service Engineering: Entwicklungsbegleitende Normung für Dienstleistungen. *DIN-Fachbericht 75:1998*.

Gewald, Heiko, & Helbig, Kay. 2006. A Governance Model for Managing Outsourcing Partnerships: A View from Practice. *In: Proceedings of the 39th Annual Hawaii International Conference on System Sciences*. Koloa, Kauai, Hawaii, USA: IEEE.

Glinz, Martin. 2007. On Non-Functional Requirements. *Pages 21–26 of:* Sutcliffe, Alistair, & Jalote, Pankaj (eds), *Proceedings of the 15th International Requirements Engineering Conference*. New Delhi, India: IEEE.

Goldberg, Marius, & Satzger, Gerhard. 2015. Designing Organizational Models for Service Integration in a Multi-Sourcing Context. *In: Proceedings of the 21st Americas Conference on Information Systems (Paper 12)*.

Goldberg, Marius, Kieninger, Axel, Satzger, Gerhard, & Fromm, Hansjörg. 2014. Transition and Delivery Challenges of Retained Organizations in IT Outsourcing. *Pages 56–71 of:* Léonard, Michel, & Snene, Mehdi (eds), *Proceedings of the 5th International Conference on Exploring Services Science*. Geneva, Switzerland: Springer.

Goldberg, Marius, Satzger, Gerhard, & Kieninger, Axel. 2015a. A Capability Framework for IT Service Integration and Management in Multi-Sourcing. *In: Proceedings of the 23rd European Conference on Information Systems (Paper 58)*.

Goldberg, Marius, Kieninger, Axel, Satzger, Gerhard, & Fromm, Hansjörg. 2015b. Retained Organizations in IT Outsourcing: Linking Organization Design to Outsourcing Management Problems. *KSRI Discussion Paper*, under review.

Gonzales, Reyes, Gasco, Jose, & Llopis, Juan. 2006. Information Systems Outsourcing: A Literature Analysis. *Information & Management*, **43**(7), 821–834.

Grigoryev, Ilya. 2012. *AnyLogic 6 in Three Days : A Quick Course in Simulation Modeling*. Hampton: AnyLogic North America.

Gryna, Frank. 1999. Quality and Costs. *Chap. 8, pages 8.1–8.26 of:* Juran, Joseph, & Godfrey, Blanton (eds), *Juran's Quality Handbook*, 5th edn. New York et al.: McGraw-Hill.

Hale, Kathryn, Young, Allie, Lam, Lai-ling, Heng, Jacqueline, Tornbohm, Cathy, Roster, Jeffrey, Ng, Fred, Tramacere, Gianluca, Blackmore, Dean, Huntley, Helen, Matson, Susanne, Lo, Twiggy, Tang, Tina, Singh, TJ, Roy, Arup, Tenneson, Christine, Tan, Susan, Sawai, Misako, Petri, Gregor, Yeates, Morgan, Toombs, Douglas, & Kandaswamy, Rajesh. 2015. *Forecast: IT Services, Worldwide, 2013-2019, 1Q15 Update*. https://www.gartner.com/doc/3006724, accessed on 04/29/2015.

Hiles, Andrew. 2002. *The Complete Guide to IT Service Level Agreements: Aligning IT Services to Business Needs*. 3rd edn. Brookfield: Rothstein Associates Inc.

Hiles, Andrew. 2011. Business Impact Analysis: Building a Better Mousetrap. *Chap. 12.2, pages 183–205 of:* Hiles, Andrew (ed), *The Definitive Handbook of Business Continuity Management*, 3rd edn. Chichester: Wiley.

Hill, Peter. 1977. On Goods and Services. *Review of Income and Wealth*, **23**(4), 315–338.

Hlupic, Vlatka, & Robinson, Stewart. 1998. Business Process Modelling and Analysis Using Discrete-Event Simulation. *Pages 1363–1370 of: Proceedings of the 30th Winter Simulation Conference*. Los Alamitos, California, USA: IEEE.

Hunnebeck, Lou, Rudd, Colin, Lacy, Shirley, Hanna, Ashley, Cannon, David, Lloyd, Vernon, Orr, Anthony, Rance, Stuart, Steinberg, Randy, & Wheeldon, David. 2013. *ITIL Service Design: Ausgabe 2011*. 2nd edn. London: The Stationery Office.

Huntley, Helen, Young, Allie, Tramacere, Gianluca, Blackmore, Dean, Petri, Gregor, Anderson, Ed, & Toombs, Douglas. 2014. *Forecast Overview: IT Outsourcing, 2014*. https://www.gartner.com/doc/2815020, accessed on 06/27/2015.

Huston, Chad. 2013. *IDC's Worldwide Services Taxonomy, 2013: Industry Development and Models*. Tech. rept. International Data Corporation.

International Journal of Production Economics. 2014. *Aims and Scope*. http://www.journals.elsevier.com/international-journal-of-production-economics, accessed on 07/24/2014.

International Organization for Standardization. 2005. Quality Management Systems: Fundamentals and Vocabulary. *ISO 9000:2005-12*.

International Organization for Standardization. 2009. Risk management: Principles and Guidelines. *ISO 31000:2009*.

International Organization for Standardization. 2011. Information Technology: Service Management: Part 1: Service Management System Requirements. *ISO/IEC 20000-1:2011*.

International Organization for Standardization. 2012a. Societal Security: Business Continuity Management Systems: Requirements. *ISO 22301:2012*.

International Organization for Standardization. 2012b. Societal Security: Terminology. *ISO 22300:2012*.

Jacob, Bart, Adhia, Bhavesh, Badr, Karim, Huang, Qing, Lawrence, Carol, Marino, Martin, & Unglaub-Lloyd, Petra. 2008. *IBM Tivoli Application Dependency Discovery Manager: Capabilities and Best Practices*. Tech. rept. IBM.

Jakoubi, Stefan, Tjoa, Simon, Goluch, Sigrun, & Kitzler, Gerhard. 2010. A Formal Approach Towards Risk-Aware Service Level Analysis and Planning. *Pages 180–187 of: Proceedings of the 5th International Conference on Availability, Reliability and Security*. Krakow, Poland: IEEE.

Karten, Naomi. 1997. *How to Establish Service Level Agreements*. 1st edn. Randolph: Karten Associates.

Kelton, David, Sadowski, Randall, & Sturrock, David. 2004. *Simulation with Arena*. 3rd edn. Boston: McGraw-Hill.

Kenett, Ron, & Raanan, Yossi. 2011. *Operational Risk Management: A Practical Approach to Intelligent Data Analysis*. online edn. Chichester et al.: John Wiley & Sons.

Kern, Thomas, & Willcocks, Leslie. 2002. Exploring Relationships in Information Technology Outsourcing: The Interaction Approach. *European Journal of Information Systems*, **11**(1), 3–19.

Kieninger, Axel, Westernhagen, Jens, & Satzger, Gerhard. 2011a. The Economics of Service Level Engineering. *In: Proceedings of the 44th Hawaii International Conference on System Sciences*. Koloa, Kauai, Hawaii, USA: IEEE.

Kieninger, Axel, Baltadzhiev, Dian, Schmitz, Björn, & Satzger, Gerhard. 2011b. Towards Service Level Engineering for IT Services: Defining IT Services from a Line of Business Perspective. *In: Proceedings of the 1st Global Conference of the Service Research Innovation Initiative*. San Jose, California, USA: IEEE.

Kieninger, Axel, Satzger, Gerhard, Straeten, Detlef, Schmitz, Björn, & Baltadzhiev, Dian. 2012a. Business Cost Budgets: A Methodology to Incorporate Business Impact into Service Level Agreements. *International Journal of Service Science, Management, Engineering, and Technology*, **3**(3), 49–64.

Kieninger, Axel, Schmitz, Björn, Straeten, Detlef, & Satzger, Gerhard. 2012b. Incorporating Business Impact into Service Offers: A Procedure to Select Cost-Optimal Service Contracts. *In: Proceedings of the 18th Americas Conference on Information Systems (Paper 19)*.

Kieninger, Axel, Straeten, Detlef, Kimbrough, Steven, Schmitz, Björn, & Satzger, Gerhard. 2013a. Leveraging Service Incident Analytics to Determine Cost-Optimal Service Offers (Nominated for Best Paper Award). *Pages 1015–1029 of:* Alt, Rainer, & Franczyk, Bogdan (eds), *Proceedings of the 11th International Conference on Wirtschaftsinformatik*.

Kieninger, Axel, Berghoff, Florian, Fromm, Hansjörg, & Satzger, Gerhard. 2013b. Simulation-based Quantification of Business Impacts Caused by Service Incidents. *Pages 170–185 of:* e Cunha, João Falcão, Snene, Mehdi, & Novoa, Henriqueta (eds), *Proceedings of the 4th International Conference on Exploring Services Science*, vol. 143. Porto, Portugal: Springer.

Kinney, Michael, & Raiborn, Cecily. 2012. *Cost Accounting: Foundations and Evolutions*. 9th edn. Mason: South-Western.

Klemperer, Paul. 2004. A Survey of Auction Theory. *Chap. 1, pages 9–61 of:* Klemperer, Paul (ed), *Auctions: Theory and Practice*, 1st edn. Princeton et al.: Princeton University Press.

Krishna, Vijay. 2010. *Auction Theory*. 2nd edn. Amsterdam et al.: Academic Press.

Küchler, Peter. 2004. IT-Outsourcing: Eine Darstellung aus rechtlicher, technischer, wirtschaftlicher und vertraglicher Sicht. *Chap. 1, pages 51–160 of:* Bräutigam, Peter (ed), *IT-Outsourcing: Eine Darstellung aus rechtlicher, technischer, wirtschaftlicher und vertraglicher Sicht*, 1st edn. Berlin: Schmidt.

Kütz, Martin. 2011. *Kennzahlen in der IT: Werkzeuge für Controlling und Management*. 4th edn. Heidelberg: Dpunkt.

Lacity, Mary, Khan, Shaji, & Willcocks, Leslie. 2009. A Review of the IT Outsourcing Literature: Insights for Practice. *Journal of Strategic Information Systems*, **18**(3), 130–148.

Laux, Helmut, Gillenkirch, Robert, & Schenk-Mathes, Heike. 2014. *Entscheidungstheorie*. 9th edn. Berlin & Heidelberg: Springer Gabler.

Law, Averill. 2011. *Simulation Modeling and Analysis*. 4th, inter edn. Boston et al.: McGraw-Hill.

Leemis, Lawrence, & Park, Steve. 2006. *Discrete-Event Simulation: A First Course*. 1st edn. Upper Saddle River: Prentice Hall.

Levitt, Joel. 1997. *The Handbook of Maintenance Management*. New York: Industrial Press.

Lewis, Lundy. 1999. *Service Level Management for Enterprise Networks*. 1st edn. Norwood: Artech House.

Luftman, Jerry. 2014. *Influential IT Management Trends: An International Study*.

Luftman, Jerry, & Ben-Zvi, Tal. 2010. Key Issues for IT Executives 2010: Judicious IT Investments Continue Post-Recession. *MIS Quarterly Executive*, **9**(4), 263–273.

Luftman, Jerry, & Derksen, Barry. 2012. Key Issues for IT Executives 2012: Doing More with Less. *MIS Quarterly Executive*, **11**(4), 207–218.

Luftman, Jerry, Zadeh, Hossein S, Derksen, Barry, Santana, Martin, Rigoni, Eduardo Henrique, & Huang, Zhengwei (David). 2012. Key Information Technology and Management Issues 2011-2012: An International Study. *Journal of Information Technology*, **27**(3), 198–212.

Marques, Filipe, Sauvé, Jacques, & Moura, Antão. 2007. Service Level Agreement Design and Service Provisioning for Outsourced Services. *Pages 106–113 of: Proceedings of the Latin American Network Operations and Management Symposium*. Petrópolis, Brazil: IEEE.

Marques, Filipe, Sauvé, Jacques, & Moura, Antão. 2009. SLA Design and Service Provisioning for Outsourced Services. *Journal of Network and Systems Management*, **17**(1-2), 73–90.

Mason, Paul. 2002. *A New Culture for Service-Level Management: Business Impact Management*. Tech. rept. International Data Corporation, Framingham, USA.

McAfee, Robert, & McMillan, John. 1987. Auctions and Bidding. *Journal of Economic Literature*, **25**(2), 699–738.

Merriam-Webster Online Dictionary. 2014. *Merriam-Webster Online Dictionary*. http://www.merriam-webster.com/dictionary, accessed on 07/05/2015.

Meyer, Konrad, & Eul, Marcus. 2013. IT-Outsourcing-Trends: Und wie Unternehmen erfolgreiche Partnerschaften gestalten. *Pages 71–89 of:* Rickmann, Hagen, Diefenbach, Stefan, & Brüning, Kai T. (eds), *IT-Outsourcing*. Berlin et al.: Springer.

Mora, Manuel, Raisinghani, Mahesh, O'Connor, Rory V., Gomez, Jorge Marx, & Gelman, Ovsei. 2014. An Extensive Review of IT Service Design in Seven International ITSM Processes Frameworks: Part I. *International Journal of Information Technologies and Systems Approach*, **7**(2), 83–107.

Moura, Antão, Sauvé, Jacques, Jornada, João, & Radziuk, Eduardo. 2006. A Quantitative Approach to IT Investment Allocation to Improve Business Results. *Pages 87–95 of: Proceedings of the 7th International Workshop on Policies for Distributed Systems and Networks*. London: IEEE Computer Society.

Moura, Antão, Sauvé, Jacques, & Bartolini, Claudio. 2007. Research Challenges of Business-Driven IT Management. *Pages 19–28 of:* Bartolini, Claudio, Sahai, Akhil, & Sauvé, Jacques (eds), *Proceedings of the 2nd International Workshop on Business-Driven IT Management*. Munich,Germany: IEEE.

Moura, Antão, Sauvé, Jacques, & Bartolini, Claudio. 2008. Business-Driven IT Management: Upping the Ante of IT - Exploring the Linkage between IT and Business to Improve Both IT and Business Results. *IEEE Communications Magazine*, **46**(10), 148–153.

Niessen, Jan, & Oldenburg, Paul. 1997. *Service Level Management: Customer Focused*. London et al.: The Stationery Office.

O'Sullivan, Justin. 2006. *Towards a Precise Understanding of Service Properties*. Ph.D. thesis, Queensland University of Technology.

Pang, Guodong, & Whitt, Ward. 2009. Service Interruptions in Large-Scale Service Systems. *Management Science*, **55**(9), 1499–1512.

Pantry, Sheila, & Griffiths, Peter. 1997. *The Complete Guide to Preparing and Installing Service Level Agreements*. 1st edn. London: Library Association Publishing.

Paul, Ray, Hlupic, Vlatka, & Giaglis, George. 1998. Simulation Modelling of Business Processes. *Pages 311–320 of: Proceedings of the 3rd UK Academy of Information Systems Conference*. Lincoln, United Kingdom: McGraw-Hill.

Putnoki, Hans, Schwadorf, Heike, & Then Bergh, Friedrich. 2011. *Investition und Finanzierung*. Munich: Vahlen.

Rana, Omer, Warnier, Martijn, Quillinan, Thomas, & Brazier, Frances. 2008. Monitoring and Reputation Mechanisms for Service Level Agreements. *Pages 125–139 of:* Altmann, Jörn, Neumann, Dirk, & Fahringer, Thomas (eds), *Proceedings of the 5th International Workshop on Grid Economics and Business Models*. Lecture Notes in Computer Science, vol. 5206. Berlin et al.: Springer.

Rasmussen, Svend. 2013. *Production Economics: The Basic Theory of Production Optimisation*. 2nd edn. Springer Texts in Business and Economics. Berlin and Heidelberg: Springer.

Rosenkranz, Friedrich. 2006. *Geschäftsprozesse: Modell- und computergestützte Planung*. 2nd edn. Berlin et al.: Springer.

Salmela, Hannu. 2007. Analysing Business Losses Caused by Information Systems Risk: A Business Process Analysis Approach. *Journal of Information Technology*, **23**(3), 185–202.

Satzger, Gerhard, & Kieninger, Axel. 2011. Risk-Reward Sharing in IT Service Contracts: A Service System View. *In: Proceedings of the 44th Hawaii International Conference on System Sciences*. Koloa, Kauai, Hawaii, USA: IEEE.

Sauvé, Jacques, Marques, Filipe, Moura, Antão, Sampaio, Marcus, Jornada, João, & Radziuk, Eduardo. 2005. SLA Design from a Business Perspective. *Pages 72–83 of: Proceedings of the 16th International Workshop on Distributed Systems - Operations and Management*. Barcelona, Spain: Springer.

Sauvé, Jacques, Moura, Antão, Sampaio, Marcus, Jornada, João, & Radziuk, Eduardo. 2006a. An Introductory Overview and Survey of Business-Driven IT Management. *Pages 1–10 of:* Bartolini, Claudio, Sahai, Akhil, & Sauvé, Jacques (eds), *Proceedings of the 1st International Workshop on Business-Driven IT Management*. Vancouver, Canada: IEEE.

Sauvé, Jacques, Bartolini, Claudio, & Moura, Antão. 2009. Looking at business through a keyhole. *Pages 48–51 of: Proceedings of the International Symposium on Integrated Network Management-Workshops*. New York, New York, USA: IEEE.

Sauvé, Jaques, Marques, Filipe, Moura, Antão, Sampaio, Marcus, Jornada, João, & Radziuk, Eduardo. 2006b. Optimal Design of E-Commerce Site Infrastructure from a Business Perspective. *In: Proceedings of the 39th Hawaii International Conference on System Sciences*. Koloa, Kauai, Hawaii, USA: IEEE.

Schmitz, Björn, Kieninger, Axel, Satzger, Gerhard, & Fromm, Hansjörg. 2014. Towards the Consideration of Performance Risks for the Design of Service Offers. *Pages 108–123 of:* Léonard, Michel, & Snene, Mehdi (eds), *Proceedings of the 5th International Conference on Exploring Services Science*. Geneva, Switzerland: Springer.

Setzer, Thomas, Bhattacharya, Kamal, & Ludwig, Heiko. 2010. Change Scheduling based on Business Impact Analysis of Change-Related Risk. *IEEE Transactions on Network and Service Management*, **7**(1), 58–71.

Sickles, Robin. 2008. Energy Economics. *Pages 842–848 of:* Durlauf, Steven, & Blume, Lawrence (eds), *The New Palgrave Dictionary of Economics*, 2nd edn. Basingstoke: Palgrave Macmillan.

Sommerville, Ian. 2007. *Software Engineering*. 8th edn. Harlow et al.: Addison-Wesley.

Stephens, Matthew. 2004. *Productivity and Reliability-Based Maintenance Management*. Upper Saddle River: Prentice Hall.

Stiglitz, Joseph. 2003. Information and the Change in the Paradigm in Economics: Part 1. *The American Economist*, **47**(2), 6–26.

Sturm, Rick, Morris, Wayne, & Jander, Mary. 2000. *Foundations of Service Level Management*. Indianapolis: SAMS.

Suh, Bomil, & Han, Ingoo. 2003. The IS Risk Analysis Based on a Business Model. *Information & Management*, **41**(2), 149–158.

Sullivan, Michael, Vardell, Terry, & Johnson, Mark. 1996. Power Interruption Costs to Industrial and Commercial Consumers of Electricity. *Pages 23–35 of: Proceedings of the Industrial and Commercial Power Systems Technical Conference*, vol. 33. New Orleans, Louisiana, USA: IEEE.

Tattam, David. 2011. *A Short Guide to Operational Risk*. online edn. Farnham & Burlington: Gower.

Taylor, Richard, & Tofts, Chris. 2005. *Death by a Thousand SLAs: A Short Study of Commercial Suicide Pacts*. Tech. rept. Hewlett-Packard Laboratories, Bristol, United Kingdom.

Taylor, Sharon, Lloyd, Vernon, Rudd, Colin, Bronkhorst, Jereon, Cannon, David, Case, Gary, Hannah, Ashley, Iqbal, Majid, Lacy, Shirley, Macfarlane, Ivor, Nieves, Michael, Rance, Stuart, Spalding, George, & Wheeldon, David. 2007a. *ITIL Service Design*. London: The Stationery Office.

Taylor, Sharon, Cannon, David, Wheeldon, David, Bronkhorst, Jereon, Case, Gary, Hannah, Ashley, Iqbal, Majid, Lacy, Shirley, Lloyd, Vernon, Macfarlane, Ivor, Nieves, Michael, Rance, Stuart, Rudd, Colin, & Spalding, George. 2007b. *ITIL Service Operation*. London: The Stationery Office.

Taylor, Sharon, Lacy, Shirley, Macfarlane, Ivor, Bronkhorst, Jereon, Cannon, David, Case, Gary, Hannah, Ashley, Iqbal, Majid, Lloyd, Vernon, Nieves, Michael, Rance, Stuart, Rudd, Colin, Spalding, George, & Wheeldon, David. 2007c. *ITIL Service Transition*. London: The Stationery Office.

ten Hompel, Michael, & Schmidt, Thorsten. 2010. *Warehouse Management: Organisation und Steuerung von Lager- und Kommissioniersystemen*. 4th edn. Heidelberg et al.: Springer.

Thalheim, Bernhard. 2000. *Entity Relationship Modeling: Foundations of Database Technology*. Berlin et al.: Springer.

Thorenz, Lynn, & Zacher, Mathias. 2013. Cloud-Computing: Neue Chancen für das Outsourcing. *Pages 25–38 of:* Rickmann, Hagen, Diefenbach, Stefan, & Brüning, Kai T. (eds), *IT-Outsourcing*. Berlin et al.: Springer.

Unterharnscheidt, Philipp, & Kieninger, Axel. 2010. Service Level Management: Challenges and their Relevance from the Customers' Point of View. *In: Proceedings of the 16th Americas Conference on Information Systems (Paper 540)*.

van der Aalst, Wil M. P. 2013. Business Process Management: A Comprehensive Survey. *International Scholarly Research Notices Software Engineering*, **2013**(1), 1–37.

van Dinther, Clemens. 2007. *Adaptive Bidding in Single-Sided Auctions under Uncertainty: An Agent-Based Approach in Market Engineering*. Basel et al.: Birkhäuser.

van Dinther, Clemens, Knapper, Rico, Blau, Benjamin, Conte, Tobias, & Anandasivam, Arun. 2010. Managing the Quality of Modular Services: A Process-Oriented Aggregation of Expected Service Levels Based on Probability Distributions. *Pages 162–167 of: Proceedings of the Twelfth Conference on Commerce and Enterprise Computing*. Shanghai, China: IEEE.

Varian, Hal. 2010. *Intermediate Microeconomics: A Modern Approach*. 8th, int. edn. New York & London: Norton.

vom Brocke, Jan, Simons, Alexander, Niehaves, Björn, Riemer, Kai, Plattfaut, Ralf, & Cleven, Anne. 2009. Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. *Pages 2206–2217 of: Proceedings of the 17th European Conference on Information Systems*.

von Rössing, Rolf. 2005. *Betriebliches Kontinuitätsmanagement*. Bonn: mitp.

Webster, Jane, & Watson, Richard. 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, **26**(2), xiii–xxiii.

Wiedemann, Jochen. 2008. *IT-Notfallvorsorge im betrieblichen Risikomanagement: Entwicklung eines Gestaltungsmodells unter Berücksichtigung ökonomischer Aspekte am Beispiel einer TK-Unternehmung*. Ph.D. thesis, Ruhr-Universität Bochum, Germany.

Wiener, Martin, & Saunders, Carol. 2014. Forced Coopetition in IT Multi-Sourcing. *Journal of Strategic Information Systems*, **23**(3), 210–225.

Wurman, Peter, Wellman, Michael, & Walsh, William. 2001. A Parametrization of the Auction Design Space. *Games and Economic Behavior*, **35**(1-2), 304–338.

Yang, Ching-Chow. 2008. Improving the Definition and Quantification of Quality Costs. *Total Quality Management*, **19**(3).

Zambon, Emmanuele, Bolzoni, Damiano, Etalle, Sandro, & Salvato, Marco. 2007. Model-Based Mitigation of Availability Risks. *Pages 75–83 of: Proceedings of the International Workshop on Business-Driven IT Management*. Munich, Germany: IEEE.

Zarnekow, Rüdiger, Brenner, Walter, & Pilgram, Uwe. 2006. *Integrated Information Management: Applying Successful Industrial Concepts in IT*. Berlin et al.: Springer.

Zülch, Henning (ed). 2012. *International Financial Reporting Standards: Deutsch-Englische Textausgabe der von der EU gebilligten Standards und Interpretationen*. 6th edn. Weinheim: Wiley.

# Appendix

## A. SLE Optimization Problems in Case of Non-Constant Marginal Business Cost Functions

This appendix presents different forms of the IT department's SLE optimization problem given non-constant marginal business cost functions. It thus complements the discussion in Section 4.2 on this subject. Table A.1 describes the variables used in this appendix.

Table A.1.: Description of Functions and Variables

| Function / Variable | Description |
|---|---|
| $s$ | service solution $s$ from among the set $S$ of service offers |
| $c_n$ | service incident class $n$ |
| $x_n$ | value of the only attribute of the service incident class $n$ (single-attribute service incident class) |
| $x_{nm}$ | value of attribute $m$ of service incident class $n$ (multi-attribute service incident class) |
| $t_i$ | time period $i$ within reference period $T$ |
| $p_s$ | price of service solution $s$ |
| $w_{s,c_n,t_i}(x_n)$ | distribution function characterizing service solution $s$ with regard to service incident class $n$ and time period $i$ |
| $r_{s,c_n,t_i}$ | reference value of service solution $s$ with regard to service incident class $n$ and time period $i$ |
| $b_{c_n,t_i}(x_n)$ | business cost function with regard to service incident class $n$ and time period $i$ within $T$ |

First, we specify equations to be considered in case of service incident attributes with continuous value ranges. Afterwards, we provide equations to be regarded in case of service incident attributes with discrete value ranges.

### A.1. Continuous Value Range of Service Incident Attributes

In this section, we first describe the IT department's optimization problem considering one service incident class and multiple service incident classes in case of continuous service incident distribution functions. Afterwards, we regard the case in which a certain type of service incident is specified through multiple attributes.[11] Finally, we consider the case that service incident patterns and business cost functions may vary over time and, therefore, refer to certain time periods within a reference period only.

### A.1.1. One Service Incident Class

One service incident class $c_n$ with one attribute $x_n$:

$$find : s^* := \underset{s \in S}{\operatorname{argmin}} \left( \left( \int_{x_{n_{min}}}^{x_{n_{max}}} w_s(x_n) \cdot r_s \cdot b(x_n)\, dx \right) + p_s \right) \tag{A.1}$$

### A.1.2. Multiple Service Incident Classes

Multiple service incident classes $c_j$ $(j \in [1; n])$ with one attribute $x_j$ each:

$$\begin{aligned} find : s^* := \underset{s \in S}{\operatorname{argmin}} \Bigg( \Bigg( & \int_{x_{1_{min}}}^{x_{1_{max}}} w_{s,c_1}(x_1) \cdot r_{s,c_1} \cdot b_{c_1}(x_1) dx_1 + \\ & \int_{x_{2_{min}}}^{x_{2_{max}}} w_{s,c_2}(x_2) \cdot r_{s,c_2} \cdot b_{c_2}(x_2) dx_2 + ... + \\ & \int_{x_{n_{min}}}^{x_{n_{max}}} w_{s,c_n}(x_n) \cdot r_{s,c_n} \cdot b_{c_n}(x_n) dx_n \Bigg) + p_s \Bigg) \end{aligned} \tag{A.2}$$

### A.1.3. Service Incident Types Specified by Multiple Attributes

One service incident class $c_n$ with multiple attributes $x_{n1}, ...x_{nm}$:

$$\begin{aligned} find : s^* := \underset{s \in S}{\operatorname{argmin}} \Bigg( \Bigg( & \int_{x_{n1_{min}}}^{x_{n1_{max}}} ... \int_{x_{nm_{min}}}^{x_{nm_{max}}} w_{s,c_n}(x_{n1}, ..., x_{nm}) \cdot r_{s,c_n} \cdot \\ & b_{c_n}(x_{n1}, ..., x_{nm})\, dx_{n1} ... dx_{nm} \Bigg) + p_s \Bigg) \end{aligned} \tag{A.3}$$

[11] We recall our example from Section 4.3.1: The attributes 'reduced throughput incident duration' and 'reduced throughput incident degree' describe service incidents of the class 'reduced throughput incident'. Then, a pair of attribute values regarding a single service incident would represent its service incident level (e.g., a duration of three minutes and a degree of 900 Mbit/s or a duration of two minutes and a degree of 500 Mbit/s).

### A.1.4. Time Dependent Business Cost Functions and Service Incident Patterns

One service incident class $c_n$ with one attribute $x_n$, considering distribution function, reference value and business cost function to change over time (time periods $t_i$, $i \in [1; z]$):

$$find : s^* := \underset{s \in S}{\operatorname{argmin}} \left( \left( \sum_{i=1}^{z} \left( \int_{x_{n_{min}}}^{x_{n_{max}}} w_{s,c_n,t_i}(x_n) \cdot r_{s,c_n,t_i} \cdot b_{c_n,t_i}(x_n)\, dx_n \right) \right) + p_s \right) \quad \text{(A.4)}$$

## A.2. Discrete Value Range of Service Incident Attributes

In this section, we first describe the IT department's optimization problem considering one service incident class and multiple service incident classes in case of discrete service incident distribution functions. Afterwards, we regard the case in which a certain type of service incident is specified through multiple attributes. Finally, we consider the case that service incident patterns and business cost functions may vary over time and, therefore, refer to certain time periods within a reference period only.

### A.2.1. One Service Incident Class

One service incident class $c_n$ with one attribute $x_n$:

$$find : s^* := \underset{s \in S}{\operatorname{argmin}} \left( \left( \sum_{x_n = x_{n_{min}}}^{x_{n_{max}}} w_{s,c_n}(x_n) \cdot r_{s,c_n} \cdot b_{c_n}(x_n) \right) + p_s \right) \quad \text{(A.5)}$$

### A.2.2. Multiple Service Incident Classes

Multiple service incident classes $c_j$ ($j \in [1; n]$) with one attribute $x_j$ each:

$$find : s^* := \underset{s \in S}{\operatorname{argmin}} \left( \left( \sum_{j=1}^{n} \sum_{x_j = x_{j_{min}}}^{x_{j_{max}}} w_{s,c_j}(x_j) \cdot r_{s,c_j} \cdot b_{c_j}(x_j) \right) + p_s \right) \quad \text{(A.6)}$$

### A.2.3. Service Incident Types Specified by Multiple Attributes

Multiple service incident classes $c_j$ ($j \in [1; n]$) with multiple attributes $x_jk$ ($k \in [1; m]$) each:

$$find : s^* := \underset{s \in S}{\operatorname{argmin}} \left( \left( \sum_{j=1}^{n} \sum_{k=1}^{m} \sum_{x_{jk} = x_{jk_{min}}}^{x_{jk_{max}}} w_{s,c_j}(x_{jk}) \cdot r_{s,c_j} \cdot b_{c_j}(x_{jk}) \right) + p_s \right) \quad \text{(A.7)}$$

### A.2.4. Time Dependent Business Cost Functions and Service Incident Patterns

Multiple service incident classes $c_j$ ($j \in [1; n]$) with multiple attributes $x_j k$ ($k \in [1; m]$) each, considering distribution function, reference value and business cost function to change over time (time periods $t_i$, $i \in [1; z]$):

$$find : s^* := \underset{s \in S}{\operatorname{argmin}} \left( \left( \sum_{i=1}^{z} \sum_{j=1}^{n} \sum_{k=1}^{m} \sum_{x_{jk}=x_{jk_{min}}}^{x_{jk_{max}}} w_{s,c_j,t_i}(x_{jk}) \cdot r_{s,c_j,t_i} \cdot b_{c_j,t_i}(x_{jk}) \right) + p_s \right) \tag{A.8}$$

# B. Accuracy of the Sample Mean as Point Estimate for an Expected Value

Based on Kelton *et al.* (2004), this appendix elaborates on the accuracy of the sample mean as point estimate for the expected value of a random variable. It thus complements the remarks in Section 6.3 (more precisely, in Step 3.D, 'Set the Number of Replications') on this subject.

One way to measure the accuracy of a point estimate is to form a confidence interval around it (Kelton *et al.*, 2004, p.611), which contains the actual parameter value regarding the corresponding random variable with a predefined probability (i.e., with a certain confidence level, $1-\alpha$). The confidence interval for the expected value of a random variable $E(Y)$ is calculated as follows (e.g., Bol, 2003, p.105; Kelton *et al.*, 2004, p.612):[12]

$$\left[\bar{Y} - t_{n-1,1-\alpha/2}\frac{S}{\sqrt{n}}\ ;\ \bar{Y} + t_{n-1,1-\alpha/2}\frac{S}{\sqrt{n}}\right] \tag{B.1}$$

In Interval B.1 the term $t_{n-1,1-\alpha/2}$ represents the $(1-\alpha/2)$ quantile of the Student's t distribution with n-1 degrees of freedom. Furthermore, $\bar{Y}$ is the sample mean, $S$ the sample standard deviation and $n$ the sample size. For normally-distributed random variables, the interval provides an accurate probability of coverage.

However, the Central Limit Theorem states that the sample mean $\bar{Y}$ of independent, identically distributed realizations of a random variable follows a normal distribution with mean $\mu$ and variance $\sigma^2/n$ as $n \to \infty$ (Birta & Arbez, 2013, p.369). Therefore, the confidence interval presented above may be applied to approximately determine the confidence interval of point estimates for expected values of random variables which follow other distributions as well (Bol, 2003, p.197).

In Section 6.3 we specified the half width $h$ of an interval around the point estimate $\bar{Y}$ in order to calculate the number of replications required. This half width corresponds to the term $t_{n-1,1-\alpha/2}\frac{S}{\sqrt{n}}$ in the interval above. Hence, we can set

$$h = t_{n-1,1-\alpha/2}\frac{S}{\sqrt{n}} \tag{B.2}$$

and solve for $n$. We obtain

$$n = t^2_{n-1,1-\alpha/2}\frac{S^2}{h^2}. \tag{B.3}$$

For sample sizes $n > 30$ the $1-\alpha/2$ quantile of the Student's t distribution may be replaced by the $1-\alpha/2$ quantile of the standard normal distribution — since the

[12] A derivation of Interval B.1 is provided, for instance, in Birta & Arbez (2013, pp.371-376).

Student's t distribution approximates the standard normal distribution (Bol, 2003, p.224). The resulting equation

$$n \approx z_{1-\alpha/2}^2 \frac{S^2}{h^2} \tag{B.4}$$

corresponds to Equation 6.2 which we have presented in Section 6.3.

# C. Common Random Numbers in Simulation Experiments

In this appendix we discuss an approach to more accurately estimate the difference in expected values of random variables between two configurations of a simulation model (e.g., between a service incident scenario and the base case) without increasing the number of replications.

One way to achieve this goal is to conduct simulation experiments with both configurations "under conditions that are as similar as possible, except for the model changes made" (Kelton *et al.*, 2004, p.510). Therefore, the approach of 'Common Random Numbers'[13] (e.g., Kelton *et al.*, 2004, pp.509-515; Banks *et al.*, 2010, pp.486-492) suggests to use the same random numbers for simulation experiments with the first and the second model configuration. Thus, it can be ascertained that differences in expected values of random variables observed are due to differences in model configurations and not a result of random fluctuation (Banks *et al.*, 2010, p.481) — i.e., not a consequence of "random numbers having bounced differently" (Kelton *et al.*, 2004, p.511).

Practically, a dedicated stream of random numbers is generated for each random variable which is defined in the simulation model (Kelton *et al.*, 2004, p.511). These streams are then utilized in simulation experiments with every configuration of the simulation model (if possible in consideration of model logic and differences between configurations). That is, random numbers are 'synchronized' so that "for each replication, the same random numbers are used" in all model configurations (Banks *et al.*, 2010, p.486).

The application of common random numbers aims to induce positive correlation between corresponding realizations of random variables across model configurations (Banks *et al.*, 2010, p.486; Kelton *et al.*, 2004, p.515). For instance, using common random numbers in our setting, $Y_{m,b,r}$ and $Y_{m,i_j,r}$ are rather correlated and not independent any more. This positive correlation influences the covariance of two random variables $Y_A$ and $Y_B$ and, thus, decreases the variance of the point estimate for their mean difference (see Equation C.1).[14]

$$Var(\bar{Y}_A - \bar{Y}_B) = Var(\bar{Y}_A) + Var(\bar{Y}_A) - 2Cov(\bar{Y}_B; \bar{Y}_A) \tag{C.1}$$

A smaller variance, in turn, decreases the half width of the corresponding confidence interval (see Appendix B). Consequently, the point estimate for the mean difference is more precise. Further "variance reduction techniques" are mentioned, for instance, by Kelton *et al.* (2004, pp.515-516).

[13] This approach is also denoted as "correlated sampling" in literature (Banks *et al.*, 2010, p.486).

[14] In case of a comparison between a service incident scenario and the base case, Equation C.1 turns into $Var(\bar{Y}_{m,i_j} - \bar{Y}_{m,b}) = Var(\bar{Y}_{m,i_j}) + Var(\bar{Y}_{m,b}) - 2Cov(\bar{Y}_{m,i_j}; \bar{Y}_{m,b})$.

# D. Detailed Simulation Results

This appendix provides the detailed simulation results of the case study presented in Chapter 8.

## D.1. Aggregated Results of the Initial 31 Replications per Simulation Scenario

Table D.1 and Table D.2 aggregate the changes in business process metrics (sample mean, confidence interval and standard deviation) which were calculated based on the results of the initial 31 simulation runs (replications) of each simulation scenario.

Table D.1.: Simulation Results with Regard to the Number of Additional Shipping Fee Waivers (Confidence Level 95%; Half Width 12.5 Waivers; 31 Replications)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation | Number of Replications Required |
|---|---|---|---|---|
| (10:00; 10) | 1.03 | [0.45; 1.62] | 1.66 | 0.1 |
| (10:00; 20) | 0.61 | [0.18; 1.05] | 1.23 | 0.0 |
| (10:00; 30) | 1.32 | [0.6; 2.05] | 2.06 | 0.1 |
| (10:00; 40) | 3.16 | [1.69; 4.63] | 4.18 | 0.4 |
| (10:00; 50) | 6.55 | [4.25; 8.85] | 6.53 | 1.0 |
| (10:00; 60) | 12.13 | [8.12; 16.14] | 11.39 | 3.2 |
| (10:00; 70) | 29.10 | [23.47; 34.72] | 15.98 | 6.3 |
| (10:00; 80) | 59.23 | [49.29; 69.16] | 28.22 | 19.6 |
| (10:30; 10) | 0.55 | [0.06; 1.04] | 1.39 | 0.0 |
| (10:30; 20) | 0.26 | [0.08; 0.44] | 0.51 | 0.0 |
| (10:30; 30) | 0.84 | [0.47; 1.2] | 1.04 | 0.0 |
| (10:30; 40) | 3.48 | [1.45; 5.52] | 5.77 | 0.8 |
| (10:30; 50) | 4.77 | [3.2; 6.35] | 4.46 | 0.5 |
| (10:30; 60) | 12.58 | [9.26; 15.9] | 9.43 | 2.2 |
| (10:30; 70) | 30.55 | [24.22; 36.88] | 17.98 | 7.9 |
| (10:30; 80) | 52.52 | [44.76; 60.28] | 22.05 | 11.9 |
| (11:00; 10) | 1.77 | [0.91; 2.64] | 2.46 | 0.1 |
| (11:00; 20) | 3.94 | [2.53; 5.34] | 3.98 | 0.4 |
| (11:00; 30) | 6.10 | [4.03; 8.16] | 5.87 | 0.8 |
| (11:00; 40) | 12.65 | [8.94; 16.35] | 10.51 | 2.7 |
| (11:00; 50) | 32.45 | [25.54; 39.36] | 19.63 | 9.5 |
| (11:00; 60) | 75.03 | [66.66; 83.4] | 23.77 | 13.9 |
| (11:00; 70) | 129.55 | [116.12; 142.98] | 38.15 | 35.8 |
| (11:00; 80) | 235.39 | [220.32; 250.45] | 42.79 | 45.0 |

(continued on next page)

Table D.1 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation | Number of Replications Required |
|---|---|---|---|---|
| (11:30; 10) | 1.35 | [0.45; 2.26] | 2.56 | 0.2 |
| (11:30; 20) | 3.19 | [1.55; 4.83] | 4.66 | 0.5 |
| (11:30; 30) | 7.94 | [4.82; 11.05] | 8.85 | 1.9 |
| (11:30; 40) | 16.03 | [11.24; 20.83] | 13.62 | 4.6 |
| (11:30; 50) | 36.16 | [28.92; 43.4] | 20.58 | 10.4 |
| (11:30; 60) | 70.48 | [59.74; 81.23] | 30.52 | 22.9 |
| (11:30; 70) | 133.45 | [118.85; 148.05] | 41.47 | 42.3 |
| (11:30; 80) | 231.00 | [216.12; 245.88] | 42.28 | 43.9 |
| (12:00; 10) | 2.55 | [0.6; 4.5] | 5.54 | 0.8 |
| (12:00; 20) | 2.65 | [1.66; 3.63] | 2.81 | 0.2 |
| (12:00; 30) | 7.42 | [5.08; 9.76] | 6.66 | 1.1 |
| (12:00; 40) | 19.06 | [14.89; 23.24] | 11.85 | 3.5 |
| (12:00; 50) | 40.68 | [31.26; 50.09] | 26.74 | 17.6 |
| (12:00; 60) | 67.23 | [56.2; 78.25] | 31.32 | 24.1 |
| (12:00; 70) | 132.58 | [118.34; 146.82] | 40.44 | 40.2 |
| (12:00; 80) | 245.61 | [231.46; 259.76] | 40.20 | 39.7 |
| (12:30; 10) | 1.10 | [0.52; 1.67] | 1.64 | 0.1 |
| (12:30; 20) | 2.84 | [1.25; 4.43] | 4.52 | 0.5 |
| (12:30; 30) | 8.29 | [4.99; 11.59] | 9.39 | 2.2 |
| (12:30; 40) | 11.10 | [8.2; 13.99] | 8.22 | 1.7 |
| (12:30; 50) | 27.35 | [20.84; 33.87] | 18.51 | 8.4 |
| (12:30; 60) | 61.29 | [52.14; 70.44] | 26.00 | 16.6 |
| (12:30; 70) | 132.90 | [118.9; 146.9] | 39.77 | 38.9 |
| (12:30; 80) | 223.06 | [207.02; 239.11] | 45.57 | 49.3 |
| (01:00; 10) | 1.00 | [0.55; 1.45] | 1.26 | 0.0 |
| (01:00; 20) | 2.94 | [2.05; 3.83] | 2.53 | 0.2 |
| (01:00; 30) | 6.42 | [4.05; 8.79] | 6.74 | 1.1 |
| (01:00; 40) | 18.06 | [14.19; 21.94] | 11.01 | 3.0 |
| (01:00; 50) | 28.39 | [22.63; 34.14] | 16.35 | 6.6 |
| (01:00; 60) | 67.39 | [58.4; 76.37] | 25.53 | 16.0 |
| (01:00; 70) | 131.81 | [118.54; 145.08] | 37.70 | 34.9 |
| (01:00; 80) | 220.90 | [205.57; 236.24] | 43.56 | 46.6 |
| (01:30; 10) | 1.06 | [0.36; 1.77] | 2.00 | 0.1 |
| (01:30; 20) | 3.29 | [1.94; 4.65] | 3.85 | 0.4 |
| (01:30; 30) | 7.84 | [4.51; 11.17] | 9.46 | 2.2 |
| (01:30; 40) | 15.68 | [11.48; 19.88] | 11.94 | 3.5 |
| (01:30; 50) | 31.58 | [26.69; 36.47] | 13.88 | 4.7 |
| (01:30; 60) | 72.90 | [61.78; 84.02] | 31.59 | 24.5 |
| (01:30; 70) | 145.39 | [133.17; 157.61] | 34.71 | 29.6 |

(continued on next page)

Table D.1 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation | Number of Replications Required |
|---|---|---|---|---|
| (01:30; 80) | 229.87 | [214.04; 245.7] | 44.97 | 49.7 |
| (02:00; 10) | 1.65 | [0.68; 2.61] | 2.74 | 0.2 |
| (02:00; 20) | 3.58 | [1.66; 5.5] | 5.45 | 0.7 |
| (02:00; 30) | 5.32 | [3.36; 7.29] | 5.58 | 0.8 |
| (02:00; 40) | 14.03 | [9.44; 18.63] | 13.05 | 4.2 |
| (02:00; 50) | 38.32 | [29.56; 47.09] | 24.91 | 15.2 |
| (02:00; 60) | 79.87 | [69.13; 90.62] | 30.53 | 22.9 |
| (02:00; 70) | 142.00 | [126.28; 157.72] | 44.66 | 49.0 |
| (02:00; 80) | 252.52 | [236.99; 268.04] | 44.11 | 47.8 |
| (02:30; 10) | 1.23 | [0.35; 2.1] | 2.49 | 0.2 |
| (02:30; 20) | 2.81 | [1.79; 3.82] | 2.88 | 0.2 |
| (02:30; 30) | 8.16 | [4.92; 11.4] | 9.20 | 2.1 |
| (02:30; 40) | 14.90 | [11.04; 18.76] | 10.97 | 3.0 |
| (02:30; 50) | 32.26 | [25.85; 38.67] | 18.20 | 8.1 |
| (02:30; 60) | 68.90 | [59.52; 78.29] | 26.67 | 17.5 |
| (02:30; 70) | 135.48 | [122.07; 148.9] | 38.10 | 35.7 |
| (02:30; 80) | 231.10 | [216.61; 245.59] | 41.16 | 41.6 |
| (03:00; 10) | 0.74 | [0.27; 1.21] | 1.34 | 0.0 |
| (03:00; 20) | 3.00 | [1.73; 4.27] | 3.61 | 0.3 |
| (03:00; 30) | 8.00 | [4.9; 11.1] | 8.81 | 1.9 |
| (03:00; 40) | 13.61 | [10.1; 17.12] | 9.98 | 2.4 |
| (03:00; 50) | 31.03 | [24.63; 37.44] | 18.20 | 8.1 |
| (03:00; 60) | 74.16 | [65.6; 82.72] | 24.31 | 14.5 |
| (03:00; 70) | 74.06 | [62.69; 85.44] | 32.32 | 25.7 |
| (03:00; 80) | 67.19 | [58.61; 75.78] | 24.39 | 14.6 |
| (03:30; 10) | 1.71 | [0.7; 2.71] | 2.85 | 0.2 |
| (03:30; 20) | 3.42 | [1.76; 5.08] | 4.72 | 0.5 |
| (03:30; 30) | 9.65 | [6.11; 13.18] | 10.03 | 2.5 |
| (03:30; 40) | 6.52 | [4.65; 8.38] | 5.30 | 0.7 |
| (03:30; 50) | 7.42 | [5.01; 9.83] | 6.84 | 1.1 |
| (03:30; 60) | 5.55 | [3.18; 7.92] | 6.73 | 1.1 |
| (03:30; 70) | 4.65 | [3.26; 6.03] | 3.95 | 0.4 |
| (03:30; 80) | 5.35 | [3.5; 7.21] | 5.28 | 0.7 |
| (04:00; 10) | 1.06 | [0.42; 1.71] | 1.82 | 0.1 |
| (04:00; 20) | 0.58 | [0.08; 1.08] | 1.43 | 0.1 |
| (04:00; 30) | 1.35 | [0.36; 2.35] | 2.83 | 0.2 |
| (04:00; 40) | 0.55 | [0.13; 0.96] | 1.18 | 0.0 |
| (04:00; 50) | 0.29 | [0; 0.58] | 0.82 | 0.0 |
| (04:00; 60) | 0.87 | [0.34; 1.41] | 1.52 | 0.1 |

(continued on next page)

Table D.1 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation | Number of Replications Required |
|---|---|---|---|---|
| (04:00; 70) | 0.48 | [-0.1; 1.06] | 1.65 | 0.1 |
| (04:00; 80) | 0.32 | [0.06; 0.59] | 0.75 | 0.0 |
| (04:30; 10) | 0.81 | [0.34; 1.27] | 1.33 | 0.0 |
| (04:30; 20) | 0.58 | [0.24; 0.92] | 0.96 | 0.0 |
| (04:30; 30) | 0.77 | [0.22; 1.32] | 1.56 | 0.1 |
| (04:30; 40) | 0.68 | [0.04; 1.32] | 1.81 | 0.1 |
| (04:30; 50) | 0.81 | [0.21; 1.4] | 1.68 | 0.1 |
| (04:30; 60) | 0.61 | [0.29; 0.94] | 0.92 | 0.0 |
| (04:30; 70) | 1.16 | [0.53; 1.79] | 1.79 | 0.1 |
| (04:30; 80) | 0.35 | [0.12; 0.59] | 0.66 | 0.0 |
| (05:00; 10) | 0.68 | [0.32; 1.03] | 1.01 | 0.0 |
| (05:00; 20) | 0.42 | [0.15; 0.69] | 0.76 | 0.0 |
| (05:00; 30) | 0.45 | [0.09; 0.81] | 1.03 | 0.0 |
| (05:00; 40) | 0.74 | [0.18; 1.3] | 1.59 | 0.1 |
| (05:00; 50) | 0.58 | [0.24; 0.92] | 0.96 | 0.0 |
| (05:00; 60) | 1.00 | [0.46; 1.54] | 1.53 | 0.1 |
| (05:00; 70) | 0.97 | [0.45; 1.49] | 1.47 | 0.1 |
| (05:00; 80) | 0.61 | [0.14; 1.08] | 1.33 | 0.0 |
| (05:30; 10) | 0.00 | [0; 0] | 0.00 | 0.0 |
| (05:30; 20) | 0.71 | [0.24; 1.18] | 1.35 | 0.0 |
| (05:30; 30) | 0.68 | [0.05; 1.31] | 1.80 | 0.1 |
| (05:30; 40) | 0.52 | [0.23; 0.8] | 0.81 | 0.0 |
| (05:30; 50) | 0.42 | [0.02; 0.82] | 1.15 | 0.0 |
| (05:30; 60) | 0.68 | [0.05; 1.31] | 1.80 | 0.1 |
| (05:30; 70) | 0.55 | [0.25; 0.85] | 0.85 | 0.0 |
| (05:30; 80) | 0.87 | [-0.15; 1.89] | 2.90 | 0.2 |
| (06:00; 10) | 0.61 | [-0.1; 1.32] | 2.01 | 0.1 |
| (06:00; 20) | 0.87 | [0.14; 1.6] | 2.08 | 0.1 |
| (06:00; 30) | 0.58 | [0.14; 1.02] | 1.26 | 0.0 |
| (06:00; 40) | 0.52 | [0.08; 0.95] | 1.23 | 0.0 |
| (06:00; 50) | 0.65 | [0.23; 1.06] | 1.17 | 0.0 |
| (06:00; 60) | 0.35 | [0.14; 0.57] | 0.61 | 0.0 |
| (06:00; 70) | 0.32 | [0.08; 0.57] | 0.70 | 0.0 |
| (06:00; 80) | 0.87 | [0.07; 1.67] | 2.26 | 0.1 |
| (06:30; 10) | 0.81 | [0.28; 1.33] | 1.49 | 0.1 |
| (06:30; 20) | 0.81 | [0.23; 1.38] | 1.64 | 0.1 |
| (06:30; 30) | 0.45 | [0.06; 0.85] | 1.12 | 0.0 |
| (06:30; 40) | 0.68 | [0.23; 1.13] | 1.28 | 0.0 |
| (06:30; 50) | 0.81 | [0.32; 1.29] | 1.38 | 0.0 |

(continued on next page)

Table D.1 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation | Number of Replications Required |
|---|---|---|---|---|
| (06:30; 60) | 0.87 | [0.26; 1.49] | 1.75 | 0.1 |
| (06:30; 70) | 0.84 | [0.26; 1.42] | 1.66 | 0.1 |
| (06:30; 80) | 0.52 | [0.15; 0.88] | 1.03 | 0.0 |

As analysts choose the highest number of replications required as the target number of simulation runs for all simulation scenarios, (at least) 50 replications per incident scenario are required in order to achieve the predefined accuracy of estimate.

Table D.2.: Simulation Results with Regard to Overtime in Minutes (Confidence Level 95%; Half Width 16.67 Minutes; 31 Replications)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation | Number of Replications Required |
|---|---|---|---|---|
| (10:00; 10) | 0.71 | [-0.14; 1.57] | 2.43 | 0.1 |
| (10:00; 20) | 1.21 | [0.44; 1.97] | 2.18 | 0.1 |
| (10:00; 30) | 3.49 | [2.21; 4.77] | 3.63 | 0.2 |
| (10:00; 40) | 12.22 | [10.42; 14.03] | 5.12 | 0.4 |
| (10:00; 50) | 23.21 | [21.02; 25.4] | 6.23 | 0.5 |
| (10:00; 60) | 30.63 | [28.88; 32.38] | 4.97 | 0.3 |
| (10:00; 70) | 42.63 | [40.69; 44.57] | 5.52 | 0.4 |
| (10:00; 80) | 51.99 | [50.22; 53.76] | 5.03 | 0.4 |
| (10:30; 10) | 0.90 | [0.11; 1.69] | 2.24 | 0.1 |
| (10:30; 20) | 0.46 | [0.02; 0.89] | 1.24 | 0.0 |
| (10:30; 30) | 3.25 | [2.15; 4.35] | 3.12 | 0.1 |
| (10:30; 40) | 10.83 | [8.79; 12.87] | 5.79 | 0.5 |
| (10:30; 50) | 22.15 | [20.41; 23.89] | 4.95 | 0.3 |
| (10:30; 60) | 32.80 | [31.62; 33.98] | 3.35 | 0.2 |
| (10:30; 70) | 42.84 | [41.25; 44.42] | 4.51 | 0.3 |
| (10:30; 80) | 51.55 | [50.19; 52.91] | 3.86 | 0.2 |
| (11:00; 10) | 8.17 | [6.58; 9.77] | 4.52 | 0.3 |
| (11:00; 20) | 17.06 | [15.28; 18.83] | 5.04 | 0.4 |
| (11:00; 30) | 26.34 | [24.85; 27.82] | 4.22 | 0.2 |
| (11:00; 40) | 36.62 | [34.81; 38.42] | 5.14 | 0.4 |
| (11:00; 50) | 48.29 | [46.34; 50.25] | 5.56 | 0.4 |

(continued on next page)

Table D.2 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation | Number of Replications Required |
|---|---|---|---|---|
| (11:00; 60) | 58.03 | [56.21; 59.84] | 5.15 | 0.4 |
| (11:00; 70) | 67.56 | [66.19; 68.93] | 3.89 | 0.2 |
| (11:00; 80) | 78.98 | [76.94; 81.03] | 5.80 | 0.5 |
| (11:30; 10) | 7.84 | [6.13; 9.55] | 4.86 | 0.3 |
| (11:30; 20) | 18.97 | [17.14; 20.81] | 5.21 | 0.4 |
| (11:30; 30) | 26.87 | [25.07; 28.68] | 5.13 | 0.4 |
| (11:30; 40) | 37.28 | [35.27; 39.29] | 5.71 | 0.5 |
| (11:30; 50) | 46.84 | [45.38; 48.31] | 4.17 | 0.2 |
| (11:30; 60) | 56.97 | [55.29; 58.66] | 4.79 | 0.3 |
| (11:30; 70) | 66.60 | [65.14; 68.05] | 4.14 | 0.2 |
| (11:30; 80) | 76.87 | [75.39; 78.34] | 4.20 | 0.2 |
| (12:00; 10) | 7.04 | [5.33; 8.75] | 4.87 | 0.3 |
| (12:00; 20) | 16.60 | [14.89; 18.31] | 4.85 | 0.3 |
| (12:00; 30) | 27.18 | [25.46; 28.9] | 4.89 | 0.3 |
| (12:00; 40) | 37.03 | [35.35; 38.71] | 4.76 | 0.3 |
| (12:00; 50) | 47.35 | [45.72; 48.98] | 4.62 | 0.3 |
| (12:00; 60) | 55.92 | [53.75; 58.09] | 6.17 | 0.5 |
| (12:00; 70) | 66.86 | [65.17; 68.55] | 4.81 | 0.3 |
| (12:00; 80) | 76.50 | [74.76; 78.23] | 4.92 | 0.3 |
| (12:30; 10) | 5.82 | [4.48; 7.15] | 3.79 | 0.2 |
| (12:30; 20) | 16.55 | [14.58; 18.52] | 5.59 | 0.4 |
| (12:30; 30) | 27.69 | [25.98; 29.4] | 4.87 | 0.3 |
| (12:30; 40) | 35.72 | [33.68; 37.76] | 5.80 | 0.5 |
| (12:30; 50) | 46.12 | [43.95; 48.3] | 6.18 | 0.5 |
| (12:30; 60) | 55.67 | [53.6; 57.74] | 5.88 | 0.5 |
| (12:30; 70) | 65.89 | [63.42; 68.36] | 7.01 | 0.7 |
| (12:30; 80) | 76.76 | [74.77; 78.75] | 5.65 | 0.4 |
| (01:00; 10) | 7.11 | [5.82; 8.39] | 3.65 | 0.2 |
| (01:00; 20) | 17.81 | [16.24; 19.38] | 4.46 | 0.3 |
| (01:00; 30) | 28.07 | [26.2; 29.94] | 5.31 | 0.4 |
| (01:00; 40) | 37.65 | [35.74; 39.56] | 5.43 | 0.4 |
| (01:00; 50) | 46.05 | [44.45; 47.65] | 4.55 | 0.3 |
| (01:00; 60) | 57.24 | [55.82; 58.67] | 4.05 | 0.2 |
| (01:00; 70) | 67.28 | [65.9; 68.67] | 3.94 | 0.2 |
| (01:00; 80) | 76.01 | [74.14; 77.88] | 5.31 | 0.4 |
| (01:30; 10) | 5.72 | [4.27; 7.16] | 4.11 | 0.2 |
| (01:30; 20) | 17.42 | [15.4; 19.44] | 5.73 | 0.5 |
| (01:30; 30) | 27.76 | [25.27; 30.25] | 7.08 | 0.7 |
| (01:30; 40) | 37.33 | [35.45; 39.22] | 5.35 | 0.4 |

(continued on next page)

Table D.2 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation | Number of Replications Required |
|---|---|---|---|---|
| (01:30; 50) | 46.85 | [45.49; 48.21] | 3.86 | 0.2 |
| (01:30; 60) | 57.22 | [55.04; 59.39] | 6.19 | 0.5 |
| (01:30; 70) | 67.19 | [65.27; 69.12] | 5.47 | 0.4 |
| (01:30; 80) | 76.64 | [74.95; 78.32] | 4.79 | 0.3 |
| (02:00; 10) | 7.92 | [6.31; 9.53] | 4.57 | 0.3 |
| (02:00; 20) | 18.14 | [15.92; 20.37] | 6.33 | 0.6 |
| (02:00; 30) | 26.49 | [24.62; 28.35] | 5.30 | 0.4 |
| (02:00; 40) | 37.50 | [36.08; 38.92] | 4.03 | 0.2 |
| (02:00; 50) | 48.89 | [46.86; 50.91] | 5.75 | 0.5 |
| (02:00; 60) | 58.33 | [56.59; 60.07] | 4.95 | 0.3 |
| (02:00; 70) | 67.16 | [64.95; 69.37] | 6.28 | 0.5 |
| (02:00; 80) | 77.76 | [76.02; 79.5] | 4.94 | 0.3 |
| (02:30; 10) | 6.80 | [5.52; 8.08] | 3.65 | 0.2 |
| (02:30; 20) | 17.20 | [15.7; 18.7] | 4.26 | 0.3 |
| (02:30; 30) | 28.56 | [26.81; 30.32] | 4.99 | 0.3 |
| (02:30; 40) | 36.51 | [34.71; 38.31] | 5.13 | 0.4 |
| (02:30; 50) | 46.69 | [44.82; 48.57] | 5.33 | 0.4 |
| (02:30; 60) | 56.25 | [54.83; 57.67] | 4.03 | 0.2 |
| (02:30; 70) | 67.74 | [66.06; 69.43] | 4.78 | 0.3 |
| (02:30; 80) | 77.68 | [75.89; 79.47] | 5.09 | 0.4 |
| (03:00; 10) | 6.73 | [5.07; 8.38] | 4.70 | 0.3 |
| (03:00; 20) | 16.50 | [14.98; 18.02] | 4.32 | 0.3 |
| (03:00; 30) | 25.65 | [23.9; 27.39] | 4.96 | 0.3 |
| (03:00; 40) | 35.62 | [34.15; 37.09] | 4.18 | 0.2 |
| (03:00; 50) | 45.38 | [43.41; 47.35] | 5.59 | 0.4 |
| (03:00; 60) | 56.50 | [54.6; 58.4] | 5.39 | 0.4 |
| (03:00; 70) | 67.10 | [65.01; 69.19] | 5.94 | 0.5 |
| (03:00; 80) | 75.51 | [73.67; 77.35] | 5.22 | 0.4 |
| (03:30; 10) | 7.73 | [5.66; 9.81] | 5.90 | 0.5 |
| (03:30; 20) | 17.16 | [15.26; 19.06] | 5.39 | 0.4 |
| (03:30; 30) | 27.59 | [25.71; 29.48] | 5.35 | 0.4 |
| (03:30; 40) | 37.42 | [36.09; 38.74] | 3.76 | 0.2 |
| (03:30; 50) | 47.05 | [45.32; 48.78] | 4.92 | 0.3 |
| (03:30; 60) | 55.88 | [54.11; 57.64] | 5.02 | 0.3 |
| (03:30; 70) | 66.36 | [64.46; 68.27] | 5.42 | 0.4 |
| (03:30; 80) | 77.38 | [75.5; 79.26] | 5.34 | 0.4 |
| (04:00; 10) | 6.77 | [5.32; 8.22] | 4.13 | 0.2 |
| (04:00; 20) | 16.12 | [14.67; 17.57] | 4.12 | 0.2 |
| (04:00; 30) | 28.58 | [26.5; 30.65] | 5.90 | 0.5 |

(continued on next page)

Table D.2 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation | Number of Replications Required |
|---|---|---|---|---|
| (04:00; 40) | 37.24 | [35.7; 38.77] | 4.36 | 0.3 |
| (04:00; 50) | 45.74 | [44.19; 47.3] | 4.42 | 0.3 |
| (04:00; 60) | 56.54 | [55.19; 57.89] | 3.83 | 0.2 |
| (04:00; 70) | 66.02 | [64.25; 67.8] | 5.05 | 0.4 |
| (04:00; 80) | 76.30 | [74.37; 78.24] | 5.50 | 0.4 |
| (04:30; 10) | 6.90 | [5.56; 8.23] | 3.80 | 0.2 |
| (04:30; 20) | 16.56 | [15; 18.13] | 4.45 | 0.3 |
| (04:30; 30) | 27.70 | [25.63; 29.77] | 5.89 | 0.5 |
| (04:30; 40) | 36.49 | [34.77; 38.21] | 4.90 | 0.3 |
| (04:30; 50) | 46.91 | [45.19; 48.63] | 4.89 | 0.3 |
| (04:30; 60) | 56.22 | [54.79; 57.65] | 4.07 | 0.2 |
| (04:30; 70) | 68.07 | [66.45; 69.69] | 4.61 | 0.3 |
| (04:30; 80) | 76.63 | [75.28; 77.98] | 3.83 | 0.2 |
| (05:00; 10) | 8.23 | [6.85; 9.62] | 3.94 | 0.2 |
| (05:00; 20) | 16.19 | [14.46; 17.92] | 4.91 | 0.3 |
| (05:00; 30) | 27.73 | [26.06; 29.4] | 4.74 | 0.3 |
| (05:00; 40) | 36.44 | [34.36; 38.52] | 5.91 | 0.5 |
| (05:00; 50) | 47.88 | [46.33; 49.43] | 4.40 | 0.3 |
| (05:00; 60) | 57.13 | [55.52; 58.73] | 4.56 | 0.3 |
| (05:00; 70) | 69.10 | [67.79; 70.41] | 3.73 | 0.2 |
| (05:00; 80) | 77.40 | [75.19; 79.6] | 6.27 | 0.5 |
| (05:30; 10) | 0.00 | [0; 0] | 0.00 | 0.0 |
| (05:30; 20) | 10.99 | [8.88; 13.1] | 5.99 | 0.5 |
| (05:30; 30) | 19.76 | [17.3; 22.22] | 7.00 | 0.7 |
| (05:30; 40) | 29.25 | [27.23; 31.28] | 5.76 | 0.5 |
| (05:30; 50) | 37.80 | [35.79; 39.82] | 5.73 | 0.5 |
| (05:30; 60) | 48.36 | [45.64; 51.09] | 7.75 | 0.8 |
| (05:30; 70) | 59.03 | [56.93; 61.12] | 5.95 | 0.5 |
| (05:30; 80) | 68.37 | [66.41; 70.33] | 5.57 | 0.4 |
| (06:00; 10) | 3.88 | [1.64; 6.12] | 6.37 | 0.6 |
| (06:00; 20) | 8.68 | [4.69; 12.68] | 11.35 | 1.8 |
| (06:00; 30) | 11.42 | [5.86; 16.99] | 15.80 | 3.5 |
| (06:00; 40) | 9.37 | [3.12; 15.62] | 17.75 | 4.4 |
| (06:00; 50) | 18.26 | [9.42; 27.1] | 25.10 | 8.7 |
| (06:00; 60) | 19.99 | [9.61; 30.37] | 29.48 | 12.0 |
| (06:00; 70) | 35.14 | [22.12; 48.15] | 36.96 | 18.9 |
| (06:00; 80) | 19.09 | [6.42; 31.76] | 35.99 | 17.9 |
| (06:30; 10) | 0.77 | [0.2; 1.34] | 1.62 | 0.0 |
| (06:30; 20) | 0.59 | [0; 1.18] | 1.68 | 0.0 |

(continued on next page)

Table D.2 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation | Number of Replications Required |
|---|---|---|---|---|
| (06:30; 30) | 1.05 | [0.41; 1.68] | 1.79 | 0.0 |
| (06:30; 40) | 0.88 | [0.27; 1.49] | 1.72 | 0.0 |
| (06:30; 50) | 0.65 | [0.09; 1.21] | 1.60 | 0.0 |
| (06:30; 60) | 0.29 | [-0.06; 0.65] | 1.01 | 0.0 |
| (06:30; 70) | 1.54 | [0.43; 2.65] | 3.15 | 0.1 |
| (06:30; 80) | 1.42 | [0.48; 2.37] | 2.67 | 0.1 |

As analysts choose the highest number of replications required as the target number of simulation runs for all simulation scenarios, (at least) 19 replications per incident scenario would be required in order to achieve the predefined accuracy of estimate.

## D.2. Aggregated Results of the 50 Replications per Simulation Scenario

Analogously, Table D.3 and Table D.4 summarize the changes in the number of additional shipping fee waivers granted or minutes of additional overtime required, which were calculated based on the results of the 50 simulation runs of each simulation scenario.

Table D.3.: Simulation Results with Regard to the Number of Additional Shipping Fee Waivers (Confidence Level 95%; 50 Replications)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation |
|---|---|---|---|
| (10:00; 10) | 0.70 | [0.26; 1.14] | 1.58 |
| (10:00; 20) | 1.02 | [0.38; 1.66] | 2.32 |
| (10:00; 30) | 0.78 | [0.37; 1.19] | 1.47 |
| (10:00; 40) | 3.06 | [1.88; 4.24] | 4.24 |
| (10:00; 50) | 4.64 | [3.22; 6.06] | 5.13 |
| (10:00; 60) | 14.60 | [11.75; 17.45] | 10.29 |
| (10:00; 70) | 29.96 | [25.32; 34.6] | 16.74 |
| (10:00; 80) | 63.78 | [55.26; 72.3] | 30.73 |
| (10:30; 10) | 0.78 | [0.38; 1.18] | 1.46 |
| (10:30; 20) | 1.06 | [0.6; 1.52] | 1.65 |

(continued on next page)

Table D.3 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation |
|---|---|---|---|
| (10:30; 30) | 1.58 | [0.99; 2.17] | 2.14 |
| (10:30; 40) | 2.74 | [1.73; 3.75] | 3.65 |
| (10:30; 50) | 6.24 | [4.53; 7.95] | 6.16 |
| (10:30; 60) | 16.34 | [13.03; 19.65] | 11.94 |
| (10:30; 70) | 26.66 | [21.74; 31.58] | 17.76 |
| (10:30; 80) | 65.36 | [57.07; 73.65] | 29.92 |
| (11:00; 10) | 1.20 | [0.72; 1.68] | 1.75 |
| (11:00; 20) | 2.80 | [0.84; 4.76] | 7.07 |
| (11:00; 30) | 7.22 | [5.55; 8.89] | 6.01 |
| (11:00; 40) | 15.18 | [12.57; 17.79] | 9.43 |
| (11:00; 50) | 29.70 | [24.4; 35] | 19.12 |
| (11:00; 60) | 66.88 | [56.8; 76.96] | 36.37 |
| (11:00; 70) | 150.66 | [140.9; 160.42] | 35.21 |
| (11:00; 80) | 239.24 | [225.19; 253.29] | 50.70 |
| (11:30; 10) | 1.78 | [0.94; 2.62] | 3.03 |
| (11:30; 20) | 2.62 | [1.68; 3.56] | 3.39 |
| (11:30; 30) | 5.38 | [3.47; 7.29] | 6.89 |
| (11:30; 40) | 12.50 | [9.18; 15.82] | 11.96 |
| (11:30; 50) | 31.30 | [26.65; 35.95] | 16.78 |
| (11:30; 60) | 72.52 | [62.14; 82.9] | 37.44 |
| (11:30; 70) | 140.54 | [129.15; 151.93] | 41.10 |
| (11:30; 80) | 229.82 | [219.73; 239.91] | 36.40 |
| (12:00; 10) | 1.24 | [0.73; 1.75] | 1.82 |
| (12:00; 20) | 2.58 | [1.64; 3.52] | 3.41 |
| (12:00; 30) | 7.42 | [5.61; 9.23] | 6.54 |
| (12:00; 40) | 15.98 | [11.98; 19.98] | 14.43 |
| (12:00; 50) | 35.94 | [29.66; 42.22] | 22.65 |
| (12:00; 60) | 69.64 | [61.21; 78.07] | 30.41 |
| (12:00; 70) | 136.14 | [126.88; 145.4] | 33.39 |
| (12:00; 80) | 232.20 | [220.35; 244.05] | 42.75 |
| (12:30; 10) | 1.54 | [0.97; 2.11] | 2.04 |
| (12:30; 20) | 3.52 | [2.31; 4.73] | 4.36 |
| (12:30; 30) | 7.90 | [6.28; 9.52] | 5.83 |
| (12:30; 40) | 16.86 | [13.98; 19.74] | 10.40 |
| (12:30; 50) | 40.86 | [35.2; 46.52] | 20.44 |
| (12:30; 60) | 69.70 | [59.9; 79.5] | 35.34 |
| (12:30; 70) | 129.92 | [117.93; 141.91] | 43.27 |
| (12:30; 80) | 228.02 | [216.81; 239.23] | 40.44 |
| (01:00; 10) | 1.16 | [0.71; 1.61] | 1.63 |
| (01:00; 20) | 3.30 | [1.97; 4.63] | 4.79 |

(continued on next page)

Table D.3 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation |
|---|---|---|---|
| (01:00; 30) | 7.56 | [5.46; 9.66] | 7.56 |
| (01:00; 40) | 14.64 | [11.34; 17.94] | 11.91 |
| (01:00; 50) | 33.28 | [28.31; 38.25] | 17.94 |
| (01:00; 60) | 67.06 | [57.19; 76.93] | 35.60 |
| (01:00; 70) | 141.46 | [131.99; 150.93] | 34.15 |
| (01:00; 80) | 229.72 | [217.95; 241.49] | 42.46 |
| (01:30; 10) | 2.02 | [0.96; 3.08] | 3.83 |
| (01:30; 20) | 2.62 | [1.79; 3.45] | 2.99 |
| (01:30; 30) | 6.88 | [4.94; 8.82] | 6.98 |
| (01:30; 40) | 15.14 | [12.21; 18.07] | 10.58 |
| (01:30; 50) | 34.30 | [28.15; 40.45] | 22.18 |
| (01:30; 60) | 73.54 | [64.74; 82.34] | 31.74 |
| (01:30; 70) | 137.26 | [127.37; 147.15] | 35.68 |
| (01:30; 80) | 243.44 | [232.49; 254.39] | 39.51 |
| (02:00; 10) | 0.70 | [0.32; 1.08] | 1.39 |
| (02:00; 20) | 2.40 | [1.34; 3.46] | 3.82 |
| (02:00; 30) | 7.10 | [5.19; 9.01] | 6.90 |
| (02:00; 40) | 18.14 | [14.54; 21.74] | 12.99 |
| (02:00; 50) | 36.78 | [30.08; 43.48] | 24.17 |
| (02:00; 60) | 67.18 | [59.39; 74.97] | 28.11 |
| (02:00; 70) | 154.58 | [143; 166.16] | 41.77 |
| (02:00; 80) | 228.32 | [216.84; 239.8] | 41.41 |
| (02:30; 10) | 1.06 | [0.55; 1.57] | 1.86 |
| (02:30; 20) | 3.48 | [2.11; 4.85] | 4.93 |
| (02:30; 30) | 8.02 | [5.86; 10.18] | 7.79 |
| (02:30; 40) | 12.10 | [9.57; 14.63] | 9.12 |
| (02:30; 50) | 35.46 | [30.32; 40.6] | 18.55 |
| (02:30; 60) | 70.48 | [61.49; 79.47] | 32.42 |
| (02:30; 70) | 149.20 | [138.62; 159.78] | 38.18 |
| (02:30; 80) | 244.22 | [231.42; 257.02] | 46.18 |
| (03:00; 10) | 1.40 | [0.77; 2.03] | 2.27 |
| (03:00; 20) | 3.22 | [2.13; 4.31] | 3.92 |
| (03:00; 30) | 7.08 | [5.12; 9.04] | 7.08 |
| (03:00; 40) | 14.94 | [12.02; 17.86] | 10.55 |
| (03:00; 50) | 26.46 | [22.66; 30.26] | 13.70 |
| (03:00; 60) | 71.44 | [62.77; 80.11] | 31.28 |
| (03:00; 70) | 72.28 | [63.09; 81.47] | 33.15 |
| (03:00; 80) | 68.12 | [60.6; 75.64] | 27.12 |
| (03:30; 10) | 0.84 | [0.31; 1.37] | 1.90 |
| (03:30; 20) | 1.62 | [1.05; 2.19] | 2.05 |

(continued on next page)

Table D.3 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation |
|---|---|---|---|
| (03:30; 30) | 8.06 | [5.96; 10.16] | 7.58 |
| (03:30; 40) | 6.40 | [4.43; 8.37] | 7.12 |
| (03:30; 50) | 7.00 | [4.51; 9.49] | 8.98 |
| (03:30; 60) | 6.26 | [4.19; 8.33] | 7.47 |
| (03:30; 70) | 7.16 | [4.79; 9.53] | 8.56 |
| (03:30; 80) | 5.70 | [4.07; 7.33] | 5.88 |
| (04:00; 10) | 0.60 | [0.3; 0.9] | 1.09 |
| (04:00; 20) | 0.56 | [0.19; 0.93] | 1.34 |
| (04:00; 30) | 0.54 | [0.18; 0.9] | 1.28 |
| (04:00; 40) | 0.74 | [0.38; 1.1] | 1.29 |
| (04:00; 50) | 0.60 | [0.21; 0.99] | 1.40 |
| (04:00; 60) | 0.44 | [0.19; 0.69] | 0.91 |
| (04:00; 70) | 0.52 | [0.22; 0.82] | 1.07 |
| (04:00; 80) | 0.90 | [0.45; 1.35] | 1.61 |
| (04:30; 10) | 0.42 | [0.15; 0.69] | 0.99 |
| (04:30; 20) | 0.46 | [0.15; 0.77] | 1.11 |
| (04:30; 30) | 1.02 | [0.54; 1.5] | 1.73 |
| (04:30; 40) | 0.82 | [0.45; 1.19] | 1.35 |
| (04:30; 50) | 0.88 | [0.4; 1.36] | 1.73 |
| (04:30; 60) | 0.84 | [0.36; 1.32] | 1.74 |
| (04:30; 70) | 0.38 | [0.18; 0.58] | 0.73 |
| (04:30; 80) | 0.60 | [0.25; 0.95] | 1.28 |
| (05:00; 10) | 0.32 | [0.12; 0.52] | 0.71 |
| (05:00; 20) | 0.54 | [0.3; 0.78] | 0.86 |
| (05:00; 30) | 0.52 | [0.29; 0.75] | 0.84 |
| (05:00; 40) | 0.50 | [0.15; 0.85] | 1.27 |
| (05:00; 50) | 0.68 | [0.26; 1.1] | 1.52 |
| (05:00; 60) | 0.90 | [0.51; 1.29] | 1.40 |
| (05:00; 70) | 0.58 | [0.18; 0.98] | 1.43 |
| (05:00; 80) | 0.62 | [0.35; 0.89] | 0.99 |
| (05:30; 10) | 0.00 | [0; 0] | 0.00 |
| (05:30; 20) | 1.06 | [0.47; 1.65] | 2.11 |
| (05:30; 30) | 0.62 | [0.04; 1.2] | 2.11 |
| (05:30; 40) | 0.82 | [0.38; 1.26] | 1.57 |
| (05:30; 50) | 0.78 | [0.36; 1.2] | 1.53 |
| (05:30; 60) | 0.32 | [0.11; 0.53] | 0.77 |
| (05:30; 70) | 0.78 | [0.37; 1.19] | 1.49 |
| (05:30; 80) | 0.50 | [0.25; 0.75] | 0.91 |
| (06:00; 10) | 0.78 | [0.34; 1.22] | 1.59 |
| (06:00; 20) | 1.02 | [0.4; 1.64] | 2.25 |

(continued on next page)

Table D.3 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation |
|---|---|---|---|
| (06:00; 30) | 0.46 | [0.16; 0.76] | 1.07 |
| (06:00; 40) | 0.42 | [0.16; 0.68] | 0.95 |
| (06:00; 50) | 0.60 | [0.27; 0.93] | 1.18 |
| (06:00; 60) | 0.50 | [0.22; 0.78] | 1.02 |
| (06:00; 70) | 0.74 | [0.31; 1.17] | 1.55 |
| (06:00; 80) | 0.70 | [0.32; 1.08] | 1.39 |
| (06:30; 10) | 0.66 | [0.35; 0.97] | 1.12 |
| (06:30; 20) | 0.90 | [0.37; 1.43] | 1.93 |
| (06:30; 30) | 0.82 | [0.42; 1.22] | 1.45 |
| (06:30; 40) | 0.44 | [0.15; 0.73] | 1.03 |
| (06:30; 50) | 0.56 | [0.27; 0.85] | 1.03 |
| (06:30; 60) | 0.52 | [0.24; 0.8] | 1.01 |
| (06:30; 70) | 0.56 | [0.29; 0.83] | 0.97 |
| (06:30; 80) | 0.30 | [0.14; 0.46] | 0.58 |

Table D.4.: Simulation Results with Regard to Overtime in Minutes (Confidence Level 95%; 50 Replications)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation |
|---|---|---|---|
| (10:00; 10) | 0.82 | [0.34; 1.3] | 1.73 |
| (10:00; 20) | 1.24 | [0.47; 2.01] | 2.77 |
| (10:00; 30) | 2.80 | [1.78; 3.82] | 3.69 |
| (10:00; 40) | 12.97 | [11.71; 14.24] | 4.55 |
| (10:00; 50) | 20.69 | [19.1; 22.27] | 5.72 |
| (10:00; 60) | 32.31 | [30.75; 33.87] | 5.62 |
| (10:00; 70) | 43.23 | [41.89; 44.57] | 4.85 |
| (10:00; 80) | 52.95 | [51.33; 54.56] | 5.82 |
| (10:30; 10) | 1.05 | [0.45; 1.64] | 2.16 |
| (10:30; 20) | 1.13 | [0.62; 1.64] | 1.84 |
| (10:30; 30) | 3.36 | [2.34; 4.39] | 3.71 |
| (10:30; 40) | 11.69 | [10.37; 13] | 4.75 |
| (10:30; 50) | 22.28 | [20.96; 23.6] | 4.77 |
| (10:30; 60) | 32.73 | [31.31; 34.14] | 5.11 |
| (10:30; 70) | 41.96 | [40.6; 43.31] | 4.88 |
| (10:30; 80) | 52.45 | [51.14; 53.76] | 4.74 |

(continued on next page)

Table D.4 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation |
|---|---|---|---|
| (11:00; 10) | 7.03 | [5.79; 8.28] | 4.50 |
| (11:00; 20) | 16.06 | [14.88; 17.25] | 4.28 |
| (11:00; 30) | 26.49 | [25.26; 27.71] | 4.43 |
| (11:00; 40) | 37.98 | [36.7; 39.27] | 4.63 |
| (11:00; 50) | 46.57 | [45.11; 48.04] | 5.28 |
| (11:00; 60) | 57.38 | [55.86; 58.9] | 5.49 |
| (11:00; 70) | 68.74 | [67.42; 70.06] | 4.78 |
| (11:00; 80) | 76.75 | [75.44; 78.06] | 4.72 |
| (11:30; 10) | 6.60 | [5.4; 7.8] | 4.32 |
| (11:30; 20) | 17.65 | [16.08; 19.21] | 5.65 |
| (11:30; 30) | 26.54 | [25.24; 27.84] | 4.70 |
| (11:30; 40) | 35.95 | [34.22; 37.69] | 6.26 |
| (11:30; 50) | 46.50 | [45.17; 47.83] | 4.79 |
| (11:30; 60) | 55.74 | [54.12; 57.35] | 5.82 |
| (11:30; 70) | 67.94 | [66.58; 69.3] | 4.92 |
| (11:30; 80) | 76.98 | [75.62; 78.34] | 4.91 |
| (12:00; 10) | 7.72 | [6.46; 8.98] | 4.56 |
| (12:00; 20) | 15.64 | [14.27; 17.01] | 4.93 |
| (12:00; 30) | 27.66 | [26.24; 29.08] | 5.12 |
| (12:00; 40) | 36.19 | [34.74; 37.64] | 5.23 |
| (12:00; 50) | 46.62 | [45.1; 48.14] | 5.49 |
| (12:00; 60) | 57.10 | [55.61; 58.59] | 5.36 |
| (12:00; 70) | 66.57 | [65.13; 68.01] | 5.19 |
| (12:00; 80) | 77.26 | [75.74; 78.79] | 5.51 |
| (12:30; 10) | 6.79 | [5.7; 7.88] | 3.93 |
| (12:30; 20) | 16.70 | [15.31; 18.09] | 5.01 |
| (12:30; 30) | 27.24 | [25.82; 28.65] | 5.10 |
| (12:30; 40) | 37.22 | [36.03; 38.41] | 4.29 |
| (12:30; 50) | 48.18 | [46.61; 49.75] | 5.66 |
| (12:30; 60) | 57.71 | [55.96; 59.47] | 6.33 |
| (12:30; 70) | 66.55 | [65.26; 67.83] | 4.63 |
| (12:30; 80) | 76.65 | [75.17; 78.13] | 5.35 |
| (01:00; 10) | 7.35 | [6.13; 8.57] | 4.42 |
| (01:00; 20) | 16.74 | [15.4; 18.08] | 4.83 |
| (01:00; 30) | 27.04 | [25.89; 28.18] | 4.13 |
| (01:00; 40) | 37.19 | [35.77; 38.61] | 5.12 |
| (01:00; 50) | 48.04 | [46.43; 49.66] | 5.81 |
| (01:00; 60) | 57.50 | [56.03; 58.98] | 5.32 |
| (01:00; 70) | 67.80 | [66.52; 69.07] | 4.61 |
| (01:00; 80) | 76.74 | [75.35; 78.13] | 5.02 |

(continued on next page)

Table D.4 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation |
|---|---|---|---|
| (01:30; 10) | 7.71 | [6.18; 9.24] | 5.52 |
| (01:30; 20) | 16.48 | [15.32; 17.65] | 4.20 |
| (01:30; 30) | 27.08 | [25.66; 28.49] | 5.10 |
| (01:30; 40) | 36.77 | [35.62; 37.93] | 4.16 |
| (01:30; 50) | 46.65 | [45.27; 48.03] | 4.99 |
| (01:30; 60) | 56.73 | [55.34; 58.12] | 5.02 |
| (01:30; 70) | 67.38 | [66.1; 68.65] | 4.62 |
| (01:30; 80) | 76.45 | [75.03; 77.87] | 5.12 |
| (02:00; 10) | 6.65 | [5.47; 7.82] | 4.25 |
| (02:00; 20) | 16.43 | [15.24; 17.63] | 4.30 |
| (02:00; 30) | 27.28 | [25.88; 28.68] | 5.05 |
| (02:00; 40) | 37.59 | [36.16; 39.02] | 5.17 |
| (02:00; 50) | 47.23 | [45.66; 48.79] | 5.64 |
| (02:00; 60) | 55.76 | [54.27; 57.26] | 5.40 |
| (02:00; 70) | 67.22 | [65.95; 68.48] | 4.56 |
| (02:00; 80) | 75.94 | [74.39; 77.49] | 5.58 |
| (02:30; 10) | 7.54 | [6.26; 8.82] | 4.61 |
| (02:30; 20) | 17.53 | [16.3; 18.75] | 4.42 |
| (02:30; 30) | 28.33 | [26.84; 29.82] | 5.37 |
| (02:30; 40) | 36.89 | [35.54; 38.25] | 4.89 |
| (02:30; 50) | 47.16 | [45.95; 48.37] | 4.36 |
| (02:30; 60) | 57.04 | [56.04; 58.05] | 3.61 |
| (02:30; 70) | 67.68 | [66.41; 68.95] | 4.60 |
| (02:30; 80) | 77.93 | [76.45; 79.4] | 5.31 |
| (03:00; 10) | 6.92 | [5.82; 8.03] | 4.00 |
| (03:00; 20) | 17.22 | [15.89; 18.55] | 4.79 |
| (03:00; 30) | 26.96 | [25.63; 28.29] | 4.81 |
| (03:00; 40) | 37.14 | [35.86; 38.42] | 4.63 |
| (03:00; 50) | 46.15 | [44.56; 47.73] | 5.71 |
| (03:00; 60) | 57.01 | [55.7; 58.32] | 4.71 |
| (03:00; 70) | 67.99 | [66.68; 69.31] | 4.75 |
| (03:00; 80) | 77.15 | [75.64; 78.66] | 5.45 |
| (03:30; 10) | 6.19 | [5.15; 7.23] | 3.77 |
| (03:30; 20) | 16.05 | [14.63; 17.48] | 5.15 |
| (03:30; 30) | 26.63 | [25.21; 28.05] | 5.11 |
| (03:30; 40) | 35.51 | [34.17; 36.85] | 4.83 |
| (03:30; 50) | 46.38 | [45.13; 47.63] | 4.51 |
| (03:30; 60) | 56.18 | [54.97; 57.38] | 4.35 |
| (03:30; 70) | 67.42 | [66.28; 68.56] | 4.11 |
| (03:30; 80) | 76.56 | [75.25; 77.87] | 4.72 |

(continued on next page)

Table D.4 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation |
|---|---|---|---|
| (04:00; 10) | 8.14 | [6.91; 9.36] | 4.41 |
| (04:00; 20) | 16.46 | [15.34; 17.58] | 4.05 |
| (04:00; 30) | 27.46 | [26.26; 28.65] | 4.30 |
| (04:00; 40) | 37.15 | [35.56; 38.74] | 5.74 |
| (04:00; 50) | 46.99 | [45.74; 48.24] | 4.51 |
| (04:00; 60) | 57.60 | [56.43; 58.76] | 4.19 |
| (04:00; 70) | 66.93 | [65.84; 68.03] | 3.96 |
| (04:00; 80) | 77.74 | [76.27; 79.21] | 5.29 |
| (04:30; 10) | 8.46 | [7.22; 9.71] | 4.48 |
| (04:30; 20) | 17.81 | [16.42; 19.2] | 5.01 |
| (04:30; 30) | 27.99 | [26.68; 29.3] | 4.72 |
| (04:30; 40) | 37.58 | [36.13; 39.03] | 5.24 |
| (04:30; 50) | 48.20 | [46.92; 49.48] | 4.62 |
| (04:30; 60) | 57.98 | [56.53; 59.43] | 5.23 |
| (04:30; 70) | 67.29 | [66.17; 68.4] | 4.01 |
| (04:30; 80) | 77.95 | [76.73; 79.18] | 4.42 |
| (05:00; 10) | 7.38 | [6.27; 8.49] | 4.01 |
| (05:00; 20) | 16.98 | [15.71; 18.25] | 4.58 |
| (05:00; 30) | 27.55 | [26.27; 28.84] | 4.62 |
| (05:00; 40) | 37.09 | [35.67; 38.51] | 5.13 |
| (05:00; 50) | 47.15 | [45.84; 48.47] | 4.75 |
| (05:00; 60) | 56.31 | [54.86; 57.76] | 5.24 |
| (05:00; 70) | 66.83 | [65.48; 68.18] | 4.88 |
| (05:00; 80) | 78.09 | [76.74; 79.45] | 4.90 |
| (05:30; 10) | 0.00 | [0; 0] | 0.00 |
| (05:30; 20) | 9.58 | [7.77; 11.39] | 6.52 |
| (05:30; 30) | 19.95 | [18.07; 21.82] | 6.77 |
| (05:30; 40) | 29.43 | [27.76; 31.1] | 6.03 |
| (05:30; 50) | 39.06 | [37.07; 41.04] | 7.16 |
| (05:30; 60) | 49.09 | [47.29; 50.9] | 6.52 |
| (05:30; 70) | 59.16 | [57.64; 60.68] | 5.48 |
| (05:30; 80) | 68.89 | [66.96; 70.83] | 6.99 |
| (06:00; 10) | 4.36 | [2.52; 6.21] | 6.64 |
| (06:00; 20) | 9.38 | [6.27; 12.49] | 11.22 |
| (06:00; 30) | 9.09 | [4.98; 13.19] | 14.80 |
| (06:00; 40) | 9.39 | [4.43; 14.35] | 17.89 |
| (06:00; 50) | 18.87 | [11.81; 25.93] | 25.47 |
| (06:00; 60) | 19.63 | [11.62; 27.65] | 28.93 |
| (06:00; 70) | 26.44 | [16.56; 36.32] | 35.64 |
| (06:00; 80) | 28.21 | [17.2; 39.22] | 39.73 |

(continued on next page)

Table D.4 (continued from previous page)

| Simulation Scenario | Sample Mean | Confidence Interval | Standard Deviation |
|---|---|---|---|
| (06:30; 10) | 1.04 | [0.43; 1.64] | 2.19 |
| (06:30; 20) | 0.79 | [0.33; 1.25] | 1.67 |
| (06:30; 30) | 1.45 | [0.75; 2.15] | 2.52 |
| (06:30; 40) | 0.71 | [0.26; 1.15] | 1.60 |
| (06:30; 50) | 0.65 | [0.23; 1.07] | 1.52 |
| (06:30; 60) | 0.83 | [0.33; 1.32] | 1.79 |
| (06:30; 70) | 1.00 | [0.35; 1.65] | 2.35 |
| (06:30; 80) | 0.66 | [0.27; 1.05] | 1.41 |

## D.3. Calculation of Business Costs Induced by Additional Shipping Fee Waivers due to Service Outage Incidents

Table D.5 and Figure D.1 illustrate the numbers of shipping waivers which have to be granted due to a single outage incident with a specific duration that occurs at a certain time. Based on these numbers, Figure D.2 and Table D.6 present the corresponding business costs induced by each combination of outage duration and time of outage occurrence.

Table D.5.: Number of Additional Shipping Fee Waivers due to a Service Outage Incident (Sample Mean)

| Time of Occurrence | Outage Duration [Minutes] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
| 10:00 a.m. | 0.7 | 1.0 | 0.8 | 3.1 | 4.6 | 14.6 | 30.0 | 63.8 |
| 10:30 a.m. | 0.8 | 1.1 | 1.6 | 2.7 | 6.2 | 16.3 | 26.7 | 65.4 |
| 11:00 a.m. | 1.2 | 2.8 | 7.2 | 15.2 | 29.7 | 66.9 | 150.7 | 239.2 |
| 11:30 a.m. | 1.8 | 2.6 | 5.4 | 12.5 | 31.3 | 72.5 | 140.5 | 229.8 |
| 12:00 noon | 1.2 | 2.6 | 7.4 | 16.0 | 35.9 | 69.6 | 136.1 | 232.2 |
| 12:30 p.m. | 1.5 | 3.5 | 7.9 | 16.9 | 40.9 | 69.7 | 129.9 | 228.0 |
| 01:00 p.m. | 1.2 | 3.3 | 7.6 | 14.6 | 33.3 | 67.1 | 141.5 | 229.7 |
| 01:30 p.m. | 2.0 | 2.6 | 6.9 | 15.1 | 34.3 | 73.5 | 137.3 | 243.4 |
| 02:00 p.m. | 0.7 | 2.4 | 7.1 | 18.1 | 36.8 | 67.2 | 154.6 | 228.3 |
| 02:30 p.m. | 1.1 | 3.5 | 8.0 | 12.1 | 35.5 | 70.5 | 149.2 | 244.2 |
| 03:00 p.m. | 1.4 | 3.2 | 7.1 | 14.9 | 26.5 | 71.4 | 72.3 | 68.1 |
| 03:30 p.m. | 0.8 | 1.6 | 8.1 | 6.4 | 7.0 | 6.3 | 7.2 | 5.7 |
| 04:00 p.m. | 0.6 | 0.6 | 0.5 | 0.7 | 0.6 | 0.4 | 0.5 | 0.9 |
| 04:30 p.m. | 0.4 | 0.5 | 1.0 | 0.8 | 0.9 | 0.8 | 0.4 | 0.6 |
| 05:00 p.m. | 0.3 | 0.5 | 0.5 | 0.5 | 0.7 | 0.9 | 0.6 | 0.6 |
| 05:30 p.m. | 0.0 | 1.1 | 0.6 | 0.8 | 0.8 | 0.3 | 0.8 | 0.5 |
| 06:00 p.m. | 0.8 | 1.0 | 0.5 | 0.4 | 0.6 | 0.5 | 0.7 | 0.7 |
| 06:30 p.m. | 0.7 | 0.9 | 0.8 | 0.4 | 0.6 | 0.5 | 0.6 | 0.3 |

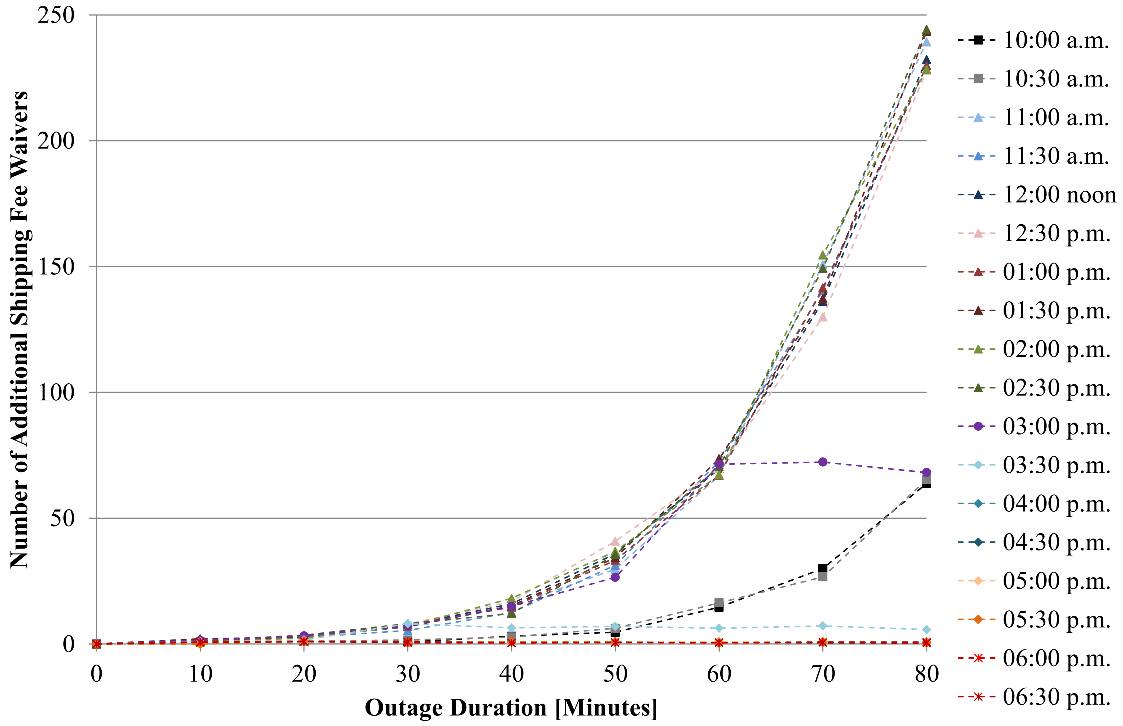


Figure D.1.: Number of Additional Shipping Fee Waivers due to a Service Outage Incident (Sample Mean)

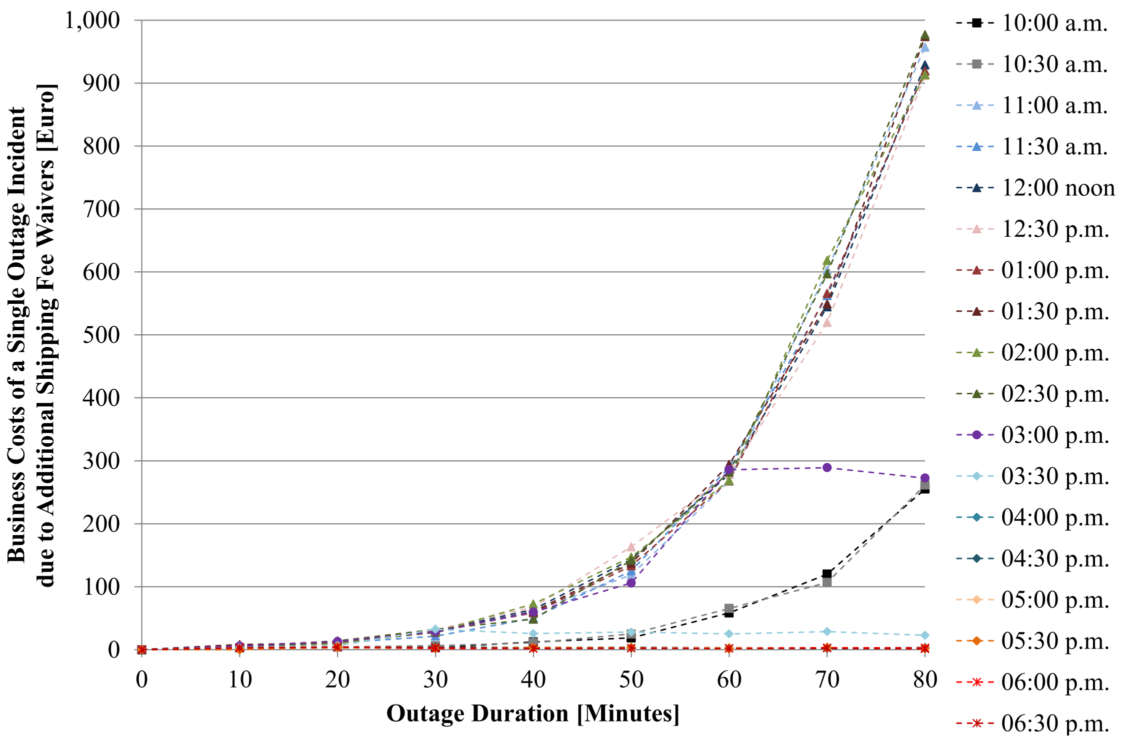


Figure D.2.: Business Costs Induced by Additional Shipping Fee Waivers due to a Service Outage Incident

Table D.6.: Business Costs Induced by Additional Shipping Fee Waivers due to a Service Outage Incident (in Euro)

| Time of Occurrence | Outage Duration [Minutes] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
| 10:00 a.m. | 2.8 | 4.1 | 3.1 | 12.2 | 18.6 | 58.4 | 119.8 | 255.1 |
| 10:30 a.m. | 3.1 | 4.2 | 6.3 | 11.0 | 25.0 | 65.4 | 106.6 | 261.4 |
| 11:00 a.m. | 4.8 | 11.2 | 28.9 | 60.7 | 118.8 | 267.5 | 602.6 | 957.0 |
| 11:30 a.m. | 7.1 | 10.5 | 21.5 | 50.0 | 125.2 | 290.1 | 562.2 | 919.3 |
| 12:00 noon | 5.0 | 10.3 | 29.7 | 63.9 | 143.8 | 278.6 | 544.6 | 928.8 |
| 12:30 p.m. | 6.2 | 14.1 | 31.6 | 67.4 | 163.4 | 278.8 | 519.7 | 912.1 |
| 01:00 p.m. | 4.6 | 13.2 | 30.2 | 58.6 | 133.1 | 268.2 | 565.8 | 918.9 |
| 01:30 p.m. | 8.1 | 10.5 | 27.5 | 60.6 | 137.2 | 294.2 | 549.0 | 973.8 |
| 02:00 p.m. | 2.8 | 9.6 | 28.4 | 72.6 | 147.1 | 268.7 | 618.3 | 913.3 |
| 02:30 p.m. | 4.2 | 13.9 | 32.1 | 48.4 | 141.8 | 281.9 | 596.8 | 976.9 |
| 03:00 p.m. | 5.6 | 12.9 | 28.3 | 59.8 | 105.8 | 285.8 | 289.1 | 272.5 |
| 03:30 p.m. | 3.4 | 6.5 | 32.2 | 25.6 | 28.0 | 25.0 | 28.6 | 22.8 |
| 04:00 p.m. | 2.4 | 2.2 | 2.2 | 3.0 | 2.4 | 1.8 | 2.1 | 3.6 |
| 04:30 p.m. | 1.7 | 1.8 | 4.1 | 3.3 | 3.5 | 3.4 | 1.5 | 2.4 |
| 05:00 p.m. | 1.3 | 2.2 | 2.1 | 2.0 | 2.7 | 3.6 | 2.3 | 2.5 |
| 05:30 p.m. | 0.0 | 4.2 | 2.5 | 3.3 | 3.1 | 1.3 | 3.1 | 2.0 |
| 06:00 p.m. | 3.1 | 4.1 | 1.8 | 1.7 | 2.4 | 2.0 | 3.0 | 2.8 |
| 06:30 p.m. | 2.6 | 3.6 | 3.3 | 1.8 | 2.2 | 2.1 | 2.2 | 1.2 |

## D.4. Calculation of Business Costs Induced by Additional Minutes of Overtime due to Service Outage Incidents

Analogously, Table D.7 and Figure D.3 illustrate the numbers of additional minutes of overtime which have to be granted due to a single outage incident with a specific duration that occurs at a certain time. Based on these numbers, Figure D.4 and Table D.8 present the corresponding business costs induced by each combination of outage duration and time of outage occurrence.

Table D.7.: Additional Minutes of Overtime due to a Service Outage Incident (Sample Mean)

| Time of Occurrence | Outage Duration [Minutes] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
| 10:00 a.m. | 0.8 | 1.2 | 2.8 | 13.0 | 20.7 | 32.3 | 43.2 | 52.9 |
| 10:30 a.m. | 1.0 | 1.1 | 3.4 | 11.7 | 22.3 | 32.7 | 42.0 | 52.5 |
| 11:00 a.m. | 7.0 | 16.1 | 26.5 | 38.0 | 46.6 | 57.4 | 68.7 | 76.8 |
| 11:30 a.m. | 6.6 | 17.6 | 26.5 | 36.0 | 46.5 | 55.7 | 67.9 | 77.0 |
| 12:00 noon | 7.7 | 15.6 | 27.7 | 36.2 | 46.6 | 57.1 | 66.6 | 77.3 |
| 12:30 p.m. | 6.8 | 16.7 | 27.2 | 37.2 | 48.2 | 57.7 | 66.5 | 76.6 |
| 01:00 p.m. | 7.4 | 16.7 | 27.0 | 37.2 | 48.0 | 57.5 | 67.8 | 76.7 |
| 01:30 p.m. | 7.7 | 16.5 | 27.1 | 36.8 | 46.6 | 56.7 | 67.4 | 76.4 |
| 02:00 p.m. | 6.6 | 16.4 | 27.3 | 37.6 | 47.2 | 55.8 | 67.2 | 75.9 |
| 02:30 p.m. | 7.5 | 17.5 | 28.3 | 36.9 | 47.2 | 57.0 | 67.7 | 77.9 |
| 03:00 p.m. | 6.9 | 17.2 | 27.0 | 37.1 | 46.1 | 57.0 | 68.0 | 77.1 |
| 03:30 p.m. | 6.2 | 16.1 | 26.6 | 35.5 | 46.4 | 56.2 | 67.4 | 76.6 |
| 04:00 p.m. | 8.1 | 16.5 | 27.5 | 37.2 | 47.0 | 57.6 | 66.9 | 77.7 |
| 04:30 p.m. | 8.5 | 17.8 | 28.0 | 37.6 | 48.2 | 58.0 | 67.3 | 78.0 |
| 05:00 p.m. | 7.4 | 17.0 | 27.6 | 37.1 | 47.2 | 56.3 | 66.8 | 78.1 |
| 05:30 p.m. | 0.0 | 9.6 | 19.9 | 29.4 | 39.1 | 49.1 | 59.2 | 68.9 |
| 06:00 p.m. | 4.4 | 9.4 | 9.1 | 9.4 | 18.9 | 19.6 | 26.4 | 28.2 |
| 06:30 p.m. | 1.0 | 0.8 | 1.5 | 0.7 | 0.6 | 0.8 | 1.0 | 0.7 |

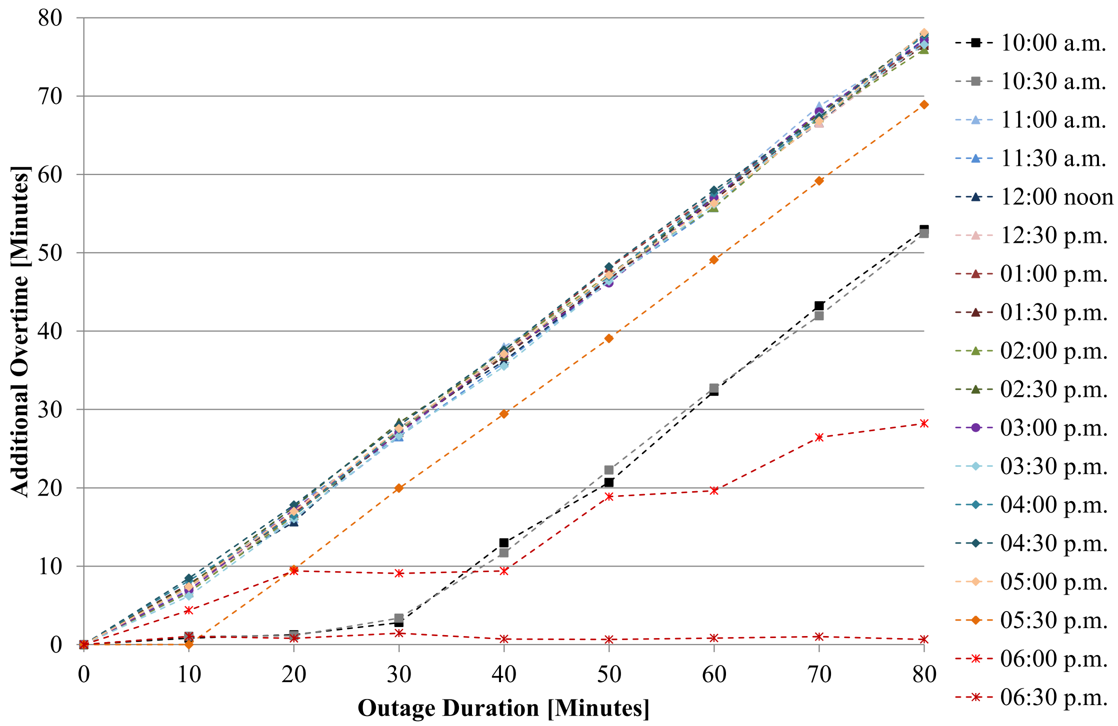


Figure D.3.: Additional Minutes of Overtime due to a Service Outage Incident (Sample Mean)

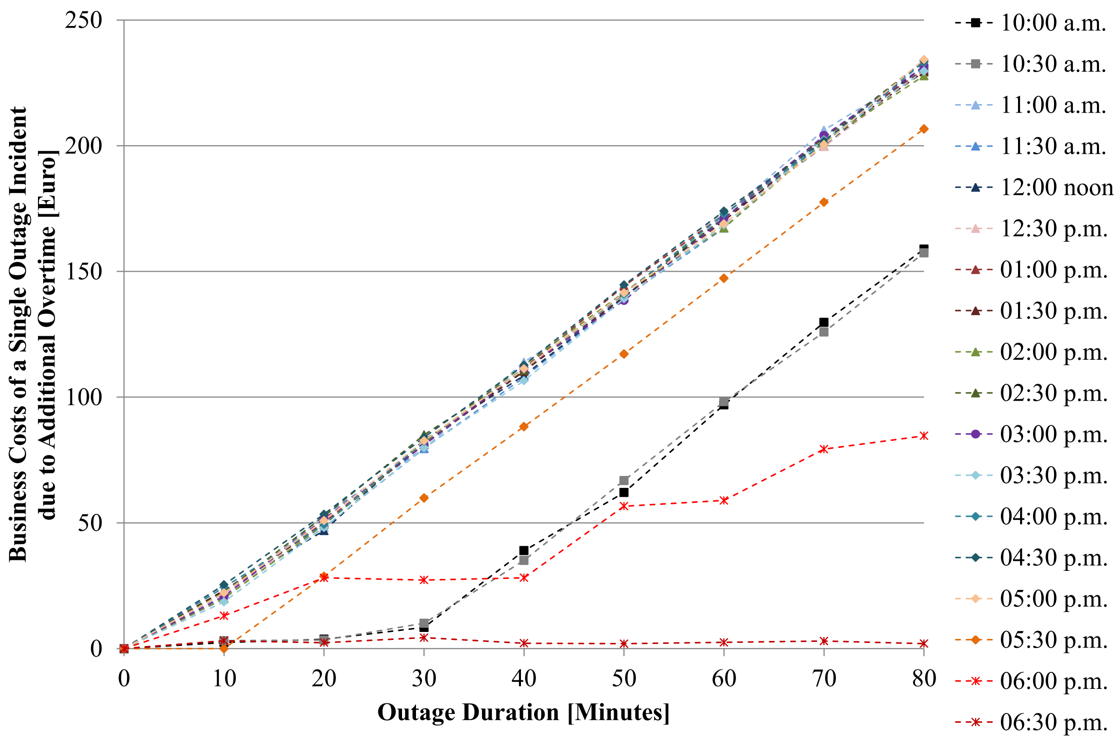


Figure D.4.: Business Costs Induced by Additional Overtime due to a Service Outage Incident

Table D.8.: Business Costs Induced by Additional Overtime due to a Service Outage Incident (in Euro)

| Time of Occurrence | Outage Duration [Minutes] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
| 10:00 a.m. | 2.4 | 3.7 | 8.4 | 38.9 | 62.1 | 96.9 | 129.7 | 158.8 |
| 10:30 a.m. | 3.1 | 3.4 | 10.1 | 35.1 | 66.8 | 98.2 | 125.9 | 157.4 |
| 11:00 a.m. | 21.1 | 48.2 | 79.5 | 113.9 | 139.7 | 172.1 | 206.2 | 230.3 |
| 11:30 a.m. | 19.8 | 52.9 | 79.6 | 107.9 | 139.5 | 167.2 | 203.8 | 230.9 |
| 12:00 noon | 23.2 | 46.9 | 83.0 | 108.6 | 139.9 | 171.3 | 199.7 | 231.8 |
| 12:30 p.m. | 20.4 | 50.1 | 81.7 | 111.7 | 144.5 | 173.1 | 199.6 | 229.9 |
| 01:00 p.m. | 22.1 | 50.2 | 81.1 | 111.6 | 144.1 | 172.5 | 203.4 | 230.2 |
| 01:30 p.m. | 23.1 | 49.4 | 81.2 | 110.3 | 139.9 | 170.2 | 202.1 | 229.3 |
| 02:00 p.m. | 19.9 | 49.3 | 81.9 | 112.8 | 141.7 | 167.3 | 201.7 | 227.8 |
| 02:30 p.m. | 22.6 | 52.6 | 85.0 | 110.7 | 141.5 | 171.1 | 203.0 | 233.8 |
| 03:00 p.m. | 20.8 | 51.7 | 80.9 | 111.4 | 138.4 | 171.0 | 204.0 | 231.4 |
| 03:30 p.m. | 18.6 | 48.2 | 79.9 | 106.5 | 139.1 | 168.5 | 202.3 | 229.7 |
| 04:00 p.m. | 24.4 | 49.4 | 82.4 | 111.5 | 141.0 | 172.8 | 200.8 | 233.2 |
| 04:30 p.m. | 25.4 | 53.4 | 84.0 | 112.7 | 144.6 | 173.9 | 201.9 | 233.9 |
| 05:00 p.m. | 22.1 | 50.9 | 82.7 | 111.3 | 141.5 | 168.9 | 200.5 | 234.3 |
| 05:30 p.m. | 0.0 | 28.7 | 59.8 | 88.3 | 117.2 | 147.3 | 177.5 | 206.7 |
| 06:00 p.m. | 13.1 | 28.2 | 27.3 | 28.2 | 56.6 | 58.9 | 79.3 | 84.6 |
| 06:30 p.m. | 3.1 | 2.4 | 4.4 | 2.1 | 1.9 | 2.5 | 3.0 | 2.0 |

# Index

## Affidavit

I hereby affirm truthfully that this thesis has been written only by the undersigned and without any assistance from third parties. Furthermore, I confirm that no resources have been used in the preparation of this thesis other than those indicated in the thesis itself, including quoted and adapted contents of publications from other authors and myself.

## Eidesstattliche Erklärung

Ich versichere wahrheitsgemäß, die Dissertation bis auf die in der Abhandlung angegebene Hilfe selbständig angefertigt, alle benutzten Hilfsmittel vollständig und genau angegeben und genau kenntlich gemacht zu haben, was aus Arbeiten anderer und aus eigenen Veröffentlichungen unverändert oder mit Abänderungen entnommen wurde.

Herrenberg, October 9th, 2015

Axel Christoph Kieninger