**TECHNISCHE UNIVERSITÄT MÜNCHEN**

**Lehrstuhl für Dienstleistungs- und Technologiemarketing**

# The Effects of Failure and Recovery on Customer Purchase Behavior

**Christian G. Heumann**

Vollständiger Abdruck der von der Fakultät für Wirtschaftswissenschaften
der Technischen Universität München zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften (Dr. rer. pol.)
genehmigten Dissertation.

| | | |
|---|---|---|
| Vorsitzender: | | Univ.-Prof. Dr. Rainer Kolisch |
| Prüfer der Dissertation: | 1. | Univ.-Prof. Dr. Florian von Wangenheim |
| | 2. | Prof. Katherine N. Lemon, Ph.D., Boston College University / USA |

Die Dissertation wurde am 22.05.2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Wirtschaftswissenschaften am 15.07.2012 angenommen.

# Summary

Prior research indicates that damaged customer relationships can be repaired on an attitudinal level, but behavioral evidence is lacking. Using a unique data set incorporating retail purchase data over three years and repeated survey measures capturing customer pre- and postfailure relationship perceptions, this thesis investigates postfailure purchase behavior and its determinants. Three important research gaps are identified and addressed in three empirical projects.

*Project I* focuses on the research question of how performance failures affect relationship outcomes. The study aims to contribute to prior research by (1) comprehensively assessing the average relationship damage of performance failures on attitudinal and behavioral outcomes, (2) clearly establishing causality, and (3) estimating the financial impact in terms of postfailure purchase behavior. Building on equity theory and Hirschman's theory of exit, voice, and loyalty, a negative causal effect of performance failure on six relationship outcomes—that is, satisfaction, repurchase intent, word-of-mouth intent, share of wallet, average transaction value, and annual customer purchase spending—is hypothesized and tested using a matching methodology combined with difference-in-differences estimation. The results suggest a negative effect of performance failure on satisfaction, word-of-mouth intent, average value per transaction, and annual purchase spending. A projection of financial effects shows that performance failure has a strong negative impact on customer equity. The quantification of the monetary impact of performance failure can help managers justify investments in service functions to enhance high-quality complaint handling capabilities.

*Project II* concentrates on the research question of how perceived justice—that is, perceptions of interactional, procedural, and distributive justice—affects postfailure behavioral loyalty. The study aims to contribute to prior research by (1) investigating the effect of perceived justice dimensions on postfailure satisfaction and purchase behavior, (2) analyzing whether satisfaction mediates the effect of justice perceptions on purchase behavior, and (3) accounting for prefailure levels of satisfaction, which are examined for potential carryover effects. Building on justice theory, hypotheses are derived and tested in a dynamic, multiple equation model with seemingly unrelated regression estimation. The results show that interactional justice plays a crucial role as it affects both postfailure satisfaction and purchase behavior. Moreover, satisfaction fully mediates the link between interactional justice and purchase behavior. In addition, carryover effects are present as prefailure outcomes turn out to be a good predictor of postfailure outcomes. No significant effects emerge for procedural and distributive justice. This suggests that elements of personal interaction in organizational response to failures are of greater relevance for postfailure loyalty than processes and compensation. The results highlight the importance of consumers' perceived justice with complaint handling as well as their responsiveness to different justice dimensions and thereby enhance understanding the drivers of postfailure purchase behavior, which helps companies evaluate complaint handling strategies and obtain guidance for resource allocation.

*Project III* strives to answer the research questions of whether and how a damaged customer relationship can be restored to its prefailure state in terms of actual purchase behavior. Overall, the study aims to make three key contributions: (1) to develop a dynamic, integrative model of postfailure purchase behavior, (2) to examine how the prefailure relationship state affects postfailure purchase behavior, and (3) to analyze the moderating effects of recovery, relationship, and marketplace characteristics on the link between failure resolution and postfailure purchase behavior. Building on the theories of relationship marketing and switching costs, a conceptual model, which suggests a contingency approach to postfailure purchase behavior, is developed and subsequently tested with hierarchical regression analysis. The results indicate that postfailure purchase behavior is influenced by failure resolution, resolution speed, switching costs, locational convenience, and prefailure affective commitment. Overall, successful and speedy failure resolution can effectively restore purchase activity to its prefailure level. However, in cases of low switching costs and low locational convenience, investments in recovery are at risk because even successful recoveries can lead to a reduction in purchase spending, particularly when delayed. Moreover, customers with high prior affective commitment significantly reduce their repurchase spending regardless of recovery success, which provides behavioral evidence of the dark side of strong customer relationships. The results reveal the relative (monetary) impact of different configurations of situational factors on recovery strategies. Managers should try to account for these contingencies and pursue an adaptive approach to complaint handling.

In summary, this dissertation contributes to an enhanced understanding of postfailure processes. From a theoretical perspective, it contributes to current knowledge by investigating the determinants of postfailure purchase behavior. More specifically, it highlights the role of the boundary conditions in a failure/recovery situation and reveals that recovery effectiveness is contingent on several factors, with some of them beyond company control. From a managerial perspective, the assessment of financial consequences resulting from altered customer purchase behavior after failure helps make investments in complaint management and recovery more accountable.

# Table of Contents

**Summary** **I**

**Table of Contents** **V**

**List of Figures** **VIII**

**List of Tables** **IX**

**List of Abbreviations** **XI**

**1 Introduction** **1**

1.1 Research Motivation . . . . . 1

1.2 Research Questions . . . . . 3

1.3 Structure of the Thesis . . . . . 7

**2 Conceptual Basis** **9**

2.1 Complaint Management and Relationship Marketing . . . . . 9

2.2 Terminology in the Research Domain of Complaint Management . . . . 12

2.2.1 Complaint Management and Service Recovery . . . . . 12

2.2.2 Failure and Complaint Types . . . . . 14

2.3 Recovery from Failure—Current Knowledge . . . . . 16

2.3.1 Postfailure Outcomes . . . . . . . . . . . . . . . . . . . . . . . 18

2.3.2 Outcome Determinants . . . . . . . . . . . . . . . . . . . . . . 24

**3 Research Design and Data 34**

3.1 Empirical Setting and Data Collection . . . . . . . . . . . . . . . . . . 34

3.2 Sample Description . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37

3.3 Measures . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 38

3.3.1 Database Measures . . . . . . . . . . . . . . . . . . . . . . . . 38

3.3.2 Survey Measures . . . . . . . . . . . . . . . . . . . . . . . . . 40

**4 The Causal Effects of Performance Failure on Relationship Outcomes 43**

4.1 Overall Background . . . . . . . . . . . . . . . . . . . . . . . . . . . . 43

4.2 Theoretical Basis and Hypotheses . . . . . . . . . . . . . . . . . . . . 47

4.3 Methodology . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 50

4.4 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 60

4.5 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 62

**5 The Effects of Perceived Justice on Postfailure Purchase Behavior 74**

5.1 Overall Background . . . . . . . . . . . . . . . . . . . . . . . . . . . . 74

5.2 Theoretical Basis and Hypotheses . . . . . . . . . . . . . . . . . . . . 80

5.3 Methodology . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 85

5.4 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 87

5.5 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 91

**6 The Moderating Effects of Recovery, Relationship, and Marketplace Characteristics on the Failure Resolution–Purchase Behavior Link 100**

6.1 Overall Background . . . . . . . . . . . . . . . . . . . . . . . . . . . . 100

6.2 Conceptual Development . . . . . . . . . . . . . . . . . . . . . . . . . . 105
6.3 Theoretical Basis and Hypotheses . . . . . . . . . . . . . . . . . . . . . 108
6.4 Methodology . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 119
6.5 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 123
6.6 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 129

**7 General Discussion and Conclusion 136**

7.1 Summary of the Key Results . . . . . . . . . . . . . . . . . . . . . . . 137
7.2 General Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . 140
7.3 Conclusion and Outlook . . . . . . . . . . . . . . . . . . . . . . . . . . 153

**References 156**

**Appendix 186**

# List of Figures

1.1 Research Questions and Contributions of the Thesis . . . . . . . . . . . 7
1.2 Structure of the Thesis . . . . . . . . . . . . . . . . . . . . . . . . . 8
2.1 General Complaint Research Framework . . . . . . . . . . . . . . . . . . 18
2.2 Organizational Responses . . . . . . . . . . . . . . . . . . . . . . . . 27
3.1 Research Design . . . . . . . . . . . . . . . . . . . . . . . . . . . . 35
4.1 PSM Implementation Steps . . . . . . . . . . . . . . . . . . . . . . . 52
4.2 Purchase Behavior Over Time for Treatment and Control Groups . . . . 61
4.3 Summary of Results of Hypotheses Tests (Project I) . . . . . . . . . . . 62
5.1 Hypothesized Model (Project II) . . . . . . . . . . . . . . . . . . . . . 81
5.2 Summary of Results of Hypotheses Tests (Project II) . . . . . . . . . . 90
6.1 Hypothesized Model (Project III) . . . . . . . . . . . . . . . . . . . . . 107
6.2 Significant Interaction Plots (Project III) . . . . . . . . . . . . . . . . . 127
6.3 Summary of Results of Hypotheses Tests (Project III) . . . . . . . . . . 128
7.1 Summary of Project Contributions to General Research Framework . . . 141

# List of Tables


2.1 Studies Examining Proactive Strategies to Offset Performance Failures . 30

4.1 Results of Matching Procedure . . . . . . . . . . . . . . . . . . . . . . 55

4.2 Group Means Before and After Matching and PRB . . . . . . . . . . . . 57

4.3 Causal Effects of Performance Failure . . . . . . . . . . . . . . . . . . 61

5.1 Prior Studies Investigating the JDs $\rightarrow$ SAT (C) $\rightarrow$ Loyalty Outcomes Link - Part I . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 78

5.2 Prior Studies Investigating the JDs $\rightarrow$ SAT (C) $\rightarrow$ Loyalty Outcomes Link - Part II . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 79

5.3 Descriptive Statistics and Correlations of the Study Variables (Project II) 88

6.1 Prior Field Studies Investigating Postfailure Behaviors . . . . . . . . . . 104

6.2 Descriptive Statistics and Correlations of the Study Variables (Project III) 120

6.3 Regression Estimates $\Delta$ Revenue . . . . . . . . . . . . . . . . . . . . 122

A.1 *Appendix*: Descriptive Sample Statistics - Part I . . . . . . . . . . . . . 187

A.2 *Appendix*: Descriptive Sample Statistics - Part II . . . . . . . . . . . . . 188

B.1 *Appendix*: Evaluation of Overall Model Fit . . . . . . . . . . . . . . . . 189

B.2 *Appendix*: Evaluation Criteria of Latent Constructs . . . . . . . . . . . 190

C.1 *Appendix*: Multi-Item Survey Measures (Project II) . . . . . . . . . . . . 192
C.2 *Appendix*: Discriminant Validity (Project II) . . . . . . . . . . . . . . . . 192
C.3 *Appendix*: Multi-Item Survey Measures (Project III) . . . . . . . . . . . . 194
C.4 *Appendix*: Discriminant Validity (Project III) . . . . . . . . . . . . . . . . 194

D.1 *Appendix*: Single-Item Survey Measures . . . . . . . . . . . . . . . . . . . 195

E.1 *Appendix*: Results of the Logistic Regression . . . . . . . . . . . . . . . . 196

## List of Abbreviations

B2B ............................... Business-to-business
B2C ............................... Business-to-consumer
CE ............................... Customer equity
CFA ............................... Confirmatory factor analysis
CFI ............................... Comparative fit index
ch. ............................... Chapter
CLV ............................... Customer lifetime value
CRM ............................... Customer relationship management
d.f. ............................... Degrees of freedom
DID ............................... Difference-in-differences
DJ ............................... Distributive justice
e.g. ............................... Exempli gratia (for example)
EFA ............................... Exploratory factor analysis
i.e. ............................... Id est (that is)
IJ ............................... Interactional justice
JDs ............................... Justice dimensions
M ............................... Mean
MD ............................... Mean difference

MI . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Monetary impact

n.s. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Not significant

NA . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Not applicable

OLS . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Ordinary least squares

p. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Page

PJ . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Procedural justice

PRB . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Percentage reduction in bias

PSM . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Propensity score matching

RMSEA . . . . . . . . . . . . . . . . . . . . . . . . . . . . Root mean square error of approximation

ROC . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Return on complaint management

RQ . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Relationship quality

SD . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Standard deviation

sec. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Section

SRMR . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Standardized root mean square residual

SUR . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Seemingly unrelatead regression

TLI . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Tucker-Lewis index

VIF . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Variance inflation factor

vs. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Versus

# 1 Introduction

## 1.1 Research Motivation

The effective management of customer complaints has become more critical than ever. Consumer power has increased with the emergence of the Internet Economy because today customers can "easily band together against companies and impose sanctions via exit and voice" (Rezabakhsh et al. 2006, p. 3). High levels of market transparency and low switching costs facilitate consumer defection "with just a few mouse clicks" (Porter 2001, p. 8). "Technologies are changing the nature of the interactions ... by amplifying the speed and impact of customer complaints" (DeVine, Lal, and Zea 2012, p. 2) and also by fostering negative customer engagement (van Doorn 2011): Increasingly, companies face the rising threat that *Unhappy Customers Strike Back on the Internet* (Tripp and Grégoire 2011), voice their dissatisfaction online, and publicly *Complain to the Masses* (Ward and Ostrom 2006). However, not only Internet companies are challenged by this decisive shift from supplier power to consumer power; rather, even traditional "retailers fear, with reason, the vengeful customer" (Mindlin 2009, p. 1). In the wake of performance failures, all kinds of businesses might have to deal with negative consequences, which may ultimately lead to greater customer churn and substantially affect a company's overall sales. Thus, today the potential damage a business-to-consumer

(B2C) company may risk with bad complaint handling and poor failure recovery can be many times greater than it was two decades ago.

That said, however, companies often disregard complaint management (Homburg and Fürst 2007), and the topic has frequently lacked managerial attention (Stauss and Seidel 2004). Such disregard is difficult to understand considering the evidence that "effective complaint handling can have a dramatic impact on customer retention rates, deflect the spread of damaging word of mouth, and improve bottom-line performance" (Tax, Brown, and Chandrashekaran 1998, p. 60). Fornell and Wernerfelt (1987) show that defensive marketing (e.g., complaint management) can lower the total marketing expenditure by substantially reducing the cost of offensive marketing (e.g., advertising). Moreover, several studies report that complaint management can be highly profitable. For example, the Technical Assistance Research Program (1986) finds that the return on complaint management (ROC) can sometimes exceed 100%. In addition, Brown (2000) notes that investments in service recovery can yield returns of 30% to 150%, and Fornell and Wernerfelt (1988) report an ROC of even 400% for some retailers.

One reason for the neglect of complaint management in many companies may be that 81% of company executives do not know the cost associated with complaints (Strativity 2007). Moreover, 78% do not assess the repurchase behavior of complainants (Stauss and Schoeler 2004). Thus, managers often have no clear picture of the monetary consequences of performance failures and cannot assess the profitability of investments in complaint management. Consequently, complaint management is frequently perceived as a mere cost center and not as a profit center (Stauss and Schoeler 2004). "Unless decision makers fully understand customer complaint behavior and can quantify the return on investment of complaint handling, they won't see the link between complaint handling and loyalty and profits, and it's unlikely they will ever allocate adequate resources for change" (Goodman 2006, p. 28).

## 1.2 Research Questions

In view of these arguments and the notion that *repurchase behavior* is the key determinant of complaint management profitability (Stauss and Seidel 2004), it is suprising that research has largely neglected investigating customer purchase behavior after failure and recovery. Thus, this dissertation pursues the overarching goal of studying postfailure purchase behavior and its determinants. In light of this, three important research gaps can be identified, and correspondingly, this thesis comprises three projects that aim to address these voids. The background and research questions for these projects are outlined as follows.

First, as a *dependent variable*, postfailure purchase behavior has scarcely been investigated (Evanschitzky, Brock, and Blut 2011; Gilly 1987; Gilly and Gelb 1982; von Wangenheim and Bayón 2007), and evidence of the behavioral consequences in terms of their monetary impact remains lacking. Until now, no work has examined postfailure purchase behavior in the popular B2C retail setting. Prior work has predominantly assessed postfailure outcomes using attitudinal loyalty and behavioral intentions as dependent variables (e.g., Maxham and Netemeyer 2002b; Tax, Brown, and Chandrashekaran 1998). However, attitudes and intentions are weak predictors of *actual purchase behavior* (e.g., Chandon, Morwitz, and Reinartz 2005; Morwitz and Schmittlein 1992). Moreover, attitudinal data cannot satisfactorily answer the question of how much to spend on a recovery and how to allocate resources. Parasuraman (2006, p. 590) notes that "extant research on service recovery is, by and large, characterized by a conspicuous dearth of analytical modeling efforts, especially in terms of providing insights that could inform the design of optimal recovery strategies." Therefore, researchers have called for a database approach to complaint management (Rust and Chung 2006) because a quantification of the effects of failure and recovery on customer purchase behavior helps trade off efforts and plan efficient and effective recovery strategies (Davidow 2003b). In addition, most extant research has produced evidence in a piecemeal manner, examining only a few outcome variables at a time. In their meta-analysis, Orsingher, Valentini, and de Angelis (2010, p. 183) strongly recommend "the inclusion of all relevant outcome variables" because, otherwise, researchers risk obtaining only a partial picture of the complex structures in complaint handling. Thus, *Project I* focuses on comprehensively assessing the average relationship damage caused by a performance failure. Overall, the project aims to answer the research question of how performance failures affect both attitudinal and behavioral outcomes, including postfailure purchase behavior.

Second, as *independent variables*, extensive research has investigated how *organizational response* (Davidow 2003a) and *perceived justice* (Orsingher, Valentini, and de Angelis 2010) affect postfailure satisfaction and loyalty outcomes. A recent meta-analysis identified more than 140 empirical studies examining such outcome determinants (Gelbrich and Roschk 2011). However, little research has studied their impact on actual purchase behavior. Evanschitzky, Brock, and Blut (2011) and Gilly and Gelb (1982) analyze the effect of satisfaction with complaint handling on purchase behavior, and von Wangenheim and Bayón (2007) assess the behavioral consequences of downgrading and denied boarding of airline customers. Nevertheless, the most agreed-on theoretical framework for explaining postfailure outcomes—that is, justice theory—has not been studied in conjunction with actual purchase behavior. More specifically, research that investigates how the dimensions of perceived justice translate into postfailure purchase behavior is nonexistent. As Blodgett, Hill, and Tax (1997, p. 187) note, "limited effort has been expended in developing a theoretical understanding of how different facets of justice affect consumers' postcomplaint behavior." Accordingly, *Project II* concentrates on answering the research question of how perceived justice—that is, perceptions of interactional, procedural, and distributive justice—affects postfailure satisfaction and purchase behavior.

Third, as *moderating variables* of the prominently studied recovery–loyalty outcome link, research has investigated the effects of failure-related characteristics (e.g., Smith, Bolton, and Wagner 1999), company characteristics (e.g., Homburg and Fürst 2005), customer characteristics (e.g., Homburg, Fürst, and Koschate 2010), relationship characteristics (e.g., Grégoire and Fisher 2006), and marketplace characteristics (e.g., Chebat, Davidow, and Borges 2011). However, overall, many relevant factors have not been studied—in particular with regard to the recovery–postfailure purchase behavior link. Gilly and Gelb (1982, p. 327) recognize that there is "no evidence that once a company response is 'satisfactory,' the degree of satisfaction affects repurchase significantly. Presumably, other market factors take precedence." More recently, Homburg, Fürst, and Koschate (2010, p. 280) have noted that competition-related market conditions play a major role in failure situations and encourage researchers "to systematically consider moderating effects" in future frameworks. In addition, in their meta-analysis, Gelbrich and Roschk (2011) contend that there is a lack of studies analyzing the moderating role of relationship aspects in a failure context. Homburg, Fürst, and Koschate (2010, p. 281) also request that "research should certainly consider the perceived quality of the business relationship." Overall, little is known about the boundary conditions and contingencies under which failure and recovery can have an effect on purchase behavior. Consequently, *Project III* examines how understudied moderating factors affect the relationship between failure resolution and postfailure purchase behavior. More specifically, the project strives to answer the research question of how recovery, relationship, and marketplace characteristics affect this link.

In summary, this dissertation aims to make three contributions to the research field of failure and recovery: First, it comprehensively assesses the effect of performance failures on key relationship outcomes on both attitudinal and behavioral levels, while clearly establishing causality. Moreover, the financial impact is quantified in terms of postfailure purchase behavior. Second, it makes a theoretical contribution by analyzing and discussing the relevance of complainants' fairness perceptions of complaint handling with respect to their effects on postfailure satisfaction and purchase behavior. Third, the thesis extends current knowledge by providing an integrative framework and dynamically assessing the moderating effects of recovery, relationship, and marketplace characteristics on the relationship between failure resolution and postfailure purchase behavior. Overall, the findings provide theoretical insights into the role of different outcome determinants and contingency factors, and draw important implications for managerial practice. Figure 1.1 summarizes the research questions and contributions of the thesis.

**Overarching Research Question:**
What are the determinants of postfailure purchase behavior?

| **Research Questions** *Project I* | **Research Questions** *Project II* | **Research Questions** *Project III* |
|---|---|---|
| What is the causal effect of performance failure on<br>▪ attitudinal, and<br>▪ behavioral<br>relationship outcomes? | How does perceived fairness with complaint handling—that is<br>▪ interactional,<br>▪ procedural, and<br>▪ distributive justice<br>affect postfailure satisfaction and purchase behavior? | How do characteristics of the<br>▪ recovery,<br>▪ relationship, and<br>▪ marketplace<br>moderate the failure resolution–purchase behavior link? |

**Key Contributions**

I Comprehensive assessment of relationship damage: Estimation of the causal effect of performance failure on attitudinal and behavioral outcomes and calculation of the financial impact.

II Assessment of the effects of perceived justice dimensions on postfailure satisfaction and purchase behavior and calculation of the financial impact.

III Development of an integrative, dynamic framework. Investigation of the moderating effects of recovery, relationship, and marketplace characteristics on the failure resolution–purchase behavior link and determination of their relative (monetary) effects.

**Figure 1.1:** Research Questions and Contributions of the Thesis

## 1.3 Structure of the Thesis

The dissertation proceeds as depicted in Figure 1.2: Following this introduction, Chapter 2 establishes the conceptual basis underlying the thesis. It describes how complaint management is anchored within the relationship marketing paradigm, delineates the terminology of the research field and explains how respective terms are used throughout the thesis, and summarizes current knowledge of the research domain. Chapter 3 describes the study design, the empirical setting, and the data collection procedure. Moreover, it presents the retail purchase data and survey measures used within the individual projects of this thesis and details the results of reliability and validity analyses. Chapter 4 comprises *Project I*, Chapter 5 represents *Project II*, and Chapter 6 focuses on *Project III*. For each project, the overall research background is first outlined. Then, the theoretical basis is introduced and the hypotheses are derived. After describing the methodology, the propositions are tested before the chapters conclude with a presentation and discussion of the results. Finally, Chapter 7 synthesizes the central findings of all the projects and elaborates on general key insights for researchers and managers.

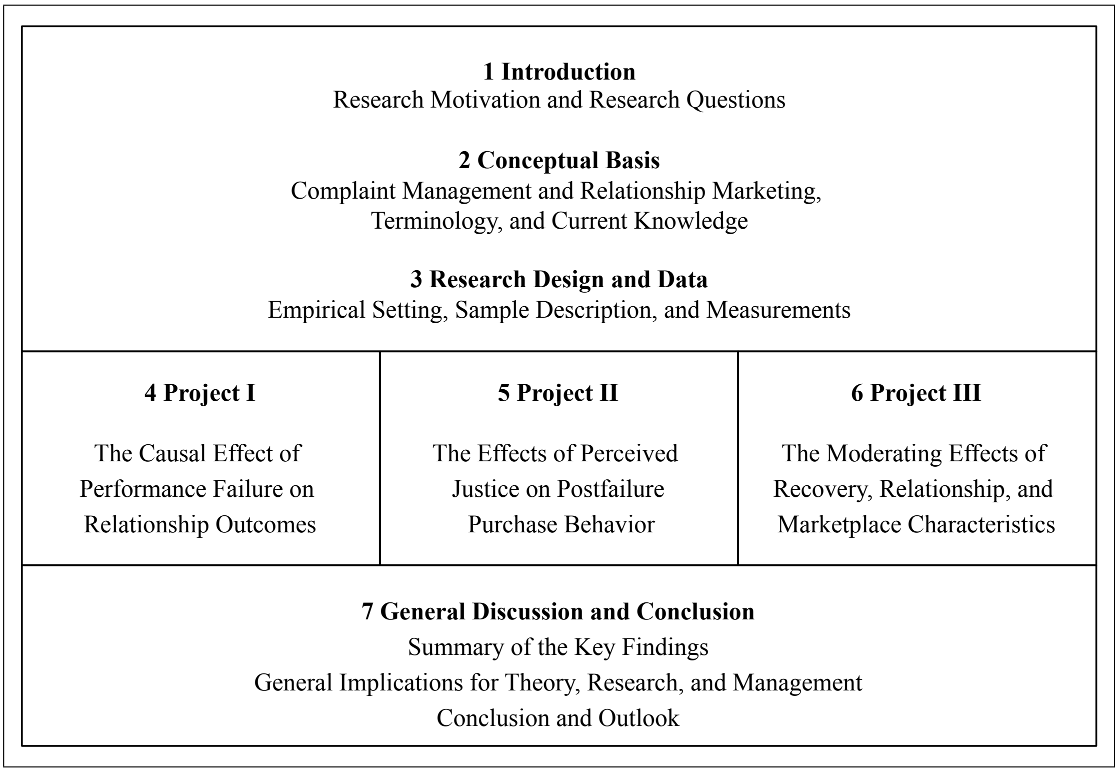


**Figure 1.2:** Structure of the Thesis

# 2 Conceptual Basis

This chapter introduces the general conceptual basis for this thesis. In the following sections, it first describes how complaint management is interrelated to the relationship marketing paradigm. Next, explanations and definitions for specific terms frequently used in the research domain of failure/recovery and complaint management are provided. Then, a description of prominently studied variables and outcomes is presented, and the current knowledge of the field that is relevant to the individual thesis projects is summarized.

## 2.1 Complaint Management and Relationship Marketing

During the past decades, marketing has experienced a paradigmatic shift from a product-focused, transactional view to a customer-centric, relational focus (e.g., Berry 2002; Grönroos 1994; Gummesson 1997; Morgan and Hunt 1994; Parvatiyar, Sheth, and Whittington 1992). The term "relationship marketing" was coined to label these new perspectives on marketing. Berry (1983, p. 25) was one of the first researchers to define relationship marketing as "attracting, maintaining, and enhancing customer relationships," and Morgan and Hunt (1994, p. 22) added to this by arguing that this comprises "all marketing activities directed towards" developing "successful relational exchanges." At the center of this development is the notion that long-term customer relationships are more profitable than short-term, transactional exchanges (e.g., Reichheld and Sasser 1990). The management of customer relationships has become increasingly perceived as critical to corporate success, and thus in the past decade, relationship marketing "experienced explosive growth" both in business and in academia (Srinivasan and Moorman 2005, p. 193). For example, a large body of research emerged that empirically tested the concept and provided evidence for its effectiveness (for a meta-analysis, see Palmatier et al. 2006). Practice widely adopted new ideas in the form of customer relationship management (CRM), implementing it with the support of software and IT systems (e.g., Payne and Frow 2005). Companies invested millions in such infrastructure (Kale 2004), and academia began researching the success factors of CRM (e.g., Jayachandran et al. 2005; Mithas, Krishnan, and Fornell 2005; Reinartz, Krafft, and Hoyer 2004). Overall, researchers and managers agree that one of the key strategic goals of marketing is to build and maintain strong customer relationships (e.g., De Wulf, Odekerken-Schröder, and Iacobucci 2001; Reichheld 2003; Harvard Business Review 2011).

Typical goals of relationship marketing include creating strong buyer–seller relationships (Dwyer, Schurr, and Oh 1987) through the development of consumer trust, commitment, and satisfaction (Hennig-Thurau, Gwinner, and Gremler 2002; Moorman, Deshpandé, and Zaltman 1993), which in turn should lead to loyal customer behavior (e.g, Verhoef 2003; Vogel, Evanschitzky, and Ramaseshan 2008). However, "trust and loyalty can ... be neither forced [nor] bought; they must be won based on positive experiences" (Stauss and Seidel 2004, p. 3). Complaint management involves handling negative events and strives to turn them into positive experiences and outcomes for customers. As such, the handling of customer complaints is strongly related to the concept of relationship marketing because both pursue similar goals. The "effective resolution of customer problems and relationship marketing are linked closely in terms of their mutual interest in customer satisfaction, trust, and commitment" (Tax, Brown, and Chandrashekaran 1998, p. 60). Complaint management strives to restore these customer perceptions after a failure, with the ultimate goal to retain customers and keep them from switching to competitors (Holloway and Beatty 2003). This goal is paramount because customer retention is considered a key driver of customer equity[1] (CE; Rust, Zeithaml, and Lemon 2000). Moreover, Reichheld and Sasser (1990) report that firms can increase their profits by 100% when customer churn is reduced by 5%, and Gupta, Lehmann, and Stuart (2004) find that a 1% improvement in the retention rate leads to a 5% increase in firm value. Thus, "complaint management is very important for the value of a customer" (Zineldin 2006, p. 435) and fulfills a critical role operating at the core of relationship marketing: customer retention. Consequently, it is no surprise that researchers view complaint management as *The Heart of CRM* (Stauss and Seidel 2004).

With the relationship marketing concept, the idea to view customers as assets emerged (e.g., Hunt and Morgan 1995; Srivastava, Shervani, and Fahey 1998). According to the *Customer Asset Management of Services* framework (Bolton, Lemon, and Verhoef 2004), relationships should be managed differently depending on how marketing instruments influence customer behavior within the relationship, thereby affecting the customer lifetime value (CLV; e.g., Hogan et al. 2002) and ultimately the financial value of the firm (Gupta, Lehmann, and Stuart 2004). Viewing customers as assets, moreover,

[1] CE is defined as "the aggregation of the expected lifetime values of a firm's entire base of existing customers and the expected future value of newly acquired customers" (Hogan et al. 2002, p. 30).

implies that resources should be allocated according to the customers' projected financial return (Mulhern 1999), and overall, marketing efforts should be directed toward maximizing the CE (Hogan, Lemon, and Rust 2002). This approach also contributes to making marketing more accountable and assessing the *Return on Marketing* (Rust, Lemon, and Zeithaml 2004). For example, Venkatesan and Kumar (2004) use the CLV as a metric for customer selection and marketing resource allocation. They show how their framework can help managers maintain and improve customer relationships and conclude that this approach leads to increased profits in future periods. However, such a value-based approach has rarely been considered within the research domain of complaint management (for an exception see Stauss and Schoeler 2004; Stauss and Seidel 2004). Thus, this thesis aims to approach this void and set a starting point by investigating behavioral consequences of failure/recovery and their financial effects.

## 2.2 Terminology in the Research Domain of Complaint Management

In more than 30 years of academic research, work in the area of failure, recovery, and complaint management has established several frequently employed frameworks, concepts, and terms. The following sections provide some general definitions and explanations of the most important terms relevant to this thesis.

### 2.2.1 Complaint Management and Service Recovery

In the literature, some authors clearly distinguish between complaint management and service failure/recovery (e.g., Michel and Meuter 2008) whereas others make no stringent distinction. How researchers decide to label these terms may depend on different factors, such as the investigated failure type, the respective study context, or the terms used in the target journal in which the authors aim to publish. For example, while clearly investigating service- *and* product-related failures, some authors use the term "service failure" for both throughout their study (e.g., Holloway and Beatty 2003; Hoffman, Kelley, and Rotalsky 1995). Naturally, within the service literature, the term "service recovery" is more prevalent, whereas in other literature streams—for example, studies in the field of relationship marketing and CRM carried out in contexts other than services—the term "complaint management" finds broader application.

In general, a *service recovery* follows a narrower conceptualization and indicates the service provider's action when something goes wrong (Grönroos 1988) or "the process by which a firm attempts to rectify a service delivery failure" (Maxham 2001, p. 12). Smith, Bolton, and Wagner (1999, p. 357) treat "service recovery as a 'bundle of resources' that an organization can employ in response to a failure." Thus, the term applies to service industries and service failures in particular. Moreover, it comprises not only reactive efforts but—in contrast with complaint management—also proactive actions (Miller, Craighead, and Karwan 2000) because in the case of a failed service encounter, firms may react immediately before the customer finds it necessary to complain (Michel and Meuter 2008).

*Complaint management* follows a broader conceptualization. On a general level, it was first described as a defensive marketing strategy (Fornell and Wernerfelt 1987). More specifically, Stauss and Seidel (2004, p. 30) contend that "complaint management encompasses the planning, execution, and controlling of all the measures taken by a firm in connection with the complaints it receives." Furthermore, as a global goal, they note that complaint management aims to increase "the profitability and competitiveness of the firm by restoring customer satisfaction, minimizing the negative effects of customer dissatisfaction on the firm, and using the indications of operational weaknesses and of market opportunities that are contained in complaints" (p. 30). Similarly, DeWitt and Brady (2003, p. 193) suggest that the objective of complaint management is "to lessen or eliminate any damage done and, ultimately, to retain a once dissatisfied customer." Although some authors suggest a proactive approach (e.g., McAlister and Erffmeyer 2003), complaint management is commonly viewed as a set of reactive strategies to resolve performance failures (e.g., Hocutt and Chakraborty 1997). Firms may "use complaint management for services as well as products" (Fornell and Wernerfelt 1988, p. 289); thus, complaint management encompasses the handling of product-related failures as well as "service recovery and involves the receipt, investigation, settlement and prevention of customer complaints and recovery of the customer" (Johnston 2001, p. 61). This thesis investigates both product- and service-related failures and thus omits the term "service recovery." When referring to recovery from failure or complaint management, I follow the broad conceptualization, which applies to various contexts (retailing in particular) and failure/complaint types.

### 2.2.2 Failure and Complaint Types

The literature uses several terms to describe the phenomenon when a customer experiences a problem at some point during the exchange relationship with a firm. For example, prior research has labeled such incidences performance lapse (Roehm and Brady 2007), performance failure (Brady et al. 2008), service failure (e.g., Hess, Ganesan, and Klein 2007), transgression (Aaker, Fournier, and Brasel 2004; Jones, Dacin, and Taylor 2011), supplier misbehavior (Ganesan et al. 2010), product-harm crisis (Klein and Dawar 2004), critical incident (van Doorn and Verhoef 2008; Johnson, Matear, and Thomson 2011), and complaint (e.g., Homburg and Fürst 2005). Depending on the study context and research design, specific aspects of such problems may be distinct, but on a general level, these events all pertain to some dissatisfying experience customers had with a firm, which potentially puts the continuance of their relationship at risk. Throughout the literature, the terms "service failure" and "complaint" are the most frequently employed. Broadly defined, the term *complaint* describes a consumer's articulation of dissatisfaction with firms and/or third-party institutions (e.g., Fornell and Wernerfelt 1987). Beyond that, it simultaneously indicates that a customer experienced some general problem with a product or service and that "the performance or the behavior of the firm does not fully comply with the customer's expectations" (Stauss and Seidel 2004, p. 16). In contrast, *service failures* only happen in the service sector and reflect "any service-related mishaps or problems (real and/or perceived) that occur during a consumer's experience with the firm" (Maxham 2001, p. 11).

With regard to the content of a failure, several typologies have been established. Researchers have developed detailed classification schemes for failures in services (Bitner, Booms, and Mohr 1994; Bitner, Booms, and Tetreault 1990; Hoffman, Kelley, and Rotalsky 1995; Keaveney 1995), retailing (Kelley, Hoffman, and Davis 1993), and online businesses (Forbes, Kelley, and Hoffmann 2006; Holloway and Beatty 2003). For example, Kelley, Hoffman, and Davis (1993) identify 15 types of retailing failures and propose three major categories based on the work of Bitner, Booms, and Tetreault (1990): (1) employee response to service and/or product failure (e.g., slow or unavailable service, product defect, repairs, packaging errors), (2) employee response to customer needs and requests (e.g., order/request, admitted customer error), and (3) unprompted and unsolicited employee actions (e.g., mischarged, embarassments, employee attention failures). On a more abstract level, the marketing literature distinguishes between two failure types: outcome and process failures (e.g., Bitner, Booms, and Tetreault 1990; Hoffman, Kelley, and Rotalsky 1995; Smith and Bolton 1998). *Outcome failures* pertain to the core offering itself (Keaveney 1995) and are concerned with what customers receive and whether the results meet their expectations. These types of failures are frequently product related (e.g., wrong or cold dish served, product malfunction), typically involve an utilitarian exchange, and may entail economic or monetary loss (Smith, Bolton, and Wagner 1999). *Process failures* pertain to "the manner in which the service is delivered" and are concerned with how consumers perceive organizational procedures and interactions (Smith and Bolton 2002, p. 10). As such, they often represent service-related problems (e.g., waiting time, failures directly attributed to the actions of service personnel, such as impoliteness), tend to occur in symbolic exchanges (Smith, Bolton, and Wagner 1999), and may lead to social or emotional loss (Gelbrich and Roschk 2011). The current research investigates both outcome and process failures.[2] Throughout this thesis, I use the term "performance failure" to holistically capture all facets of both product- and service-related problems that customers may have experienced.

## 2.3 Recovery from Failure—Current Knowledge

Within the field of failure, recovery, and complaint management, research has investigated how best to resolve performance failures and restore damaged customer relationships, and in particular, research has undertaken three attempts to consolidate empirical findings: Davidow (2003b) reviews and summarizes the findings of 57 studies on the

[2] A further description of the nature of the failures investigated in this thesis is given in section 3.2.

effects of *organizational response*. Orsingher, Valentini, and de Angelis (2010) incorporate data from 50 articles in their meta-analytic endeavor focusing on outcomes of *perceived justice*. Finally, Gelbrich and Roschk (2011) perform a literature search covering the period from 1980 to June 2009 and find 142 empirical articles relevant to their study of both *organizational response* and *perceived justice*, 87 of which reported enough statistics to be included in their meta-analysis. The majority of studies in the research domain investigate one or several relationships of the following well-established causal chain: *organizational response* → *perceived justice* → *postfailure satisfaction* → *postfailure loyalty*. Research investigates this sequence with a detailed focus on specific aspects of individual links (e.g., Evanschitzky, Brock, and Blut 2011; Smith and Bolton 2002), examines it on the whole (e.g., Gilly 1987; Homburg and Fürst 2005), and/or analyzes the central constructs in conjunction with moderating factors, such as failure-related characteristics (e.g., Smith, Bolton, and Wagner 1999), company characteristics (e.g., Homburg and Fürst 2005), customer characteristics (e.g., Homburg, Fürst, and Koschate 2010), relationship characteristics (e.g., Grégoire and Fisher 2006, 2008; Grégoire, Tripp, and Legoux 2009), and marketplace characteristics (Chebat, Davidow, and Borges 2011; Jones, Mothersbaugh, and Beatty 2000; Valenzuela, Pearson, and Epworth 2005). Figure 2.1 depicts the general research framework of the most prominently analyzed variable categories of the field.

A primary goal of failure-related research is to investigate how negative consequences resulting from performance failures can be prevented or, in other words, how such negative events can be turned into positive outcomes for both the aggrieved customer and the company. Thus, the focal variables underlying the research stream can be broken down into two basic categories: *postfailure outcomes* and the *outcome determinants* that yield

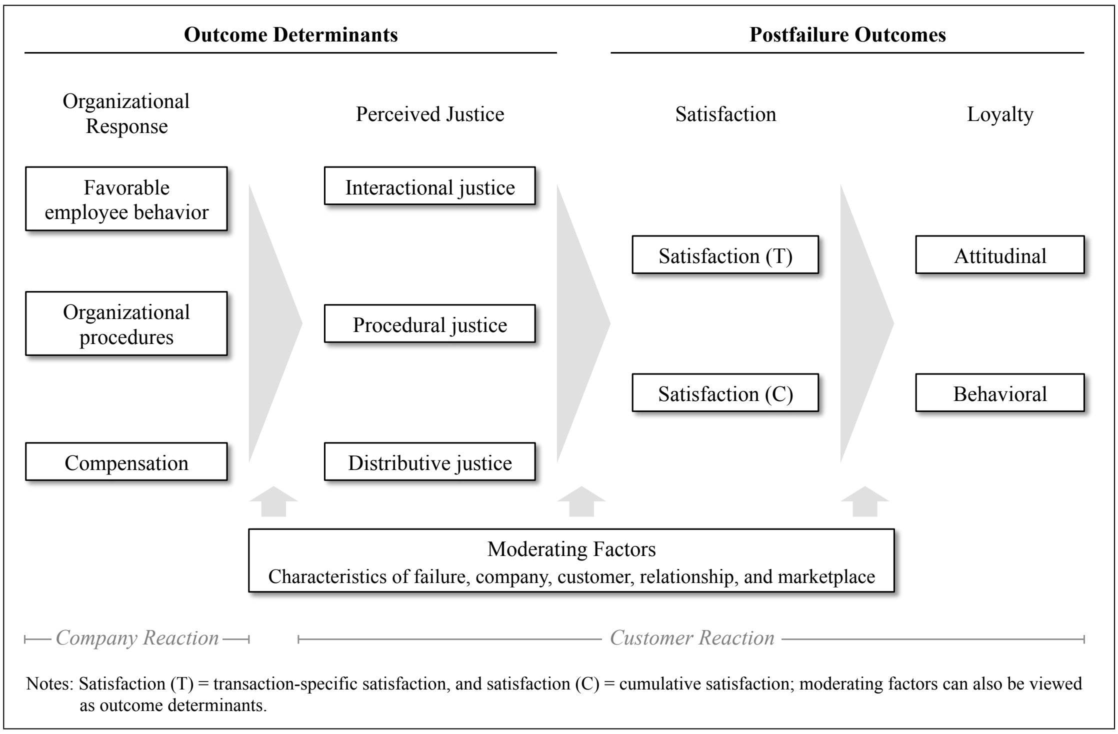


**Figure 2.1:** General Complaint Research Framework (Adapted from Davidow 2003b; Gelbrich and Roschk 2011; Homburg, Fürst, and Koschate 2010)

such results. The following sections provide a summary of the current knowledge along these two dimensions.

### 2.3.1 Postfailure Outcomes

For postfailure outcomes, research frequently investigates *satisfaction outcomes*, which act as antecedents to *loyalty outcomes*. With regard to *satisfaction outcomes*, the confirmation/disconfirmation paradigm (Oliver 1980) can serve as a theoretical basis that helps explain the formation of postfailure satisfaction. Day et al. (1981) describe satisfaction formation as a cognitive process comprising the following elements: (1) a prior basis for an evaluation (e.g., expectations), (2) an aspect of the customer–firm exchange that triggers evaluation (e.g., encounters, products, services, recoveries), and (3) a judgment of this experience (e.g., positive or negative disconfirmation of expectations). According to the confirmation/disconfirmation paradigm, customers are satisfied if their expectations of the firm's performance are met or exceeded (confirmation or positive disconfirmation), whereas dissatisfaction emerges when the company fails to meet these expectations (negative disconfirmation; Oliver 1981). In general, research has shown that postfailure satisfaction is influenced by initial disconfirmation (i.e., failure expectations vs. service performance), recovery disconfirmation (i.e., recovery expectations vs. recovery performance), and perceived justice (McCollough, Berry, and Yadav 2000).

The literature distinguishes two forms of satisfaction: transaction-specific satisfaction and cumulative satisfaction. *Transaction-specific satisfaction* refers to the judgment of single observations (Oliver 1996)–that is, it refers to a particular experience with an organization (Olsen and Johnson 2003), such as a personal encounter, a product purchase, or service consumption. In a failure context, this form of satisfaction is generally conceptualized as "the customer's evaluation of how well a ... company has handled a problem" (Orsingher, Valentini, and de Angelis 2010, p. 170) or "the degree to which the complainant perceives the company's complaint-handling performance as meeting or exceeding his or her expectations" (Homburg and Fürst 2005, p. 98). In the literature, transaction-specific satisfaction has also been labeled "recovery satisfaction" (Boshoff 1997), "satisfaction with recovery" (Maxham and Netemeyer 2002a), "satisfaction with complaint handling" (Tax, Brown, and Chandrashekaran 1998), and "complaint satisfaction" (Homburg and Fürst 2007).

*Cumulative satisfaction* represents an overall assessment of company performance; it is additive in nature and "not only takes into account the judgment of a particular recovery effort but also covers the experiences with the organization prior to these recovery efforts" (Gelbrich and Roschk 2011, p. 27). In a failure context, cumulative satisfaction refers to the degree to which complainants perceive the company's general performance as meeting or exceeding their expectations (e.g., Anderson and Sullivan 1993; Homburg and Fürst 2005). Cumulative postfailure satisfaction thus represents an evaluation on a more abstract level that captures a broader spectrum of experiences than transaction-specific satisfaction (Gelbrich and Roschk 2011; Oliver 1996). The literature frequently refers to the cumulative conceptualization as "overall customer satisfaction after the complaint" (Homburg and Fürst 2005), "overall satisfaction" (Maxham and Netemeyer 2003), or "overall firm satisfaction" (Maxham and Netemeyer 2002b).

Research on complaint handling suggests that both forms of satisfaction are related, in that transaction-specific satisfaction acts as a precursor to cumulative satisfaction. This, at least, has been demonstrated by some studies (e.g., Homburg and Fürst 2005; Maxham and Netemeyer 2002b). However, meta-analytic results find no significant effect of transaction-specific satisfaction on cumulative satisfaction (Gelbrich and Roschk 2011; Orsingher, Valentini, and de Angelis 2010). Orsingher, Valentini, and de Angelis (2010) explain that this might be because the majority of studies consider the occurrence of just one failure, which might not influence the overall satisfaction judgment. Gelbrich and Roschk (2011) offer a methodological argument and demonstrate that justice perceptions and transaction-specific satisfaction share common variance that explains cumulative satisfaction and that in a comprehensive path model, justice perceptions predominate in explaining that variance, whereas transaction-specific satisfaction does not contribute to explained variance beyond that.

In terms of antecedents to both forms of satisfaction, research has predominantly identified perceived justice dimensions (sec. 2.3.2). For example, in their meta-analyses Orsingher, Valentini, and de Angelis (2010) find that all three forms of perceived justice—interactional, procedural, and distributive justice—significantly affect transaction-specific postcomplaint satisfaction, and Gelbrich and Roschk (2011) demonstrate significant effects of all three dimensions on cumulative satisfaction. Moreover, organizational responses, such as apology, redress, timeliness, and personell attentiveness, have been frequently investigated as antecedents to satisfaction, and studies provide substantial evidence that they affect both postfailure satisfaction constructs (Davidow 2003b).

Research has also identified moderating effects in conjunction with satisfaction formation. For example, marketplace and failure characteristics play a significant role as industry type (e.g., service setting) and complaint type (e.g., monetary vs. nonmonetary complaints significantly modulate the effect of perceived justice on satisfaction; Gelbrich and Roschk 2011). Mediation analyses have revealed that the link between perceived justice and positive word of mouth is mediated by transaction-specific satisfaction (Orsingher, Valentini, and de Angelis 2010). Moreover, cumulative satisfaction fully mediates the relationships between the justice dimensions and loyalty intentions, except for the link from procedural justice to word of mouth (Gelbrich and Roschk 2011).

With regard to this thesis, *Project III* examines the transaction-specific satisfaction measure "failure resolution," which refers to how well a performance failure was resolved, for its effect on postfailure purchase behavior. *Project I* investigates cumulative postfailure satisfaction as an outcome variable, and *Project II* analyzes it for its mediating role on justice perceptions and its impact on purchase behavior.

*Loyalty outcomes* are commonly used as performance metrics in marketing research, particularly in the field of relationship marketing and service research. Traditionally, loyalty intentions, such as repurchase intent and word-of-mouth intent, serve as dependent variables in a large majority of these works. Similarly, research on failure/recovery employs these measures further distinguishing between attitudinal and behavioral outcomes. *Attitudinal outcomes* that are affected by failure, perceived justice, or postfailure satisfaction constructs include trust (DeWitt, Nguyen, and Marshall 2008; Kau and Loh 2006; Kim, Kim, and Kim 2009; Sajtos, Brodie, and Whittome 2010; Tax, Brown, and Chandrashekaran 1998), commitment (e.g., Aaker, Fournier, and Brasel 2004; Tax, Brown, and Chandrashekaran 1998; Weun, Beatty, and Jones 2004), word-of-mouth intent (e.g., Blodgett, Hill, and Tax 1997; Maxham and Netemeyer 2002b), repurchase intent[3] (e.g., Schoefer and Diamantopoulos 2008; Smith and Bolton 1998), desire for revenge (Grégoire and Fisher 2006; Grégoire, Tripp, and Legoux 2009), desire for avoidance (Grégoire, Tripp, and Legoux 2009), perceived betrayal (Grégoire and Fisher 2008; Grégoire, Tripp, and Legoux 2009), intimacy and self-connection (Aaker, Fournier, and Brasel 2004), and emotions (Chebat and Slusarczyk 2005; del Río-Lanza, Vázquez-Casielles, and Díaz-Martín 2009; DeWitt, Nguyen, and Marshall 2008; Schoefer and Diamantopoulos 2008). *Behavioral outcomes* have served as performance metrics in only a few studies. Among such measures are customer share of wallet (van Doorn and Verhoef 2008), exit behavior (Chebat and Slusarczyk 2005; Chebat, Davidow, and Borges 2011), and purchase behavior (Evanschitzky, Brock, and Blut 2011; Gilly 1987;

3 Prior research has frequently labeled self-reported intentional measures as "behavior" (e.g., Francken 1983; Gilly 1987; Gursoy, Ekiz, and Chi 2007). This thesis clearly distinguishes between *behavioral intent* and *actual behavior*. This thesis treats self-reported measures of behavioral intent as attitudinal outcomes.

Gilly and Gelb 1982; von Wangenheim and Bayón 2007).

The two most frequently studied loyalty oucomes are repurchase intent and word-of-mouth intent. These variables are well established as major satisfaction outcomes because, in general, satisfaction is considered the key mediating variable and antecedent to such loyalty measures (Oliver 1996). Although individual findings are mixed, overall, complaint research results suggest that both satisfaction constructs affect loyalty outcomes (e.g., Davidow 2000; Gelbrich and Roschk 2011; Homburg and Fürst 2005; Orsingher, Valentini, and de Angelis 2010; Weun, Beatty, and Jones 2004). However, cumulative satisfaction evaluations are viewed as better predictors of customer loyalty (e.g., Gelbrich and Roschk 2011; Olsen and Johnson 2003).

Research has also identified moderating factors that affect loyalty outcomes, such as the marketplace characteristics of switching costs (Chebat, Davidow, and Borges 2011; Jones, Mothersbaugh, and Beatty 2000) and attractiveness of alternative suppliers (Jones, Mothersbaugh, and Beatty 2000), and the relationship characteristics of commitment (Evanschitzky, Brock, and Blut 2011; Ganesan et al. 2010) and relationship quality (Grégoire, Tripp, and Legoux 2009).

In summary, the majority of studies employs self-reported, attitudinal outcome measures as dependent variables, whereas research that draws on observed, behavioral data for operationalization of dependent variables is scarce. Thus, the goal of this thesis is to confirm and expand on prior research results with regard to observed purchase behavior.

### 2.3.2 Outcome Determinants

A primary goal of failure-related research is to investigate how negative consequences resulting from performance failures can be prevented and how positive outcomes can be obtained. With regard to the *customer's perspective*, the most agreed-on framework for understanding what drives postfailure satisfaction and loyalty outcomes is *justice theory* (Orsingher, Valentini, and de Angelis 2010). Justice (or fairness) theory derives from equity theory (Adams 1965), which pertains to a person's perception of the fairness of a specific event or decision. According to this, people perceive relationships and interactions as equitable (or fair) when the ratio of their outputs (benefits) to inputs (efforts) is balanced with the output/input ratio of the other party. The research field of failure, recovery, and complaints has widely adopted the justice framework and has frequently investigated its three dimensions: interactional, procedural, and distributive justice. Interactional justice pertains to a polite and respectful way of communicating in interactions with customers (e.g., Patterson, Cowley, and Prasongsukarn 2006) and thus refers to "the manner in which people are treated during the complaint resolution process" (Blodgett, Hill, and Tax 1997, p. 189). Procedural justice reflects the perceived fairness of the complaint-handling processes (e.g., Bitner, Booms, and Tetreault 1990) and is considered fair when it is easy to access, flexible, and concluded in a convenient and timely manner (e.g., Tax, Brown, and Chandrashekaran 1998). Distributive justice "describes the fairness of the complaint outcome as the customer perceives it" (Homburg and Fürst 2005, p. 98) and mostly refers to any form of compensation, including refunds, replacements, repairs, discounts on future patronage, or some combination thereof (Blodgett, Hill, and Tax 1997).[4]

Although organizational research posits a four-factor model of justice dimensions—with interactional justice being decomposed into interpersonal and informational justice—(Colquitt et al. 2001), the model has experienced little adoption in failure-related marketing research (e.g., Ambrose, Hess, and Ganesan 2007; Kau and Loh 2006; Mattila 2006). This might also be because the distinctness of the justice dimensions has recently been called into question as a result of poor discriminant validity (Gelbrich and Roschk 2011). Thus, because consumers may be unable to clearly distinguish between individual dimensions, some researchers include perceived justice in one latent variable in their model (Blodgett, Granbois, and Walters 1993; DeWitt, Nguyen, and Marshall 2008). However, overall, the three-factor model has prevailed, and the two meta-analyses of Gelbrich and Roschk (2011) and Orsingher, Valentini, and de Angelis (2010) consolidate the findings of prior research.

As Figure 2.1 indicates, the general research framework positions the justice dimensions as precursors to postfailure outcomes.[5] Research has provided profound evidence that justice perceptions are direct antecedents to postfailure satisfaction constructs in particular (e.g., Homburg and Fürst 2005; Maxham and Netemeyer 2002b; Smith, Bolton, and Wagner 1999; Tax, Brown, and Chandrashekaran 1998). In this thesis, the justice framework finds application in *Project II* in which justice dimensions are examined for their effect on postfailure satisfaction and purchase behavior. Table 5.1 (in ch. 5) summarizes the findings of studies that investigate the role of justice dimensions and that are relevant to the project's research goals.

4 A more comprehensive description of justice theory and the three dimensions of the framework appears in *Project II* (sec. 5.2).

5 In addition, some studies show that emotions act as mediators of the perceived justice–loyalty outcomes link (Chebat and Slusarczyk 2005; Schoefer and Diamantopoulos 2008).

With regard to the *company's perspective*, several strategies can potentially offset performance failures and lead to favorable justice perceptions, satisfaction, and loyalty outcomes. The majority of prior failure-related studies has examined the effectiveness of *reactive strategies*, that is, a company's actual reactions in response to a failure or complaint. Reactive strategies are corrective actions that companies execute to restore damaged relationships to their prefailure levels (Jones, Dacin, and Taylor 2011). As such, these strategies are specifically designed to resolve a failure. Within the general research framework, these strategies are frequently conceptualized as *organizational responses*, acting as antecedents to perceived justice. Because justice perceptions are the customer's subjective assessments of "the actual action itself taken by the organization" in response to a failure or complaint (Davidow 2003b, p. 232), such organizational responses constitute the salient variables that predominantly explain justice perceptions (e.g., Homburg and Fürst 2005; Smith, Bolton, and Wagner 1999).

Research has long investigated organizational responses as a potential remedy to performance failures (e.g., Gilly and Gelb 1982; Lewis 1983). In his review, Davidow (2003b) summarizes the findings of 57 studies that investigate organizational responses. According to his conceptualization (Davidow 2000, 2003b), six dimensions (i.e., redress, apology, attentiveness, credibility, facilitation, and timeliness) need to be distinguished, whereas Estelami (2000) suggests three dimensions (i.e., compensation, employee behavior, and promptness). Gelbrich and Roschk (2011) condense these conceptualizations into three categories for their meta-analytical approach. Figure 2.2 depicts their classification and that of Davidow (2003b) and Estelami (2000).

As the general research framework (Figure 2.1) indicates, each of these organizational response dimensions corresponds with a specific justice dimension. The meta-analytic

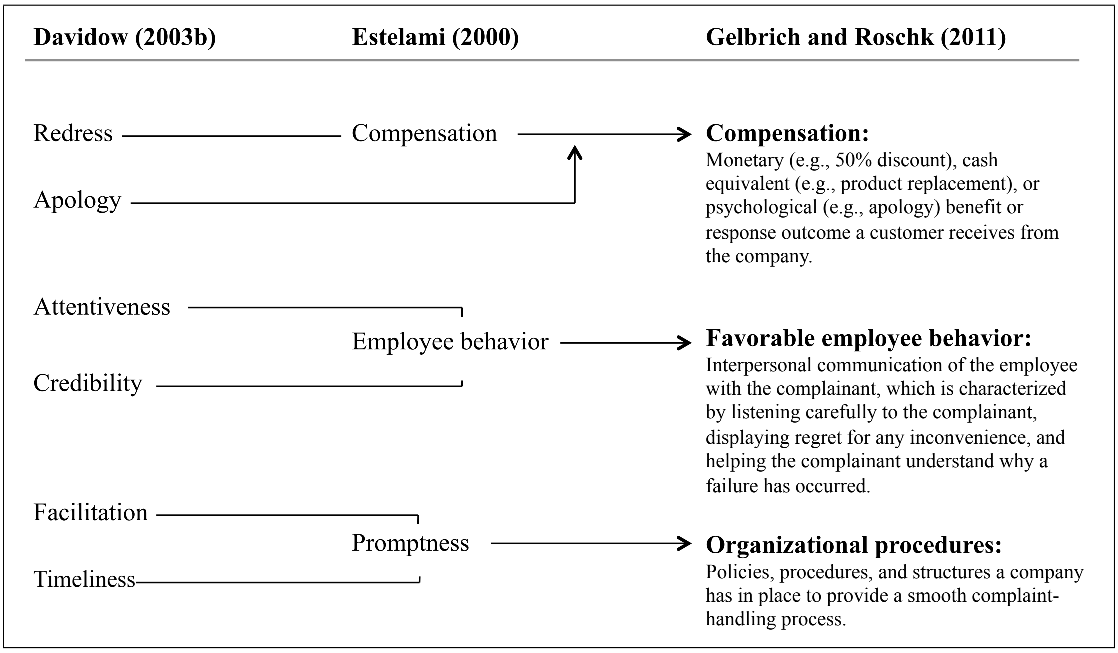


**Figure 2.2:** Organizational Responses (Source: Gelbrich and Roschk 2011)

results of Gelbrich and Roschk (2011) confirm this in that compensation is the most powerful predictor of distributive justice, favorable employee behavior is the most powerful predictor of interactional justice, and organizational procedures are the most powerful determinant of procedural justice. Moreover, the authors find that the three justice perceptions fully mediate the relationship between organizational response and cumulative satisfaction and that justice perceptions explain postfailure satisfaction better than organizational responses. For the researcher's purpose of exploring how postfailure outcomes can be determined, justice theory also offers a greater potential for generalizability of the findings because the fairness constructs operate on a more abstract level than organizational responses, which may often be subject to context-specific contingencies. For example, favorable employee behavior may play a salient role in traditional service settings and thus have a major impact on interactional justice, whereas the degree of personal interaction is typically low in online exchange relationships, may not be expected from customers, and thus may be judged accordingly in their fairness perceptions. This prevalence of the justice framework is also supported by recent studies' increasing use of the fairness theory for their research purposes; furthermore, the number of published work investigating organizational responses to failures has declined in recent years.

With the emergence of CRM, the marketing discipline shifted its focus from reactive managerial action on current customers to *proactive strategies*–that is, allocating resources to create, maintain, and enhance long-term customer loyalty behaviors (Bolton, Lemon, and Verhoef 2004). In a performance failure context, a proactive strategy is "one in which the service company invests resources in the development and strengthening of relationships with customers to attenuate the negative effects of possible service transgressions" and is characterized as a preventive action (Jones, Dacin, and Taylor 2011, p. 318). Proactive strategies typically represent more general marketing actions that are not specifically designed to recover performance failures but may still mitigate the negative impact of such transgressions. The advantage of such strategies is that they "theoretically can apply to all failures, not just those for which recovery is attempted" (Brady et al. 2008, p. 151).

Within the general nomological research framework, these strategies are oftenly conceptualized as moderating factors. For example, they are hypothesized to moderate the links between perceived justice or recovery satisfaction and postfailure outcomes (e.g., Grégoire and Fisher 2008; Tax, Brown, and Chandrashekaran 1998).

Proactive strategies that were investigated in conjunction with performance failure are, for example, brand-building strategies (Aaker, Fournier, and Brasel 2004; Brady et al.

2008), corporate social responsibility (Klein and Dawar 2004), service guarantees (Lidén and Skålén 2003), and company image (Sajtos, Brodie, and Whittome 2010). The most prominent, frequently studied proactive strategy is the development of high-quality customer relationships. Relationship marketing has long advocated that strong customer relationships lead to more favorable perceptions in failure episodes (e.g., Heskett, Sasser, and Schlesinger 1997), and research has demonstrated that this can help buffer the negative effects of performance failure or poor recovery on outcomes (e.g., Evanschitzky, Brock, and Blut 2011; Mattila 2001; Priluck 2003; Tax, Brown, and Chandrashekaran 1998). However, increasing evidence shows an "amplifying effect" of such relationship assets. That is, the establishment of strong customer relationships may backfire under certain conditions and magnify negative consequences of the failure. For example, Grégoire, Tripp, and Legoux (2009) find a "love-becomes-hate effect" for strong relationship customers on their desire for revenge and to avoid the provider, and Grégoire and Fisher (2008) detect an amplifying effect for customers who perceive low levels of fairness on their sense of betrayal. Similarly, Ganesan et al. (2010) show that affective commitment amplifies switching intentions in the case of severe opportunistic supplier behavior. The general rationale for this effect is that in strong relationships, customers perceive stronger violations of trust and an increased sense of betrayal during failure episodes (e.g., Ganesan et al. 2010; Grégoire, Tripp, and Legoux 2009). Table 2.1 summarizes the findings of studies that investigate proactive strategies and their potential to mitigate or amplify negative consequences of performance failures.

In this thesis, *Project III* investigates proactive strategies. More specifically, it examines the potential role of relationship and marketplace characteristics in enhancing or mitigating the effect of failure resolution on postfailure purchase behavior.

**Table 2.1:** Studies Examining Proactive Strategies to Offset Performance Failures

| Study | Study Context (Design) | Strategy/Variable (Effect) | Key Findings |
|---|---|---|---|
| Aaker, Fournier, and Brasel (2004) | Online photographic service (Field experiment, longitudinal) | Brand personality (Buffering for exciting brands; amplifying for sincere brands) | After transgression, relationships with sincere brands were damaged, whereas relationships with exciting brands showed signs of reinvigoration. |
| Brady et al. (2008) | Study 1: Television, amusement park<br>Study 2: Hotels, televisions, and cellular phones<br>(Experiments) | Brand equity (Buffering) | High brand equity leads to more favorable satisfaction evaluations and behavioral intentions than low brand equity. |
| Evanschitzky, Brock, and Blut (2011) | Fast-food delivery service (Field study) | Relationship characteristic: affective commitment (Buffering) | Affectively committed customers display little change in their postrecovery behavior, even after a service failure followed by an unsatisfactory recovery attempt. |
| Ganesan et al. (2010) | Studies 1, 2: Electronic equipment, B2B<br>Study 3: Fabricated metal products, industrial and commercial machinery, computer and electronic equipment, B2B<br>(Experiments) | Relationship characteristics: calculative and affective commitment (Buffering/amplifying dependent on mild/severe misbehavior) | Both calculative and affective commitment buffer suppliers against minor incidences but affective commitment amplifies the adverse effects of an supplier's flagrant opportunism in terms of switching intentions. |
| Grégoire and Fisher (2006) | Miscellaneous everyday experiences with retailers and service providers from different industries (Field study) | Relationship quality (RQ) (Retaliation depends on failure controllability) | Under low failure controllability, high RQ customers experience a lesser desire for retaliation than low RQ customers. In constrast, when high failure controllability is inferred, high RQ customers experience a greater desire for retaliation than low RQ customers. |
| Grégoire and Fisher (2008) | Airline (Field study) | Relationship quality (Amplifying) | When relationship strength is high, a violation of the fairness norm was found to have a stronger effect on the sense of betrayal experienced by customers which in turn leads to retaliation. |

| Study | Study Context (Design) | Strategy/Variable (Effect) | Key Findings |
|---|---|---|---|
| Grégoire, Tripp, and Legoux (2009) | Study 1: Online public complaining (e.g., automotive, financial services, cell phone providers, online services) (Field study, longitudinal) Study 2: Online public complaining about a restaurant (Experiment, longitudinal) | Relationship quality (Amplifying) | The desire for revenge of strong-relationship customers decreases more slowly and their desires for avoidance increases more rapidly than that of weak-relationship customers. Although high-relationship-quality customers felt more betrayed when no recovery was offered, negative perceptions are greatly attenuated by any level of recovery attempt. |
| Hess, Ganesan, and Klein (2003) | Restaurant (Experiment) | Relationship factors: quality and frequency of past experiences, expectation of continuity (Buffering) | Customers with higher expectations of relationship continuity had lower service recovery expectations and greater satisfaction with the service performance after the recovery. |
| Hess, Ganesan, and Klein (2007) | Study 1: Airline Study 2: Hotels (Experiments) | Past relationship experience: globality attributions (Buffering) | Attributions of lower globality mitigate some of the negative impact of a severe interactional failure on dissatisfaction with the organization. |
| Johnson, Matear, and Thomson (2011) | Miscellaneous, unspecified products and services (Field studies, experiment) | Strong relationships: consumer self-relevance (Amplifying) | The more self-relevant a consumer-brand relationship, the more likely are anti-brand retaliatory behaviors after the relationship ends. |
| Jones, Dacin, and Taylor (2011) | Hair salon, landscaping company (Experiment) | Extra-role interpersonal relationships (Buffering) | Creating committed extra-role interpersonal relationships between service employees and customers helps attenuate the negative effects of service transgressions. |

| Study | Study Context (Design) | Strategy/Variable (Effect) | Key Findings |
|---|---|---|---|
| Klein and Dawar (2004) | Oil company, B2C (Experiment) | Corporate social responsibility (Buffering) | Corporate social responsibility influences brand evaluations. Moreover, after product crisis it affects consumers' attributions, which in turn translate into blame for the incident that consequently influences brand evaluations and purchase intentions. |
| Lidén and Skålén (2003) | Hotels (Critical Incident Technique) | Service guarantee (Buffering) | Service guarantees can lead to more favorable perceptions of successful complaint handling and can have a mitigating effect on customer switching intentions. |
| Mattila (2001) | Restaurant (Experiment) | Relationship type: encounter/pseudo-relationships/true service relationship (Buffering) | True service relationships with the customer can mitigate the negative consequences of a failed service recovery and ensure customer loyalty. |
| Mattila (2004) | Restaurant (Experiment) | Relationship characteristic: affective commitment (Amplifying) | High affective commitment can magnify the immediate negative impact of service failures on post-recovery attitudes. Customers with lower levels of affective commitment with the service provider were more "forgiving" when the service recovery was effectively handled. |
| Priluck (2003) | Video store (Experiment) | Relational exchange (Buffering) | Relationships (vs. discrete transactions) buffer against poor product performance and mitigate negative effects of service failures in terms of satisfaction and loyalty. |

| Study | Study Context (Design) | Strategy/Variable (Effect) | Key Findings |
|---|---|---|---|
| Roehm and Brady (2007) | Sandwich catering service (Experiment) | Brand equity (Amplifying/buffering effects depending on contingencies timeliness, severity, and distraction) | High-equity brand evaluations are not adversely affected when an evaluation was made immediately after the failure and when the failure was severe or there was a distraction in the environment. |
| Sajtos, Brodie, and Whittome (2010) | Airline (Field study) | Company image and trust (Coexistence of buffering and amplifying effects) | Relationships concurrently buffer and magnify service failures. A buffering effect of company trust on customer value emerged and customer loyalty is partially protected by company image. |
| Tax, Brown, and Chandrashekaran (1998) | Miscellaneous everyday service experiences (e.g., restaurants, auto repair, banks, doctors, airlines and hotels (Field study) | Prior positive relationship experience (Buffering) | Prior positive relationship experiences mitigate, to a limited extent, the effects of poor complaint handling. |

# 3 Research Design and Data

A goal of this thesis is to test the hypotheses put forth in the individual projects by applying a descriptive research design and quantitative research methodologies. In this chapter, I describe the study design, the industry context, and the empirical basis used throughout the thesis. That is, the analyses of *Project I*, *Project II*, and *Project III* (ch. 4, 5, and 6) use data presented in the subsequent sections, which are organized as follows: First, I introduce the empirical setting in which the field study was conducted and outline the data collection procedure (sec. 3.1). Second, I provide a description of the samples (sec. 3.2). Finally, I present the measures obtained from the database and surveys and discuss the results of validity and reliability assessments for latent constructs (sec. 3.3).

## 3.1 Empirical Setting and Data Collection

Maxham and Netemeyer (2002a) suggest a need for longitudinal studies that examine the dynamics of complainant perceptions and behaviors over time. Moreover, research in the domain of failure/recovery has conducted relatively little field studies in retail settings (for an exception, see Blodgett, Granbois, and Walters 1993; Blodgett, Hill, and Tax 1997; Kelley, Hoffman, and Davis 1993). Therefore, I test my propositions in a noncontractual setting using longitudinal data from a major European retail chain. The respective company resides in a highly competitive environment and is representative of its market. The shopping category is marked by moderate purchase frequency, medium levels of involvement, and medium switching costs. The retailer's assortment includes more than 50 broad product categories, ranging from commodities to specialties. In addition, the company offers customized products and services, for which customers expect expertise and request advice from the service employees.[6]

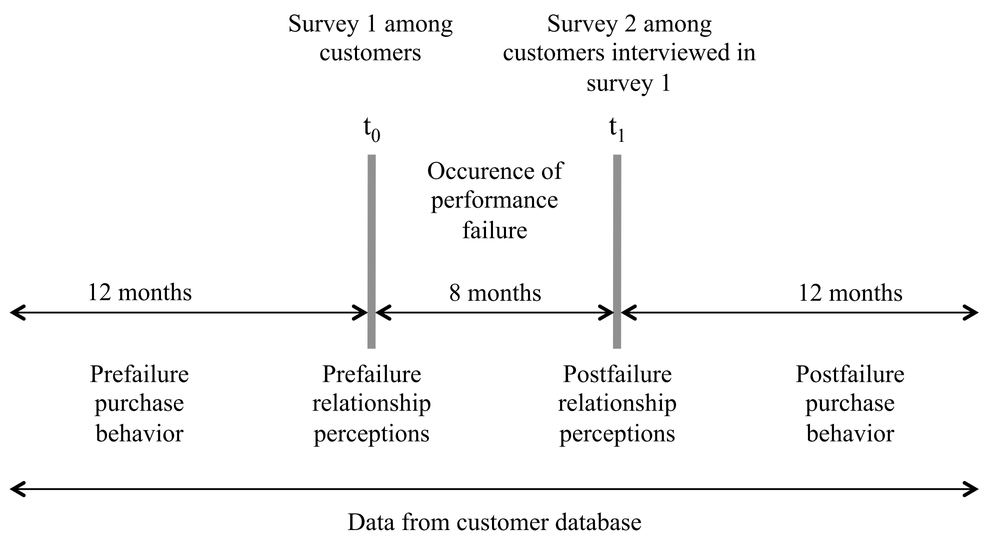


**Figure 3.1:** Research Design

The retailer granted access to its loyalty program database, which covered the period from August 2003 to October 2006. Thus, transaction information is available on a monthly basis for each customer for three-year period. A random sample of 24,015 customers was drawn for repeated surveys. Customers were contacted by mail and received a cover letter, the questionnaire, and a pre-paid return envelope. The cover letter explained the purpose of the study, assured confidentiality for the information provided,

6 Because of confidentiality agreements, no futher details about the retailer are disclosed.

and thanked the receiver for participation. As an incentive to increase the response rate, all participants were entered into a lottery to win cash prices in the range of 25 € and 500 €. The first mailing began in March 2005, and the second was sent eight months later, in December 2005. This resulted in 5688 (23.7%) responses for the first and 2435 (10.2%) responses for the second wave. Figure 3.1 depicts the research design.

The survey data was then matched to the transaction data on the basis of each customer's ID in the loyalty program. Early and late respondents were compared on key measures; no sign of nonresponse bias emerged (Armstrong and Overton 1977). Because selection effects could also introduce a bias, the data were analyzed for behavioral and attitudinal differences (e.g., purchase volume, number of transactions, interpurchase time) among the random sample (N = 24,015), participants of survey one (N = 5688), and participants of survey two (N = 2435). Comparing the random sample with the survey samples reveals some significant differences in behaviors (see Table A.2 in Appendix A). This may be due to two reasons. First, the large random sample may comprise a substantial share of occasional, low-frequency buyers. The lower average number of transactions suggests that there are less active customers in the data set, who then, expectedly, did not participate in the surveys, perhaps because they could not relate to the provider as well as regular buyers. Second, the significant differences might be due to too much statistical power. "If the sample is too large, nearly any difference, no matter how small or meaningless" will become significant (Helberg 1996, p. 2). Because absolute mean differences are small, the significant differences may also emerge due to the large sample size (N = 24,015). A comparison of behaviors and survey measures in the two survey samples does not reveal any significant differences. Thus, overall, there should be no or little selection present across the samples. In any case, because selection bias is particularly a concern with between-subjects analysis (Morimoto, White, and Newcomb 2003), which is conducted in *Project I*, a correction technique is applied that accounts for such effects. Appendix A exhibits descriptive statistics for the samples.

## 3.2 Sample Description

For the purpose of this thesis, only customers who reported having experienced a serious performance failure between the two surveys were considered.[7] In the second survey, customers needed to indicate whether they experienced a problem with the retailer since the first survey, whether they had complained about it, and how they perceived the failure resolution process and outcome. After removal of respondents who provided unusable information, 2318 questionnaires with repeated measures remained; of these 174 (7.51%) customers had complained about a major problem they experienced between the two surveys and reported their pre- and a postfailure perceptions. This group constitutes the final sample that is predominantly analyzed throughout this thesis.

Performance failures were observed over eight months (Figure 3.1). Contrary to scenario-based experimental research, failures were not artificially created, and no treatments were manipulated. Instead, performance failures occured naturally and represented product- and service-related transgressions. According to the retailer's accounts, these failures are almost equally distributed: 58% of failures were product related (e.g., product malfunction, wrong product customization, product returns and repairs), and 42% were service related (e.g., bad consultation or provision of wrong information, unfriend-

7 An exception is the analysis conducted in *Project I* (ch. 4), which generates an additional control group sample using the sample of nonfailure customers who completed both surveys (N = 2144).

liness, slow or unavailable service).[8]

The majority of the respondents in the final sample were men (76.1%), with an average age of 48.32 (SD = 12.35) years and an average relationship duration of 11.85 (SD = 6.17) years. In addition, 42.3% received vocational training, and 28.1% had a college degree. Annual incomes varied from less than 18000 € (19%), to 18000 € to 30000 € (40%), to 30000 € to 45000 € (22.3%), to greater than 45000 €. The average monthly interpurchase time was 2.20 (SD = 1.65) and the average value of transactions per month 47.09 € (SD = 36.68). The mean distance from a customer's home address to the nearest store of the retailer was 11.13 kilometers (SD = 8.51), or 6.92 miles. Further descriptive statistics for the final sample are reported in the individual projects (Table 4.2 in *Project I*, Table 5.3 in *Project II*, and (Table 6.2 in *Project III*). An overview of the descriptive statistics for the original random sample and the survey samples appears in Appendix A.

## 3.3 Measures

### 3.3.1 Database Measures

The database provided the measures for the actual customer behavior. I calculated two measures of purchase behavior, one representing the spending level of customers before they experienced the performance failure ($t_0$) and one representing their spending level after the failure/recovery ($t_1$). Similar to prior research (e.g., Bolton, Kannan, and Bramlett 2000; Mittal and Kamakura 2001; Seiders et al. 2005), the variables were op-

8 The database does not track failure type, and thus this is not available in the analyses on an individual level.

erationalized by aggregating monthly revenues over a one-year period. Use of an entire year as an aggregation basis equally accounts for seasonalities in the measures of pre- and postfailure purchase spending. This pre- and postfailure assessment of customer purchase activity facilitates analyzing whether a performance failure affected postrecovery purchase behavior by comparing it with prefailure purchase behavior. Both measures, purchase spending and lagged purchase spending, are employed in *Project II* (ch. 5).

For the purpose of *Project I* and *Project III* (ch. 4 and ch. 6), the delta of both measures–that is, the before-and-after differences–were calculated to assess the change in purchase behavior after the failure. More specifically, the difference in 12 months of post- and prefailure purchase behavior was computed ($t_1 - t_0$). This measure serves as the dependent variable and reflects the change in annual purchase spending in euros, enabling the assessment of monetary effects induced by independent variables.

In addition, two other measures were extracted from the database. First, the average value per transaction (*Project I*) was calculated by dividing annual revenue by the number of transactions per year. Second, in line with prior research (Bell, Ho, and Tang 1998), locational convenience (*Project III*) was operationalized as the distance (in kilometers) to the nearest retail store from a customer's home address. Thus, the smaller (greater) the travel distance, the greater (lesser) is the locational convenience for the respective customer. Descriptive statistics for all database measures are provided in the individual projects (Tables 4.2, 5.3 and 6.2).

### 3.3.2 Survey Measures

The survey measures for the questionnaires were developed by drawing on prior research, particularly on literature streams in the areas of retailing, failure/recovery, and relationship management. Qualitative interviews and focus-group discussions were conducted to test the initial item pool. Moreover, a pretest was run with 500 customers who did not participate in the main study, leading—after slight adjustments—to the final survey instrument. Because repeated measures were collected, the questionnaires of the first and second survey were largely similar in content, except for a few additional items that captured failure- and recovery-related consumer perceptions in the second wave (see Table D.1 in Appendix D).

*Multi-item constructs* were measured on seven-point Likert scales anchored by 1 = strongly disagree (very dissatisfied) and 7 = strongly agree (very satisfied). The constructs included the following: cumulative satisfaction, interactional justice, procedural justice, affective commitment, relationship commitment, and switching costs. *Cumulative satisfaction* was measured with three items to understand the customers' overall evaluations of their relationship experiences. In line with De Wulf, Odekerken-Schröder, and Iacobucci (2001) and Bettencourt (1997), respondents rated their relationship satisfaction, their satisfaction with the retailer, and their satisfaction relative to experiences with competing retailers. The construct is included in the analyses of *Project I* and *Project II* (ch. 4 and 5). The *interactional justice* measure consists of three items and was adapted from Homburg and Fürst (2005) and Tax, Brown, and Chandrashekaran (1998). Items that were part of the scale included perceptions of staff friendliness and courteousness exhibited during the complaint-related interaction. *Procedural justice* was also operationalized with three items and adapted from Blodgett, Hill, and Tax (1997) and Smith and Bolton (1998). The measure captures facilitation to complain, timeliness, and effort put into the process to resolve the performance failure. Both justice measures are central to *Project II*. In accordance with Fullerton (2003), *affective commitment* was measured with three items that cover emotional attachment and sense of belonging. *Relationship commitment* was measured with three items adapted from De Wulf, Odekerken-Schröder, and Iacobucci (2001) and included loyalty to the store, willingness to continue the relationship with the provider despite difficulties in reaching the store, and willingness to "go the extra mile" to maintain the relationship. However, this store loyalty item was then excluded from the instrument, because reliability and validity requirements for the construct were not met, though they improved with removal of the item. The measure for *switching costs* consists of three items from Jones, Mothersbaugh, and Beatty (2000). It captures the hassle, effort, time, and money involved with changing providers. The two forms of commitment and switching costs are part of the analyses in *Project III* (ch. 6). Appendix C provides a list of the scales.

In line with Gerbing and Anderson (1988), exploratory factor analyses (EFA) and confirmatory factor analyses (CFA) were run to assess whether the criteria for construct validity and reliability met the required thresholds (see Appendix B). All factor loadings were significant ($p < .01$), in support of convergent validity. For all constructs, Cronbach's alpha values were greater than .86 (Hair et al. 1998; Nunnally 1978), and composite reliability exceeded .87 (Bagozzi and Yi 1988). This is substantially above recommended cutoff values, demonstrating good reliability. Furthermore, discriminant validity was evaluated (Fornell and Larcker 1981) and confirmed: The average variance extracted (AVE) exceeded the square of correlations between any of the constructs.

Overall, the psychometric properties all met the recommended criteria. Appendix C provides a summary of the results.

Similar to other research on performance failures (e.g., van Doorn and Verhoef 2008), *single-item measures* were also employed to keep the survey short for the benefit of a higher response rate. According to Rossiter (2002), this is sufficient if the measured construct is "concrete singular" in the mind of the rater, meaning that it is easily and uniformly imagined. The respective measurements are purchase and word-of-mouth intent, as well as failure resolution, severity, responsibility, resolution speed, and distributive justice. The measures were adapted from prior research, which frequently employs them as single-item variables. Furthermore, because the recent results indicate good predictive validity of single-item measures (Bergkvist and Rossiter 2007), this should not be cause for concern in this research. A list of the items and their respective sources appears in Appendix D.

Several other measures were collected through the questionnaire. Share of wallet indicates the customer's share of total category purchases made at the respective retailer. Moreover, the customer and relationship characteristics of age, gender, and relationship length serve as control variables for the three projects and were also surveyed (Appendix D).

# 4 The Causal Effect of Performance Failure on Relationship Outcomes

## 4.1 Overall Background

In general, evidence shows that performance failures can have a negative impact on relationship outcomes. Several studies investigate how failures, complaints, and recoveries can affect, for example, satisfaction (e.g., Homburg and Fürst 2005), repurchase intent (e.g., Maxham and Netemeyer 2002b), word-of-mouth intent (e.g., Blodgett, Granbois, and Walters 1993), and share of wallet (e.g., van Doorn and Verhoef 2008). However, little or no work depicts the full spectrum of the negative consequences performance failures can cause. In particular, little evidence reveals the financial and behavioral consequences of performance failures. With regard to practice, a worldwide study finds that 81% of company executives do not know the cost of a customer complaint (Strativity 2007); moreover, a majority (78%) of complaint managers does not assess the repurchase behavior of complainants (Stauss and Schoeler 2004). Thus, practitioners often have no clear picture of the monetary consequences of performance failures. With regard to research, evidence of detrimental failure consequences is largely survey-based, and therefore little knowledge exists about financial and behavioral outcomes (Parasuraman 2006). In addition, the majority of existing research has produced evidence in a very piecemeal fashion, examining only a few outcome variables at a time. In their meta-analysis, Orsingher, Valentini, and de Angelis (2010, p. 183) strongly recommend "the inclusion of all relevant outcome variables" because otherwise, researchers risk obtaining only a partial picture of the complex structures in complaint handling. A comprehensive assessment of relationship damage caused by performance failures is missing; a holistic appraisal would require an assessment of both attitudinal and behavioral outcomes.

*Attitudinal outcomes* have been frequently investigated and serve as dependent variables in the majority of studies within the research domain of performance failures and complaint management. These outcomes comprise constructs such as perceived justice, satisfaction, commitment, and trust and are highly relevant because they both represent antecedents to customer loyalty and the more indirect negative consequences of performance failures, such as negative word of mouth and other forms of customer retaliation (e.g., Funches, Markley, and Davis 2009; Grégoire and Fisher 2008). These measures were collected through surveys. Similarly, *behavioral outcomes* were mostly operationalized by means of self-reported data, which helped build proxy measures for future behavior from survey items capturing, for example, purchase intentions. These outcomes represent antecedents to the more direct and monetary negative consequences of performance failures, such as exit or reduced purchase spending. Although such intentional measures are relatively easy to obtain with questionnaires, they have weak predictive power regarding actual future customer behavior (e.g., De Cannière, De Pelsmacker, and Geuens 2009; Mittal and Kamakura 2001). Moreover, they provide limited information because attitudinal and intentional data, for example, do not allow direct monetary effects of performance failures to be quantified, nor can they satisfactorily answer the questions of how much money to spend on a recovery and how to best allocate available resources.

Behavioral measures—operationalized with data that track *actual customer behavior*—can overcome these shortcomings, and therefore researchers frequently call for a database approach to complaint management (Davidow 2003b; Parasuraman 2006; Rust and Chung 2006). With the use of individual transaction information, the monetary impact resulting from altered customer purchase behavior after failure and recovery can be assessed. However, only a few studies have examined behavioral consequences in terms of purchase spending (Evanschitzky, Brock, and Blut 2011; Gilly 1987; Gilly and Gelb 1982; von Wangenheim and Bayón 2007). To date, no work has investigated postfailure purchase behavior over time in a popular B2C retail setting.

From a managerial perspective, this lack of work is surprising because postfailure purchase behavior is one of the most relevant components that determines complaint management profitability (Stauss and Schoeler 2004). On an operational level, knowledge of the effects of failure and recovery on postfailure purchase behavior enables a trade-off of failure resolution efforts as well as the planning of efficient and effective recovery strategies (Davidow 2003b; Parasuraman 2006). Moreover, assessing postfailure purchase behavior can help make the management of failures and complaint handling more accountable and thereby draw oftentimes lacking top management attention to the topic. A quantification of these effects supports strategic decision making, claims for budgets, and the justification of investments in service quality in the boardroom.

From a methodological perspective, previous research has not sufficiently considered two important aspects. First, many prior studies make causality assumptions that declining effects in outcomes are induced by performance failures. For example, Gilly (1987, p. 309) calls for "future research using alternative designs, e.g., longitudinal studies, to demonstrate causation." A longitudinal research design is recommended not only for establishing causal inference (Wooldridge 2002b) but also when studying performance failures. Prior outcome levels (e.g., prefailure satisfaction, repurchase intent) directly affect subsequent outcomes (LaBarbera and Mazursky 1983; Smith and Bolton 1998) and therefore these carryover effects should be accounted for. Second, most prior studies have used varying approaches to assess the effects of performance failures. For example, most use a (cross-sectional) *between-subjects* approach to compare a recovery group with a no-failure control group (e.g., Kau and Loh 2006). Also frequently employed is a mere *within-subjects* approach, which compares different measures from the same participant before and after a failure or recovery (e.g., Maxham and Netemeyer 2002a). To comprehensively assess the causal effect of performance failures, an integration of both approaches is necessary. As noted previously, causal inference requires a longitudinal research design. In addition, a comparison of a failure group and a no-failure control group in terms of pre- and postfailure key relationship outcome variables is necessary to obtain unbiased results. Only a few research studies meet both requirements (Maxham 2001; van Doorn and Verhoef 2008).

Against this background, the overarching goal of this project is to assess the causal effect of performance failures on key relationship outcomes. Overall, this research aims to make three key contributions: (1) to comprehensively assess the average relationship damage on both attitudinal and behavioral levels, (2) to clearly establish the causality between failure and outcomes, and (3) to estimate the financial impact in terms of postfailure purchase behavior. Unlike in prior work, longitudinal transaction and survey data from a major European retailer and a methodology that accounts for within- and between-subjects variance are employed.[9]

The remainder of the project proceeds as follows: In the next section, I elaborate on the general theoretical foundation, which serves as a basis for the investigated hypotheses. Then, I outline the research methodology and estimate the causal effect for the proposed set of outcome variables. Finally, I present the results and discuss the implications of the findings.

## 4.2 Theoretical Basis and Hypotheses

As previously mentioned, for the sake of a comprehensive assessment, two sets of post-failure relationship outcome variables are analyzed: attitudinal outcomes and behavioral outcomes. As attitudinal outcomes, I investigate satisfaction, repurchase intent, and word-of-mouth intent.[10] Several studies have demonstrated a negative effect of performance failures for satisfaction, purchase intent, and word of mouth (see Gelbrich and Roschk 2011; Orsingher, Valentini, and de Angelis 2010). As behavioral outcomes, I investigate share of wallet, average transaction value, and annual revenues. Only a few studies have linked performance failures to share of wallet (van Doorn and Verhoef 2008) and actual repurchase behavior (Evanschitzky, Brock, and Blut 2011; Gilly 1987; Gilly and Gelb 1982; von Wangenheim and Bayón 2007), and no study has investigated the average transaction value as an outcome variable. By and large, the general negative

9 For a description of the empirical setting, the data, and measures employed for this study, see Chapter 3.

10 Prior research has frequently labeled self-reported intentional measures as 'behavior' (e.g., Francken 1983; Gilly 1987; Gursoy, Ekiz, and Chi 2007). This project clearly distinguishes between *behavioral intent* and *actual behavior*. This research views measures of behavioral intent as attitudinal outcomes.

link between performance failure and relationship outcome is intuitive and well established in the literature. Thus, in the following, only a general theoretical rationale is provided that applies to individual outcomes.

Previous studies have used different theories to explain negative consequences resulting from performance failures. Several authors note that there is no single, comprehensive theory of consumer complaining behavior (e.g., Blodgett, Granbois, and Walters 1993; Goodwin and Ross 1992; Kelley and Davis 1994). Rather, several theories from different fields of study help to explain the formation of postfailure attitudes and behavior. Two of the most comprehensive theoretical foundations are Hirschman's (1970) theory of exit, voice, and loyalty and equity theory (Adams 1965). Both theories can serve to explain the detrimental effects of performance failures on attitudinal and behavioral outcomes.

Researchers frequently draw on Hirschman's (1970) theory of exit, voice, and loyalty (e.g., Singh 1990) to explain customer reactions after performance failures. It proposes three levels to predict specific consumer responses: Level I contains key outcome variables, such as negative word of mouth and exit, which in turn are functions of individual customer characteristics (Level II) and industry characteristics (Level III). This project strives to investigate the impact of performance failures on key relationship outcomes and therefore primarily builds on Level I to explain variations in key dependent constructs. More specifically, Level I describes in detail the different options for response styles of dissatisfied customers and provides three options: (1) exit, (2) voice, and (3) loyalty. For competitive firms, exit is clearly the dominant customer response to dissatisfaction (Fornell and Wernerfelt 1987). It explains a failure's negative impact on behavioral outcomes, such as purchase spending, transaction value, or share of wallet, as a failure to meet customer expectations, which is subsequently punished by customer churn or a shift in buyer patronage. The voice option is directed at management or "anyone who cares to listen" (Hirschman 1970, p. 4) and can explain the formation of word-of-mouth intent (Singh 1990).

Equity theory (Adams 1965) serves as the foundation for fairness (or justice) perceptions in service encounters (Clemmer and Schneider 1996) and helps explain customer reactions to performance failure and recovery. The concept of fairness has received widespread application in more recent studies on consumer complaints, performance failures, and recovery (for a review, see, e.g., Orsingher, Valentini, and de Angelis 2010). A customer perceives relationships and interactions as equitable (or fair) when the ratio of his or her outputs (benefits) to inputs (efforts) is balanced with the output/input ratio of the other party (Adams 1965). Customers who perceive the organizational response as unfair display lower levels in the attitudinal outcomes of satisfaction, repurchase intent, and word-of-mouth intent (Maxham and Netemeyer 2002b). In addition, after an unsuccessful failure resolution (and perceived lack of justice), negative behavioral outcomes (i.e., exit, reduction of purchase spending, customer share-of-wallet, or transaction values) are likely to follow.

Against this theoretical background, the following hypotheses are put forth for the aforementioned outcomes:

$H_1$ : *A performance failure has a negative impact on postfailure outcomes; more specifically, it has a negative effect on (a) satisfaction, (b) repurchase intent, (c) word-of-mouth intent, (d) share of wallet, (e) the average transaction value, and (f) the annual revenue.*

## 4.3 Methodology

To test the hypotheses and assess the causal effect of performance failures on key relationship outcomes, I use propensity score matching (PSM; Rosenbaum and Rubin 1983) and difference-in-differences estimation (DID). PSM is an established approach to estimate causal treatment effects (Caliendo and Kopeinig 2008) and is frequently applied in diverse fields of research, such as economics, medicine, political science, and sociology. In the past decade, it became increasingly popular (Bai 2011) and found its way into research areas of management (Campbell and Frei 2010; Xue, Hitt, and Chen 2011) and marketing (Boehm 2008; Bronnenberg, Dub, and Mela 2010; Gensler, Leeflang, and Skiera 2012; Mithas, Krishnan, and Fornell 2005; von Wangenheim and Bayón 2007). The DID estimator is typically applied to evaluate effects of treatments on relevant outcome variables (Angrist and Pischke 2009; Ashenfelter and Card 1985) and is often used in conjunction with PSM (Heckman, Ichimura, and Todd 1997).

In general, PSM is a correction strategy that attempts to reduce selection bias of treatment-effect estimates from observational studies. This is achieved by creating homogeneous, comparable samples for causal interference. For this purpose, the PSM method assembles a control group from a reservoir of nontreatment cases. In the process, each treatment recipient is matched to one "similar" nonrecipient. When a good-fitting control group has been created, posttreatment differences between treatment and control cases can be further analyzed by comparing the two groups. This methodology also helps answer the counterfactual question of how the behavior of someone who has received treatment might have developed had he or she not received the treatment (Heckman, Ichimura, and Todd 1997; Rosenbaum and Rubin 1984; Rubin 1977). By combining

PSM with the DID technique, I estimate the causal effect of the treatment, which in this study is the occurrence of a performance failure. More specifically, the analysis shows how a performance failure affects postfailure behaviors and attitudes and also answers the question of how customer behaviors and attitudes would have developed had participants not experienced a failure.

Selection bias may arise for several reasons in a performance failure research study. Systematic survey response or nonresponse may be a problem. Customers who experienced a performance failure, for example, may be more inclined to participate in a survey (to voice their dissatisfaction) than nonfailure customers. Moreover, the probability of experiencing a performance failure is not the same for all customers. Typically, a large amount of heterogeneity in individual behaviors exists in the customer base. Some customers may purchase from a provider *once a year*, whereas others purchase *once a week*. Not only are customers with dozens of encounters more likely to experience a performance failure at some point, but they are also likely to react very differently to failure and recovery attempts than those who experience a failure right after their first interaction.

Although several correction techniques can account for selection bias,[11] PSM effectively removes bias for the particular application of estimating treatment effects in observational studies (Heckman, Ichimura, and Todd 1997; Rosenbaum and Rubin 1984). Moreover, combined with DID, it is most appropriate for establishing causal inference.

[11] I also considered alternative approaches for bias correction (e.g., the Heckman selection model). When I applied Heckman's two-step procedure (Heckman 1976, 1979) to the data, the so-called Heckman correction term (or inverse Mills ratio) became significant when included in regression models. This indicated that selection bias was present, meaning that failure and nonfailure customers are different in terms of their prefailure key characteristics, and thus, these differences need to be controlled for. However, ultimately I decided to employ PSM as a correction technique because it is more suitable for estimating treatment effects in a longitudinal, nonexperimental setting.

Following the procedures in the relevant literature (Caliendo and Kopeinig 2008; von Wangenheim and Bayón 2007), I implement PSM in four steps as depicted in Figure 4.1. For a detailed description of the research design, data collection, sample description, measurements, validity and reliability analyses, please refer to Chapter 3. The measures used in the subsequent analyses are reported in section 3.3.

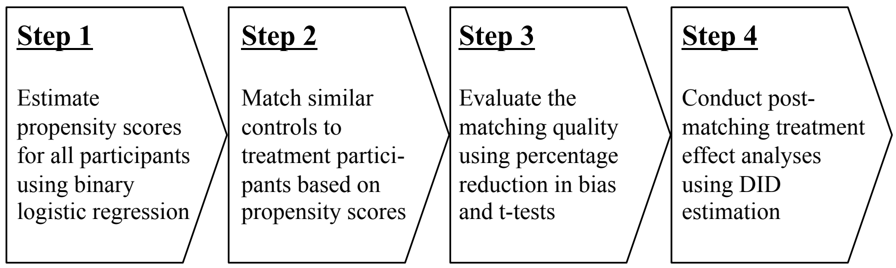


**Figure 4.1:** PSM Implementation Steps

### 4.3.1 Propensity Scores

The propensity score serves as a matching metric that determines which treatment and nontreatment participants are paired. Thus, propensity scores for all participants were estimated. This was attained by running binary logistic regression with the treatment (i.e., the occurrence of a performance failure) as a dependent variable. The selection of independent variables was driven by two considerations: First, factors that potentially increase the probability that a performance failure will occur and the corresponding customers complaints should be included to reflect the selection mechanism. As indicated previously, a failure may be dependent on the transaction frequency; moreover, the probability that a customer complains may be dependent on the transaction value. Consequently, to capture these two dimensions, the prior year's purchase spending was included in the model. Second, the goal of the matching is to obtain comparable, homogeneous treatment (failure) and control (no failure) groups; thus, key covariates in which balance is required for this particular context were also included. More specifically, as another behavioral predictor, prefailure share of wallet was included to reflect the degree of customer loyalty to the retailer. Relationship length also served as an independent variable to account for the customer's familiarity and prior experience with the provider.[12] On an attitudinal level, the prefailure perceptions included as predictors are outcomes that research studies in the domain of complaint management frequently use: satisfaction, repurchase intent, and word-of-mouth intent. In addition, the sociodemographic variables age and gender were included as relevant predictors to account for (unobserved) customer heterogeneity. The model results appear in Appendix E. Note again that the purpose of the logistic regression is not to predict performance failures; rather, the intent of the model here is only to compute the propensity scores for each participant. These are then used to perform the matching and obtain comparable customer groups, as described in the next section.

### 4.3.2 Matching

In this step, customers who experienced a failure (treatments) were matched to similar customers who did not experience a failure (controls). To accomplish this, several matching algorithms, such as nearest neighbor, kernel, and stratification, are available (Caliendo and Kopeinig 2008). In general, these various techniques yield comparable results (e.g., Heckman et al. 1998). The basic idea underlying all variants is that the treatment case is matched with a nontreatment case closest to its propensity

[12] The scope of the available transaction data includes 39 months of observed behavior. However, within the customer base, the average relationship length (in years) with the provider goes beyond that (M = 11.28, SD = 6.01) and therefore is included in the model to capture older experiences.

score. Formally, this can be described as follows: Let $\mathrm{P(X_i)}$ be subject i's propensity score. The treated subject i is then matched to the nontreated subject j, where j is $\min \left|\left|\mathrm{P(X_i)} - \mathrm{P(X_j)}\right|\right|$. The smaller the data reservoir of nontreatment participants, the more difficult it is to find appropriate matches for the treatments. In this case, nearest-neighbor matching can be disadvantageous because closest neighbors are potentially still far away. Caliper matching provides a potential remedy to this problem (Cochran and Rubin 1973). This algorithm imposes a tolerance level on the maximum propensity score distance (caliper). Thus, the nearest neighbor is only matched to the treatment case if a specified condition is met, which can be formally described as $\left|\left|\mathrm{P(X_i)} - \mathrm{P(X_j)}\right|\right| < \varepsilon$, where $\varepsilon$ is the imposed tolerance zone. As Smith and Todd (2005) note, a possible difficulty of caliper matching is determining a reasonable tolerance level. The definition of the tolerance zone significantly determines the so-called common-support region, which represents the overlap between treatment and comparison group. Consequently, a rigid tolerance definition can lead to situations in which for some treatment cases, no appropriate matching partner is available in the data. As a result, sample sizes for further analyses may be substantially reduced. In general, when studying performance failures, obtaining large sample sizes in field studies is a common problem.[13] Thus, for the purpose of obtaining as many cases as possible, the imposed tolerance levels were not as strict as, for example, Silverman (1986) suggests. Instead, a greedy matching was employed, which allows for a stepwise relaxation of the imposed tolerance level.[14] This way, the final sample size is maximized with reasonable accuracy

[13] In observational studies, failures typically cannot be artificially created and manipulated as in experiments. Instead, researchers must observe for a substantial period and let performance failures naturally occur. A combination of transaction data and repeated survey measures makes such a research design even more costly and complex; therefore, obtainable treatment samples are typically relatively small.

[14] I applied the SAS Greedy 5→1 Digit Match (Parsons 2001). As (Parsons 2001, p. 2) describes, the procedure "makes 'best' matches first and 'next-best' matches next, in a hierarchical sequence until

and goodness of matched pairs.

**Table 4.1:** Results of Matching Procedure

| Algorithm Step (5 → 1) | Completeness of Match | | Goodness of Matched Pairs | |
|---|---|---|---|---|
| | N | % Matched (of $N_T$) | Absolute Mean Difference of Matched Pairs in Propensity Score (SD) | |
| 5-digits | 21 | 12.1 | .00000385 | (.00000235) |
| 4-digits | 72 | 41.4 | .00003909 | (.00002305) |
| 3-digits | 36 | 20.7 | .00042351 | (.00025239) |
| 2-digits | 26 | 14.9 | .00281488 | (.00183808) |
| **Total** | **155** | **89.1** | **.00043572** | **(.00105000)** |

Note: Number of treatment cases $N_T$=174; number of nontreatment cases $N_N$=2144.

Table 4.1 presents the results of the matching algorithm. The procedure was able to pair 89.1% of all performance failure cases with a similar nontreatment case, which is a good quota. With a mean difference of matched pairs' propensity scores of .00043572 (.00105000), the obtained tolerance levels are within an acceptable range. In the following, samples are further evaluated to check for validity and reliability of the matching results.

### 4.3.3 Matching Quality

The matching procedure ideally balances the distribution of the relevant variables in both the treatment and the control group. Accordingly, to assess whether balance could be achieved, I evaluated the quality of the matching using two criteria. First, I assessed the percentage reduction in bias (PRB; e.g., Cochran 1968; Cochran and Rubin 1973; Rubin 1973). Therefore, for each covariate, the difference of sample means between

no more matches can be made. Best matches are those with the highest digit match on propensity score. The algorithm proceeds sequentially to the lowest digit match on propensity score. Goodness of matched pairs is defined as those with the least absolute difference in matched propensity score."

the treatment and the control group was calculated. Then, the postmatching difference, established as a fraction of the prematching difference, was subtracted from one. Similar to prior research (von Wangenheim and Bayón 2007), the PRB computation follows the adjacent formula:

$$\mathrm{PRB_n} = \left(1 - \left|\frac{\bar{x}^A_{i,n} - \bar{x}^A_{j,n}}{\bar{x}^B_{i,n} - \bar{x}^B_{j,n}}\right|\right) \times 100, \tag{4.1}$$

where

$\mathrm{PRB_n}$= the PRB for the nth predictor variable,

$\bar{x}^A_{i,n}$= the mean of the nth predictor variable for the treatment group after matching,

$\bar{x}^A_{j,n}$= the mean of the nth predictor variable for the control group after matching,

$\bar{x}^B_{i,n}$= the mean of the nth predictor variable for the treatment group before matching,

$\bar{x}^B_{j,n}$= the mean of the nth predictor variable for the control group before matching, and

N= the number of predictor variables.

Second, as Rosenbaum and Rubin (1985, p. 34) suggest, examinations of sample means "often suffice to indicate whether treated and matched control groups can be directly compared without bias due to observed covariates." Thus, I used a two-sample t-test to check whether significant differences arise in the covariate means for both groups. Before matching, differences can be expected; after matching, the covariates should be balanced across groups, and consequently no significant differences should be detectable. Evaluation of the matching quality with t-tests is particularly appropriate if statistical significance of the results in subsequent analyses is of importance (Caliendo and Kopeinig 2008).

**Table 4.2:** Group Means Before and After Matching and PRB

| | Before Matching | | | After Matching[a,b] | | |
|---|---|---|---|---|---|---|
| **Group Characteristics** | **Control** | **Performance Failure** | **MD** | **Control** | **Performance Failure** | **PRB** |
| Satisfaction ($t_0$) | 5.46 (1.16) | 4.94 (1.53) | -.51** | 4.81 (1.38) | 4.92 (1.51) | 78.8% |
| Repurchase intent ($t_0$) | 5.99 (1.12) | 5.57 (1.47) | -.42** | 5.56 (1.28) | 5.59 (1.45) | 92.9% |
| Word-of-mouth intent ($t_0$) | 6.10 (1.09) | 5.58 (1.49) | -.51** | 5.59 (1.31) | 5.60 (1.46) | 98.1% |
| % Share of wallet ($t_0$) | 74.30 (20.61) | 70.21 (24.22) | -4.09* | 68.09 (23.51) | 70.03 (24.12) | 52.6% |
| Revenue ($t_0$) | 1220.86 (1222.59) | 1325.78 (1273.17) | 104.92$^{n.s.}$ | 1221.17 (1179.37) | 1316.52 (1253.16) | 9.1% |
| % Female customers | 27.03 | 24.53 | $^{n.s.}$ | 23.29 | 24.66 | 45.2% |
| Age | 50.99 (13.69) | 48.56 (12.85) | -2.44* | 48.85 (13.68) | 48.32 (12.35) | 78.2% |
| Relationship length | 11.28 (6.01) | 11.85 (5.69) | -.57$^{n.s.}$ | 11.69 (6.21) | 11.71 (6.17) | 96.5% |
| N | 2144 | 174 | | 155 | 155 | |

$^{**}p < .01$, $^{*}p < .05$; n.s. = not significant; MD = mean difference; standard deviations are reported in parentheses.
[a]89.1% of original cases matched; [b]all mean differences are nonsignificant.

Table 4.2 displays the results of the matching evaluation. Note that the groups had significantly different prefailure characteristics before the matching. More specifically, significant mean differences are exhibited for satisfaction (MD = $-.51$; t = $-4.18$, $p < .001$), repurchase intent (MD = $-.42$; t = $-3.48$, $p < .001$), word-of-mouth intent (MD = $-.51$; t = $-4.25$, $p < .001$), share of wallet (MD = $-4.09$; t = $-2.07$, $p < .05$), and customer age (MD = $-2.44$; t = $-2.20$, $p < .05$). Thus, a comparison of both groups would have led to biased results. However, from the postmatching mean values for the treatment and control groups in the right-hand column of Table 4.2, the application of the matching procedure succeeded in removing these differences. In the postmatching state, the treatment group is similar to the control group in all key characteristics. The PRB indicates a strong reduction of bias for all previously differing predictors, and no further significant mean differences in group means could be detected. The primary goal of the matching was to create homogeneous samples; a comparable treatment and control group. Considering the results given in Table 4.2, this goal was achieved, and the overall matching quality can be judged as good.

### 4.3.4 Treatment Effect

Using the matched treatment and control groups, postmatching analyses were conducted. To estimate the average treatment effect (i.e., the causal effect of a performance failure), the conditional DID technique was applied. Research comparing the performance of matching methods and estimators has found this technique more effective than other approaches for evaluating treatment effects in nonexperimental settings (Heckman, Ichimura, and Todd 1997). In line with prior applications in marketing research (von Wangenheim and Bayón 2007), the treatment effect is estimated using the following formula:

$$\hat{\beta} = \frac{1}{n}\left\{\sum_{i\in I_1\cap S_p}^{N}(Y_{1ti} - Y_{1t'i}) - \sum_{j\in I_1\cap S_p}^{N}(Y_{0tj} - Y_{0t'j})\right\}, \tag{4.2}$$

where

$\hat{\beta}$ = the estimated treatment effect,

n = the total number of treatment cases,

$Y_{1ti} - Y_{1t'i}$ = the before-and-after difference of the treatment cases,

$Y_{0tj} - Y_{0t'j}$ = the before-and-after difference of the control cases, and

$S_p$ = the defined common support region.

This can be implemented by applying the following regression model on all matched cases:

$$(Y_{t'} - Y_t) = \alpha + \beta D + \varepsilon, \tag{4.3}$$

where

$(Y_{t'} - Y_t)$ = the before-and-after difference of the outcome variable,

$\beta$ = the estimator of the treatment effect, and

D = the treatment (where 1 = treatment case and 0 = control case).

The conditional DID technique is especially powerful because it simultaneously incorporates two important sources of variance: It reflects (1) the development over time by incorporating before-and-after differences and (2) group differences by estimating the differential effect of performance failures by comparing homogeneous failure and nonfailure customer samples. Moreover, the technique takes advantage of the longitudinal research design, as it establishes a clear causality of the treatment effect (i.e., the performance failure). The obtained results are presented in the following section.

## 4.4 Results

The research question underlying this study is whether a treatment effect of a performance failure can be observed on key outcome variables. Regarding purchase behavior, Figure 4.2 provides a first impression; it displays the time series of monthly purchase spending for the treatment and control groups. The matching removed the original bias because, in the prefailure period, both groups exhibit similar purchasing levels. Futhermore, Figure 4.2 shows that in the postfailure period, the performance failure treatment has a negative effect. That is, the performance failure group purchases significantly less than the control group.

Table 4.3 displays the overall results for the treatment effects. It shows the causal impact of a performance failure on each outcome variable for customers who experienced a failure versus those who did not experience a failure. The estimation reveals a significant negative effect for four of the six key outcome variables. The results can be interpreted, such that, in the postfailure period, customers who experienced a performance failure have satisfaction levels that are .67 rating points lower than and spend 290.10 € ($-22\%$) less than similar customers who did not experience a performance failure.

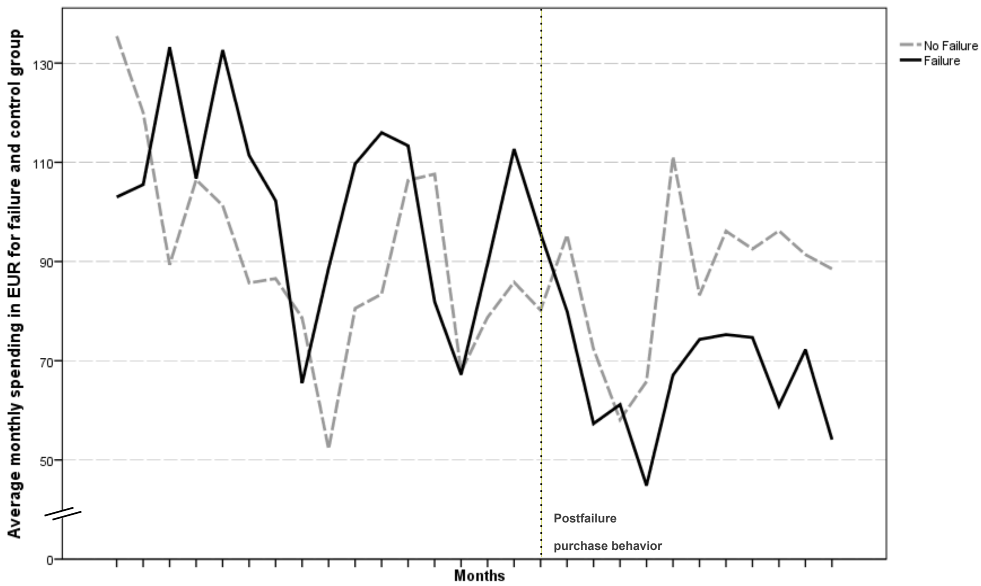


**Figure 4.2:** Purchase Behavior Over Time for Treatment and Control Groups

For the attitudinal outcomes, two of the three hypotheses find support. The strongest negative impact emerges for satisfaction ($\beta = -.67$, $p < .001$), providing support for $H_{1a}$. Word-of-mouth intent ($H_{1c}$) is also confirmed and shows the second-largest effect size ($\beta = -.49$, $p < .01$). No significant effect emerged for repurchase intent ($H_{1b}$; $\beta = -.31$, $p > .05$). However, repurchase intent is significant at the 10% level ($p = .083$) and might reach the 5% threshold if a larger sample were available. For the behavioral

**Table 4.3:** Causal Effects of Performance Failure

| **Dependent Variable** | **Difference-in-Differences** |
|---|---|
| Satisfaction | $-.67^{***}$ |
| Repurchase intent | $-.31$ |
| Word-of-mouth intent | $-.49^{**}$ |
| Share of wallet | $-5.79$ |
| Average value per transaction | $-7.52^{*}$ |
| Annual revenue | $-290.10^{*}$ |

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$.

outcomes, significant effects emerged for two of the three proposed hypotheses. Both average transaction value ($\beta = -7.52$, $p < .05$) and annual revenues ($\beta = -290.10$, $p < .05$) are significant on the 5% level. This confirms $H_{1e}$ and $H_{1f}$. No support was found for share of wallet ($H_{1d}$; $\beta = -5.79$, $p > .05$); however, it is significant at the 10% level and almost reaches the 5% threshold with a $p$-value of .052. Figure 5.2 summarizes the results of the hypotheses tests.

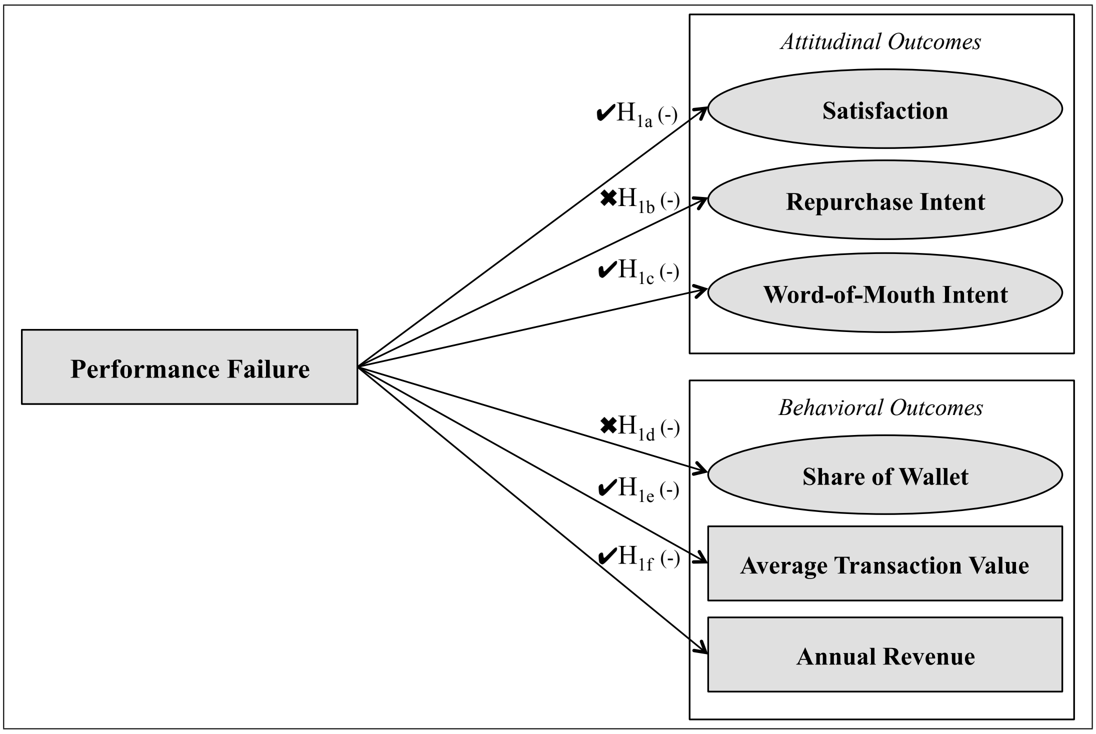


**Figure 4.3:** Summary of Results of Hypotheses Tests

## 4.5 Discussion

In this project, I comprehensively assess the average relationship damage caused by performance failure on both attitudinal and behavioral levels. The results detailed in

Table 4.3 lend substantial support to four of the six hypotheses, thus confirming that performance failures have a strong negative effect on key relationship outcomes. Prior research has not established these effects while considering within- and between-subjects variance. Moreover, research investigating the impact of performance failures on purchase behavior and assessing the financial impact is lacking. The results are discussed in more detail in the following sections that provide theoretical and managerial implications, as well as implications for future research.

### 4.5.1 Implications for Research

*Attitudinal outcomes.* As expected, this research confirms the detrimental effect of performance failure on satisfaction and word-of-mouth intent. With regard to satisfaction, the impact estimated is $-.67$ ($-13.6\%$), in line with prior work. For example, in conducting an scenario-based experiment in a restaurant setting, Hocutt, Bowers, and Donavan (2006) report mean differences across six treatment conditions ranging from .18 to $-1.52$. Averaging these means and comparing them with their no-failure control group yields a value of $-.69$ ($-13.0\%$), which is remarkably close to this study's result.[15] Averaging the values that Michel and Meuter (2008) report yields an effect of approximately $-.39$ ($-8.9\%$), which is slightly lower than the result of this study. However, because of the different operationalization of satisfaction, these figures suffer from limited comparability.[16] Van Doorn and Verhoef (2008) report service satisfaction

15 The authors employ a comparable operationalization of satisfaction, using a similar multi-item construct to that employed in this study, also measured on a seven-point scale (see section 3.3). However, no prefailure assessment of satisfaction levels was conducted; yet it is likely that there is little or no heterogeneity within respondents, because they have no history with the restaurant in the experimental scenario.

16 Michel and Meuter (2008) measure satisfaction using a five-point scale. Moreover, sample sizes of subgroups are unequal. Thus, I calculated the approximation of $-.39$ ($-8.9\%$) by including weighted means for individual subgroups. This approach yields only a rough estimate.

values of a failure and a no-failure control group. Calculating the mean difference of the two groups in the postfailure state yields a value of $-.65$ ($-14.6\%$), which again is much in line with this study's result.[17]

With regard to word of mouth, the analysis reveals a significant, negative effect of $-.49$ ($-8.8\%$). Hocutt, Bowers, and Donavan (2006) report mean differences ranging from $-.31$ to $-1.32$ among subgroups. Averaging these means yields a value of $-1.14$ ($-22.7\%$), which is substantially higher than this study's result. However, it is important to note that this large deviation is probably due to a very different operationalization of word of mouth.[18] The mean differences that Kau and Loh (2006) report range from $-.17$ to $-1.34$ for groups of complainers and noncomplainers who exhibit varying levels of satisfaction. Calculating an average value of these reported statistics yields a mean effect of $-.28$ ($-8.4\%$), which is fairly close to the relative percentage decline revealed in this study.[19] Similarly, averaging the values that Michel and Meuter (2008) report yields an effect of $-.40$ ($-9.1\%$), which also comes remarkably close to the $-.49$ ($-8.8\%$) computed in this study.[20]

Regarding repurchase intent, no significant effect is detected. However, the *p*-value is close to the 5% level, with $p = .083$. A significant impact might potentially be found when using a larger sample. In general, previous research has produced mixed findings in this regard. A few studies report failure-induced effects on repurchase intentions

17 Van Doorn and Verhoef (2008) employ a seven-point, single-item measure of service satisfaction.

18 Hocutt, Bowers, and Donavan (2006) measure negative word-of-mouth intent, whereas this study captures the future intent to recommend the respective provider to friends (see section 3.3). Thus, for computation purposes, the means reported in their study were reverse-coded to produce a comparable value.

19 Kau and Loh (2006) measure word-of-mouth intent on a five-point scale using multiple items. Thus, a comparable value in absolute terms is unexpected.

20 Note that Michel and Meuter (2008) measure recommendation intent using a five-point scale. Moreover, sample sizes of subgroups are unequal. Thus, the approximation of $-.40$ ($-9.1\%$) was calculated using weighted means for individual subgroups and then estimating the average values.

(e.g., Gelbrich and Roschk 2011), and others do not. For example, Kau and Loh (2006) find no significant mean difference between complainers and noncomplainers in loyalty intentions. In addition, Maxham and Netemeyer (2002a) detect no significant increase in repurchase intent in one of their recovery sequences and speculate that consumers' formation of repurchase intentions depends more heavily on past experiences and is less susceptible to recent encounters. This rationale is in line with the findings of Orsingher, Valentini, and de Angelis's (2010) meta-analysis, which reports a nonsignificant effect of transactional complaint satisfaction on return intent. Gelbrich and Roschk (2011, p. 37) argue that repurchase intentions are less affected because "single transactions are not salient for the decision to continue a relationship"; instead, an overall assessment of all past experiences would be more powerful in predicting future purchases. Thus, repurchase intentions are potentially more stable attitudinal outcomes, which are more resilient and not easily affected by performance failures—particularly when positive past experiences are present.

*Behavioral outcomes.* Regarding the set of behavioral outcomes, no significant impact occurs for share of wallet. However, the $p$-value is close to the 5% level, with $p$ = .052. A significant effect might potentially be detected when using a larger sample. Prior research investigating share of wallet as an outcome variable in a failure scenario is limited. To my knowledge, only one study exists (van Doorn and Verhoef 2008); however, the authors use a categorical scale[21] to measure customer share; thus, no comparison

[21] Van Doorn and Verhoef (2008) measure customer share on a six-point scale (i.e., 1: <10%; 2: 10%-20%; 3: 20%-30%; 4: 30%-40%; 5: 40%-60%; and 6: >60%). Thus, a comparison of effect sizes would involve very rough approximations. Moreover, paradoxically, in their study they report a stronger decline in share of wallet for customers in the no-failure control group than for those in the failure group, which is not only counterintuitive but also contrary to this study's results. The authors offer no specific explanation for this phenomenon. Consequently, their reported values could not be related to this study's result.

can be drawn. A potential reason for the nonsignificant effect for share of wallet might be that customers were limited in the number of alternative providers to which they could shift their patronage. Furthermore, high switching costs might have restrained dissatisfied customers from using competitive offerings.[22]

For both average transaction value and annual revenue, a significant, negative impact is found. Customers apparently reduced their purchase spending per transaction by 7.52 € ($-16.0\%$) , suggesting that they refused to repurchase high value products after a performance failure. Potentially, they would consider only the most necessary products from very basic and low-priced categories from the transgressing retailer. No prior research has assessed the impact of performance failure on average transaction value; thus, future research could further investigate the underlying reasons for this finding.

Regarding annual revenues as an outcome variable, the analysis shows that customers who had a performance failure significantly reduced their annual purchase spending. This provides causal evidence for a direct effect of a performance failure on postfailure purchase behavior. Prior work has not always been able to demonstrate this link. Gilly (1987) finds no significant direct relationship between a complaint and the complainant's actual repurchase behaviors. Rather, she suggests that the causal relationship is mediated by the complainant's cognitive processes regarding the complaint response. In contrast, von Wangenheim and Bayón (2007) find evidence for this direct relationship between service failure and postfailure purchase behavior. Theoretically, this effect is naturally mediated by cognitive processes, as shown by other research (e.g., mediation through justice perceptions on attitudinal outcomes, see Orsingher, Valentini, and de Angelis 2010). However, to prove mediation on purchase behavior, methodologi-

[22] See also Chapter 6.

cally a direct effect must be present in the first place (Baron and Kenny 1986).[23] This study's results suggest a decline in annual revenues by 290.10 € (−22%) on average for customers who reported a failure in comparison with customers who did not experience a failure. Figure 4.2 illustrates the development of monthly purchase spending over time comparing the two groups. Thus, customers in the failure group substantially reduced their business by roughly one-fifth of the past transacted purchase volume with their focal provider.[24] Although this finding is neither unexpected nor counterintuitive, this study is among the first to quantify a monetary impact of a performance failure in a popular B2C retail context. Only a few researchers have examined postfailure purchase behavior.[25] Reporting a change in purchase volume of −19.10%, Evanschitzky, Brock, and Blut (2011) reveal a similar average decline for dissatisfied customers in a fast-food delivery context, which suggests a relative detrimental effect similar to this study's finding.[26]

Overall, the results obtained from self-reported and observed data are not fully consistent.[27] While for the observed purchase-related outcomes—average transaction value

23 See also Chapter 5.

24 It is a known phenomenon of noncontractual settings that cohorts of a customer base exhibit a negative trend in purchase behavior over time (e.g., Reinartz and Kumar 2000). This is potentially due to satiation–that is, a decreasing individual-level demand in certain product categories (Voss, Godfrey, and Seiders 2010). Similarly, in the employed database, a slight negative trend in global purchase levels is observed. That is, customers of both failure and control groups reduce their patronage over time. However, the applied DID technique accounts for this by estimating group differences over time and thus reveals that customers who experienced a failure purchase 290.10 € (−22%) less than similar customers without such a dissatisfying encounter.

25 Gilly and Gelb (1982), Gilly (1987), and von Wangenheim and Bayón (2007) do not report the average monetary impact or distort financial figures for confidentiality reasons.

26 Evanschitzky, Brock, and Blut (2011) further differentiate between customers high or low in affective commitment and complaint satisfaction and report values in the range from .16% to −35.82%. However, they do only assess the relative percentage change. They do not investigate effects in absolute figures and consequently no financial impact can be obtained from their study.

27 Note that not all behavioral outcomes are based on transaction data. Share of wallet was measured using self-reported survey data (see 3.3).

and annual revenue—a significant result was obtained, the hypotheses regarding purchase intent and share of wallet remain unconfirmed. This finding illustrates that research relying merely on self-reported data may lead to wrong conclusions. Only by using purchase intent as a proxy for future purchase behavior, for example, would performance failures have had no detrimental effect on purchase activity, though customers would have indeed reduced their annual purchase volume by 22%. Theoretically, these deviations may be present because of intervening contingency factors that customers often fail to account for when predicting their own future behavior (Seiders et al. 2005).[28] However, for both, purchase intent and share of wallet, the $p$-values were close to the 5% threshold, suggesting that these effects presumably become significant when estimated with a larger sample size.

Regarding the estimated effect sizes, in summary, the obtained values are in line with prior research. Notably, some slight deviations are present, but these differences can likely be ascribed to dissimilar research designs, contexts, and variable operationalizations. In general, prior work conducts simple mean comparisons and does not account for customer heterogeneity or the prefailure status of the customer relationship in key outcome variables. Although some previous findings in the literature may be biased, it is remarkable how well the results of the different approaches converge in some cases. A potential reason may be that in experimental research, bias is generally small as a result of controlled study designs, hypothetical scenarios, and absence of prior customer history with a certain provider, leading to less pronounced heterogeneity among participants. In a field study, however, these influences should be accounted for with correction techniques to ensure unbiased parameter estimates.

[28] See also Chapter 6.

In summary, the applied methodology is especially powerful because it simultaneously captures before-and-after differences in outcomes and differences of homogeneous failure and nonfailure customer groups. Thereby, it establishes a clear causality of the effect of performance failure and and answers the counterfactual question of how behaviors and attitudes would have developed had customers not experienced a failure. Such an approach is new to the research domain of failure, recovery, and complaint management and may be particularly suitable for further research in the field. In particular, the use of a control group in field study research designs can help obtain true parameter estimates because customers who did not experience a failure but may also exhibit decreasing loyalty behaviors for other reasons, such as an increased attractiveness of competitive offerings or general economic trends.

### 4.5.2 Implications for Practice

This research conceptually and empirically contributes to a comprehensive understanding of the negative consequences that can result from performance failures. The findings illustrate that failures can harm multiple aspects of relationship outcomes. Service managers who focus only on a single dimension, for example, first-call resolution (FCR) quotas[29] or satisfaction survey results, may fail to capture the full spectrum of detrimental effects caused by performance failures. As such, managers may systematically underestimate the impact on the bottom line and underrate the importance of high-quality complaint handling and recovery capabilities. Both levels of outcomes —attitudinal and behavioral—must be considered when designing recovery strategies

[29] FCR is a key performance indicator employed by operations management, particularly in the call center. It measures whether a customer problem could have been fixed the first time he or she called into the call center.

to mitigate potential direct and indirect negative consequences of failures. Behavioral outcomes have a more direct financial impact resulting from altered purchase behavior. Attitudinal outcomes such as satisfaction and word-of-mouth intent are a precursor to the more indirect consequences, such as negative publicity. In summary, both outcome levels ultimately can lead to decreased customer equity; thus, managers should strive to avoid or mitigate all potential negative consequences to the best of their ability using available resources. By measuring and monitoring these key outcomes (e.g., through a survey after a customer complaint), managers can determine the success of their current recovery policies, implement continuous improvement initiatives, and set targets based on these metrics to further ameliorate recovery processes. In particular, assessing the impact on repurchase behavior would provide great insight for the management of failure and recovery because "recovery-managers often underestimate the profits lost when a customer departs unhappy, and therefore they undermanage ways of avoiding such losses" (Hart, Heskett, and Sasser 1990, p. 150). By calculating the potential losses in revenues after failure, companies could determine how much money to spend on recovery efforts. For each customer, the invested resources should typically not be higher than the estimated return that can be obtained from his or her future cash flows (i.e., the CLV). However, companies should consider not only the revenues generated from future purchase activities but also potential attitudinal benefits. The occurrence of a performance failure represents a 'moment of truth' in a customer relationship (Carlzon 1989; Tax, Brown, and Chandrashekaran 1998). That is, it offers the company a chance to convince the customer of its superior service quality and consequently may yield increased positive word of mouth. In general, fair policies and procedures, as well as polite, obliging employee interactions, have the potential to influence more enduring customer perceptions of overall firm satisfaction (Maxham and Netemeyer 2002b), which may lead to positive publicity, company image, and recommendations.

To accomplish this, high-quality complaint handling is a prerequisite. However, excellence in recovery capabilities is undoubtedly costly to sustain and requires substantial resources and investments. Quantifying the financial impact of performance failures can help managers justify investments in customer service functions. For the investigated retail company, my analysis reveals that customers who experienced a performance failure reduced their annual purchase spending by 290.10 € (−22%) in comparison with similar customers who did not experience a failure. I do not suggest that this effect is generalizable to other contexts and providers, but it clearly illustrates the dramatic impact of performance failures. With a customer base of 1.5 million and approximately 7% experiencing a serious failure, the retailer's average losses amount to 30.5 € million in revenues—in just the first year after the transgression. This is substantial, especially when considering that this assessment does not incorporate indirect negative consequences that may result from negative word of mouth or other activities related to negative publicity, such as online public complaining (e.g., Tripp and Grégoire 2011).[30] For marketing managers, this finding strongly underscores the importance of high-quality complaint handling and supports claims for budgets in the boardroom.

### 4.5.3 Limitations and Further Research

As with all research, this study is constrained by limitations that, at the same time, offer implications for further research. First, note that the estimated effects are average values. That is, they represent consequences of performance failures that may have been

[30] Potentially, the negative consequences are even more substantial when considering the unknown percentage of noncomplainers; that is, customers who experienced a failure but decided not to complain and defect silently (e.g., Goodman 2006).

successfully resolved or not. A further distinction or split of groups was not feasible because of the small sample size available and some methodological reasons. Thus, further research should try to obtain larger sample sizes and estimate causal effects for treatment subgroups–for example, a segment that received an excellent recovery versus segments receiving average or poor recoveries. Moreover, varying levels of further determinants critical to postfailure outcomes could be used for additional subanalyses, such as failure severity, failure type, time required for failure resolution, attributions of blame, and recovery satisfaction. Second, most prior studies comparing customers of a failure and a control group do not consider relationship dynamics (i.e., the pre- and post-failure level of the outcome variable) and customer heterogeneity in key characteristics. Moreover, prior work employes different operationalizations of outcomes (e.g., 5-, 7-, 10-point scales; measures of actual behavior instead of behavioral intent). Thus, a direct comparison this study's findings with the results of other work leads to conclusions of limited validity. Future work should use similar, comprehensive research designs, operationalizations, and methodologies to be able to synthesize the findings from different studies in various contexts. Third, this study relied on transactional data obtained from a loyalty program. Thus, the estimated effects might be somewhat attenuated, as loyalty program members tend to have decreased negative perceptions when experiencing poor service encounters (Bolton, Kannan, and Bramlett 2000). Thus, further research might use transactional data from other sources than loyalty programs to avoid such potential loyalty effects. Fourth, this study was conducted in a noncontractual setting in the retailing sector. Additional studies could replicate and extend the findings and investigate potential differences in other industries and contractual settings.

### 4.5.4 Conclusion

This study contributes to prior research by (1) comprehensively assessing the average relationship damage of performance failures on attitudinal and behavioral outcomes, (2) clearly establishing causality, and (3) estimating the financial impact in terms of postfailure purchase behavior. The applied technique enabled an analysis and comparison of transaction behavior and attitudes across customer groups before and after a performance failure for a substantial period and on a detailed level. The study shows that performance failures can have a detrimental effect on outcomes on multiple levels. In turn, these may lead to direct and indirect negative consequences. The more direct monetary consequences result from altered behaviors, such as reduced purchase activity and a potential shift of patronage to competitors. A greater difficulty is in assessing the more indirect negative consequences resulting from a decline in attitudinal outcomes. Studies suggest that "dissatisfied customers tell 10 to 20 people. The exploding Internet means this kind of damaging communication will soar" (Brown 2000, p. 9). Thus, in order to stay off negative word-of-mouth, particular care should be administered to managing complaints holistically with regard to both attitudinal and behavioral outcomes.

# 5 The Effects of Perceived Justice on Postfailure Purchase Behavior

## 5.1 Overall Background

Two research streams that investigate recoveries from failure have emerged over the years. Beyond some early studies examining the role of *organizational responses* to complaints (e.g., Gilly and Gelb 1982; Lewis 1983), Goodwin and Ross (1989) were among the first to build on fairness theory and employed *dimensions of perceived justice* (interactional, procedural, and distributive justice) to explain the formation of postrecovery satisfaction. A large number of studies investigating the complex interrelationships of organizational responses, justice perceptions, and moderating variables and how these affect postfailure outcomes followed.[31] In their recent meta-analysis, Gelbrich and Roschk (2011) incorporate both organizational response and perceived justice dimensions (JDs) and consolidate the empirical findings of 87 studies. They find that the three justice perceptions fully mediate the relationship between organizational response and cumulative satisfaction. That is, perceived justice entirely explains the link between organizational response and postfailure satisfaction and thus can be employed

31 For a more detailed review of research on organizational responses and perceived justice, see section 2.3.

to approximate organizational responses. Justice perceptions "are the subjective interpretation of service recovery efforts that are responsible, more than the recovery efforts themselves, for the subsequent satisfaction judgment" (Gelbrich and Roschk 2011, p. 37). This finding strongly underscores the crucial role of perceived justice in the recovery process.

Despite the substantial body of literature investigating the role of the three JDs, several questions remain unanswered. First, considerable evidence shows that postfailure satisfaction leads to postfailure loyalty. However, almost all prior studies have operationalized postfailure loyalty using self-reported loyalty intentions. Thus, it is not clear whether postfailure satisfaction leads to actual loyal behaviors, such as repatronage. The literature distinguishes between *transaction-specific satisfaction* and *cumulative satisfaction*.[32] Regarding transaction-specific satisfaction, two studies have established the link to postfailure purchase behavior (i.e., Evanschitzky, Brock, and Blut 2011; Gilly and Gelb 1982). With regard to cumulative satisfaction, in general, the evidence is mixed and several studies report it is not trivial to establish a significant effect of satisfaction on purchase behavior. For example, Seiders et al. (2005) and Voss, Godfrey, and Seiders (2010) find no significant main effect on purchase spending but report a significant effect when testing for variables interacting with satisfaction. Similarly, Mittal and Kamakura (2001) find that satisfaction significantly affects purchase behavior when tested in conjunction with customer characteristics as moderating variables. Although postfailure cumulative satisfaction has been a popular outcome variable in the research domain of performance failures and customer complaints, it has not been empirical assessed whether it actually translates into repurchase behavior. This is surprising, be-

[32] For a definition of transaction-specific vs. cumulative satisfaction, see section 2.3.1.

cause it is regarded as “mainly responsible for repatronage” (Gelbrich and Roschk 2011, p. 37).[33]

Second, postfailure satisfaction mediates the effects of perceived justice on postfailure *attitudinal outcomes*, such as loyalty intentions (e.g., Gelbrich and Roschk 2011; Maxham and Netemeyer 2002b). However, it is not known whether satisfaction also mediates the effects of perceived justice on *behavioral outcomes*. Although numerous studies have investigated the impact of JDs on behavioral intentions, little or no research has examined their impact on *actual behavior* (see Table 5.1). Only one study has investigated the effect of perceived justice on behavioral consequences of complaints (Chebat and Slusarczyk 2005); however, the study was limited in terms of the dependent variable (exit vs. loyalty). Research examining the effect of JDs on purchase behavior over time is scarce. Rust and Chung (2006) contend that no study employs a database/panel approach to complaint management, and Parasuraman (2006, p. 590) calls for analytical modeling efforts “that could inform the design of optimal recovery strategies.” In the same vein, Davidow (2003b, p. 246) argues that “only by quantifying the effects of each response dimension on postcomplaint customer behavior will we be able to plan efficient and effective complaint management.” Consequently, research that enables competing complaint resolution effort dimensions to be traded off on the basis of monetary return. By linking complaint resolution effort dimensions with buying behavior, a financial impact can be estimated is necessary. Such an approach can inform managers how investments in different JDs translate into future revenues and thus provide guidance for resource allocation and the development of efficient complaint management strategies.

33 Throughout the remainder of the study, when discussing satisfaction, I refer to the cumulative conceptualization unless otherwise indicated.

Third, prior work does not account for prefailure levels of satisfaction. Previous research has shown that prior satisfaction levels directly affect subsequent outcomes (LaBarbera and Mazursky 1983; Smith and Bolton 1998). Moreover, accounting for prefailure satisfaction is particularly important when studying the effects of perceived justice with complaint handling, because "prefailure conditions can influence perceptions of recovery" (DeWitt and Brady 2003, p. 204). In addition, for customers who experienced a critical incident, van Doorn and Verhoef (2008) find in their longitudinal study that high overall service satisfaction ratings persist over time, which implies that previously highly satisfied customers are more forgiving than less satisfied customers. Thus, when studying the effect of perceived justice on postfailure satisfaction, these carryover effects should be accounted for because otherwise, obtained results may suffer from omitted variable bias.

Against this background, the overarching goal of this project is to determine whether and how perceived JDs have an impact on actual postfailure purchase behavior. Overall, this research aims to make three key contributions: (1) to investigate the effect of postfailure satisfaction on purchase behavior, (2) to analyze whether satisfaction mediates the effect of justice perceptions on purchase behavior, and (3) to account for prefailure levels of satisfaction, which are examined for potential carryover effects. Table 5.1 summarizes the findings of prior research investigating the JDs $\rightarrow$ satisfaction $\rightarrow$ loyalty outcomes link and depicts how this study contributes to existing knowledge.

**Table 5.1:** Prior Studies Investigating the JDs → SAT (C) → Loyalty Outcomes Link - Part I

| Study | Study Context (Design) | Account for Prefailure Relationship State | Loyalty Outcome Variable(s) | Results: Influence of JDs | Results: Mediating Role of SAT (C) |
|---|---|---|---|---|---|
| Chebat and Slusarczyk (2005) | Banking (Field study) | No | Exit vs. loyalty[OB] | All JDs affect emotions. Only IJ directly affects exit-loyalty behavior. | NA |
| Davidow (2003a) | Misc:[a] Auto repair hospitality, retail (Field study) | No | Repurchase intentions[IS] | JDs → SAT (C): (1) DJ (+) (2) IJ (+) (3) PJ (n.s.) | No mediation test but indirect effects present for all JDs. |
| Gelbrich and Roschk (2011) | Meta-analysis | NA | Repurchase intentions[IS] WOM intentions[IS] | JDs → SAT (C): (1) IJ (+) (2) DJ (+) (3) PJ (+) | Full mediation for all JDs and outcomes, except from the PJ → SAT → WOM intent link. |
| Homburg and F"urst (2005) | Across industries: B2B, B2C, service & manufacturing (Field study, cross-sectional dyads) | No | Loyalty[IS] Repurchase & relationship continuation | JDs → SAT (C):[b] (1) DJ (+) (2) PJ (+) (3) IJ (+) | NA |
| Martinez-Tur et al. (2006) | (1) Hotel (2) Restaurant (Field studies) | No | NA | JDs → SAT (C): (1) DJ (+) (2) IJ (+) (3) PJ (+) | NA |
| Maxham and Netemeyer (2002b) | Study 1: Banking (Field study, longitudinal) | No | Repurchase intentions[IS] WOM intentions[IS] | JDs → SAT (C): (1) PJ (+) (2) IJ (+) (3) DJ (+) | Full mediation for DJ. Partial mediation for PJ. Full mediation for the IJ → SAT → WOM intent link and partial mediation for the IJ → SAT → repurchase intent link. |
| | Study 2: Home-construction (Field study, longitudinal) | No | | JDs → SAT (C): (1) PJ (+) (2) IJ (+) (3) DJ (+) | Full mediation for DJ. Partial mediation for IJ and PJ. |

**Table 5.2:** Prior Studies Investigating the JDs → SAT (C) → Loyalty Outcomes Link - Part II

| Study | Study Context (Design) | Account for prefailure relationship state | Loyalty Outcome Variable(s) | Results | |
|---|---|---|---|---|---|
| | | | | Influence of JDs | Mediating Role of SAT (C) |
| Maxham and Netemeyer (2003) | Online retailer (Field study) | No | Repurchase intentions[IS]<br>WOM intentions[IS] | JDs → SAT (C):<br>(1) DJ (+)<br>(2) PJ (+)<br>(3) IJ (+) | NA |
| Severt (2002) | Misc: Airline, automotive, hotel, restaurant, retail, etc. (Field study) | Prior experience[IS] | NA | JDs → SAT (C):<br>(1) PJ (+)<br>(2) IJ (+)<br>(3) DJ (n.s.) | NA |
| Smith and Bolton (2002) | Hotels and restaurants (Field study) | No | NA | JDs → SAT (C):<br>(1) DJ (+)<br>(2) IJ (+)<br>(3) PJ (n.s.) | NA |
| Varela-Neira et al. (2008) | Banking (Field study) | No | NA | JDs → SAT (C):<br>(1) IJ (+)<br>(2) PJ (+)<br>(3) DJ (n.s.) | NA |
| **This study** | Retailing (Field study, longitudinal) | Prefailure SAT (C)[IS]<br>Prefailure purchase behavior[OB] | Purchase behavior[OB] | JDs → SAT (C):<br>(1) IJ (+)<br>(2) PJ (n.s.)<br>(3) DJ (n.s.) | Full mediation for IJ.<br>No mediation for PJ & DJ. |

[a]Student sample, [b]based on squared correlations.

Notes: Studies displayed examine the effects of JDs on SAT (C) (cumulative satisfaction) or on behavioral loyalty. When findings are mixed or contingent on other variables, the most conclusive results are reported. (1)-(3) indicates the rank order of the relative strength of JD effects. JD = justice dimension, IJ = interactional justice, PJ = procedural justice, DJ = distributive justice, WOM = word of mouth, NA = not applicable, n.s. = not significant, IS = intentions survey measure, OB = observed behavior measure.

The remainder of the project proceeds as follows: In the next section, I describe the conceptual model, elaborate on the theoretical basis, and derive hypotheses. Then, I outline the research methodology and test the propositions using longitudinal transaction and survey data from a major European retailer.[34] Finally, I present the results and discuss the implications of the findings.

## 5.2 Theoretical Basis and Hypotheses

The unit of analysis is a retailer and its complaining customers. In line with the previous discussion, the framework (see Figure 5.1) includes constructs based on prior work that investigates the role of perceived justice in a failure/recovery context. I extend current knowledge by relating JDs and postfailure satisfaction to actual purchase behavior. Davidow (2003b, p. 247) calls for such research and asserts that "perceived justice is the customer's feeling or reaction to the organizational complaint response, and should have a major impact on satisfaction and postcomplaint customer behavior." Moreover, he requests that studies analyze such a causal chain "and the mediating effect of satisfaction and perceived justice on that relationship" (p. 246). Thus, in the model, satisfaction serves as an outcome and mediating variable, and purchase spending serves as an ultimate outcome and loyalty measure. In addition, the framework comprises the three perceived JDs, which are antecedents to the outcomes of satisfaction and purchase spending. The dynamic model also includes prefailure levels of satisfaction and purchase spending because performance failures affect customer relationships differently depending on their initial state (e.g., Smith and Bolton 1998).

[34] For a description of the research design, empirical setting, the data, and measures employed for this study, please refer to Chapter 3.

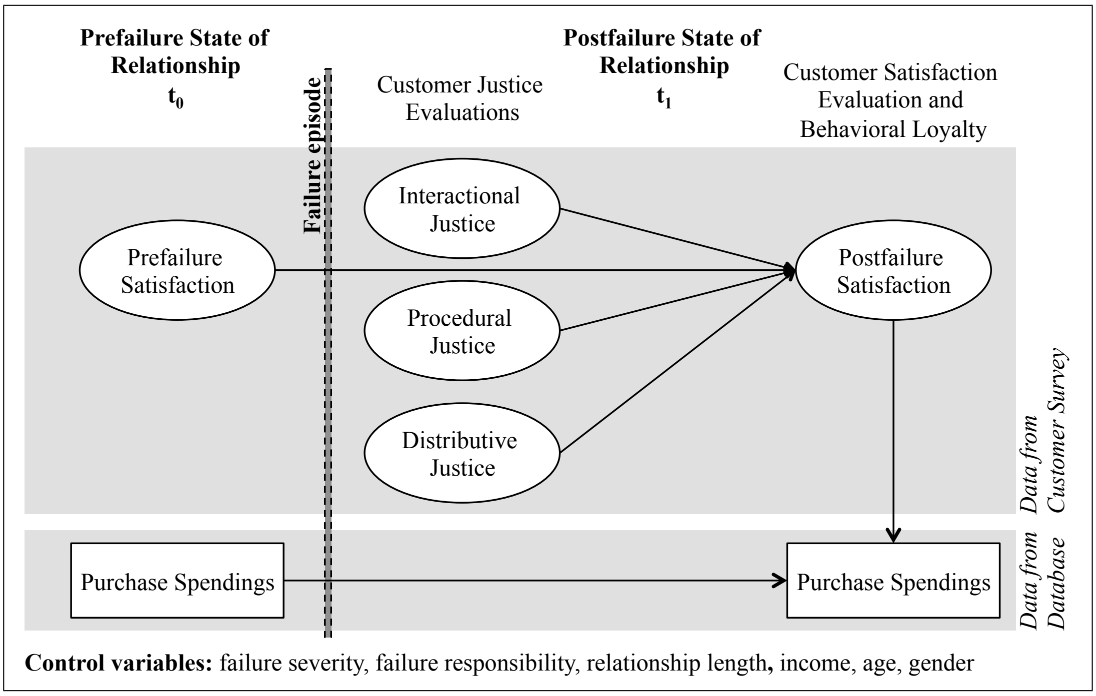


**Figure 5.1:** Hypothesized Model

Justice (or fairness) theory has its origins in social psychology and is derived from equity theory (Adams 1965), which pertains to a person's perception of the fairness of a specific event or decision. According to this, people perceive relationships and interactions as equitable (or fair) when the ratio of their outputs (benefits) to inputs (efforts) is balanced with the output/input ratio of the other party. In general, the theory explains individual reactions to a conflict situation in an exchange context. For the specific performance failure context, it has proved particularly valuable for explaining the customer's perception of fairness (Clemmer and Schneider 1996), reactions to failure/recovery, and the formation of postrecovery outcomes (e.g., Gelbrich and Roschk 2011; Goodwin and Ross 1989; Orsingher, Valentini, and de Angelis 2010; Smith, Bolton, and Wagner 1999; Tax, Brown, and Chandrashekaran 1998). Across different disciplines and contexts, research has identified and widely adopted three dimensions of perceived justice:

interactional, procedural, and distributive.

*Interactional justice* (IJ) "refers to the manner in which people are treated during the complaint resolution process" (Blodgett, Hill, and Tax 1997, p. 189) and pertains to customer interactions with the retailer's staff. That is, IJ refers to whether the provider's employees are pleasant and considerate when dealing with customers. Prior research has emphasized the importance of treating the customer politely and in a friendly manner (Tax, Brown, and Chandrashekaran 1998), behaving in a courteous way (Blodgett, Hill, and Tax 1997; Hocutt and Chakraborty 1997; Liao 2007), and making a considerable effort to solve the customer's problem (Homburg and Fürst 2005; Smith, Bolton, and Wagner 1999). Prior work has also found a positive relationship between IJ and satisfaction (e.g., Maxham and Netemeyer 2002b, 2003; Varela-Neira, Vazquez-Casielles, and Iglesias-Arguelles 2008).

*Procedural justice* (PJ) reflects the perceived fairness of the complaint handling processes (e.g., Bitner, Booms, and Tetreault 1990). The complaint handling process is meaningful because it aims to resolve conflicts in a way that encourages the continuation of the relationship between the firm and a complainant. This process comprises elements such as an easy ability to engage in complaining (e.g., Tax, Brown, and Chandrashekaran 1998) and completion of the process in a timely manner (e.g., Smith, Bolton, and Wagner 1999). Because process is an integral part of the product or service offering, companies can enhance satisfaction by engaging in activities that ameliorate perceptions of PJ (Seiders and Berry 1998). Considerable research suggests a significant influence of procedural complaint issues on satisfaction (e.g., Martínez-Tur et al. 2006; Severt 2002).

*Distributive justice* (DJ) "describes the fairness of the complaint outcome as the customer perceives it" (Homburg and Fürst 2005, p. 98). It is the result or outcome of complaint handling (Kelley, Hoffman, and Davis 1993); thus, the central component of DJ is compensation, which includes refunds, replacements, repairs, discounts on future patronage, or some combination thereof (Blodgett, Hill, and Tax 1997). According to the DJ concept, these outcomes must be fair for a positive customer perception; that is, customers must be converted back to their starting point, or otherwise, they will remain dissatisfied with the response. Previous research has found a positive relationship between DJ and satisfaction (e.g., Davidow 2003a; Maxham and Netemeyer 2003).

Note that the literature has already established the hypotheses of JDs on cumulative satisfaction. Nevertheless, I present them herein to additionally control for prefailure satisfaction, which previous research has neglected. Severt (2002) shows that past experience has an impact on justice perceptions, and thus it is important to account for it. Cumulative satisfaction is additive in nature and covers all experiences in relationship history before the failure (Anderson and Sullivan 1993). When controlling for prefailure satisfaction, a part of the unexplained variance in postfailure satisfaction can likely be ascribed to the occurrence of a performance failure. Typically, a decrease of satisfaction levels from the pre- to the postfailure state will occur. This deviation likely represents a share of variance in postfailure satisfaction that cannot be explained by prefailure satisfaction; instead, JDs should be able to contribute significantly to explaining that fragment and increase the R-square accordingly. Thus, in summary, the following hypothesis is put forth:

> $H_1$ : *(a) IJ, (b) PJ, and (c) DJ have a positive effect on postfailure satisfaction when controlling for prefailure satisfaction.*

*Postfailure satisfaction and purchase behavior.* As noted previously, cumulative satisfaction is an overall evaluation of firm performance that accounts for all experiences with a firm (e.g., Johnson, Anderson, and Fornell 1995). That is, not only does cumulative satisfaction account for the evaluation of a particular recovery effort, but it also holistically captures all other aspects and prior experiences of the customer relationship history. Gelbrich and Roschk (2011) find that cumulative satisfaction exerts predominating effects on customer loyalty intentions with an impact of $\beta = .56$, whereas transaction-specific satisfaction is less decisive ($\beta = .30$). This is also reflected, for example, in the findings of Maxham and Netemeyer (2003), who report that cumulative satisfaction has a significant impact on repurchase intent whereas transaction-specific satisfaction has no direct effect. Satisfaction with complaint handling has been linked to purchase behavior (e.g., Evanschitzky, Brock, and Blut 2011), but no prior research has conducted such an analysis with cumulative postfailure satisfaction. Although in general the link of satisfaction to loyalty behaviors is well established (e.g., relationship duration [Bolton 1998], customer retention [Verhoef 2003], share of wallet [van Doorn and Verhoef 2008]), only a few studies relate cumulative satisfaction to purchase behavior (Mittal and Kamakura 2001; Seiders et al. 2005; Voss, Godfrey, and Seiders 2010). Given the consistent evidence that cumulative satisfaction is the "primary antecedent of customer loyalty" in failure scenarios (Gelbrich and Roschk 2011, p. 24), I test the following hypothesis:

> $H_2$ : *Postfailure satisfaction has a positive effect on postfailure purchase spending.*

*The mediating role of satisfaction.* In general, satisfaction is treated in the literature as a central key mediating variable of loyalty constructs (Oliver 1996). In a failure/recovery context, prior research has identified satisfaction as a mediator of justice perceptions. For example, Maxham and Netemeyer (2002b) find that satisfaction mediates the effects of justice on repurchase intent. Similarly, with their meta-analytic approach, Gelbrich and Roschk (2011, p. 37) confirm that satisfaction fully mediates the relationships between the JDs and behavioral intentions and assert that "justice perceptions directly affect cumulative satisfaction, which in turn is mainly responsible for repatronage." A similar effect would be expected regarding purchase behavior; thus, I hypothesize the following:

$H_3$ : *Postfailure satisfaction mediates the effect of (a) IJ, (b) PJ, and (c) DJ on postfailure purchase spending.*

## 5.3 Methodology

According to the conceptual model, I formulate the econometric model as follows:

$$\mathrm{S_{i,t}} = \alpha_0 + \alpha_1 \times \mathrm{SAT_{i,t}} + \alpha_2 \times \mathrm{S_{i,t-1}} + \varepsilon_S, \text{ and} \tag{5.1}$$

$$\mathrm{SAT_{i,t}} = \beta_0 + \beta_1 \times \mathrm{IJ_{i,t}} + \beta_2 \times \mathrm{PJ_{i,t}} + \beta_3 \times \mathrm{DJ_{i,t}} + \beta_4 \times \mathrm{SAT_{i,t-1}} + \varepsilon_{SAT}, \tag{5.2}$$

where

$S$ = purchase spending,

$S_{t-1}$ = lagged purchase spending,

$SAT$ = satisfaction,

$SAT_{t-1}$ = lagged satisfaction,

$IJ$ = interactional justice,

$PJ$ = procedural justice, and

$DJ$ = distributive justice.

The model is estimated using seemingly unrelated regression (SUR). This approach is considered adequate when jointly estimating parameters in different equations (Wooldridge 2002a). Moreover, the SUR estimator accounts for contemporaneous correlations between the error terms (Kennedy 2003). Lagged variables are included in each equation because the formation of postfailure outcomes is dependent on prefailure satisfaction or spending levels. In addition, I controlled for a set of key variables that potentially affect postfailure satisfaction and purchase behavior. For eq. 5.4, these variables include the severity of the failure and failure responsibility; both of which can affect customer responses to failure recovery (e.g., Smith, Bolton, and Wagner 1999). Moreover, I controlled for the customer characteristics age, gender, and relationship length (Mittal and Kamakura 2001). Research has also identified the variables household income (Seiders et al. 2005) and relationship length (Reinartz and Kumar 2000) as influencing factors of purchase behavior, and thus these were included as controls in eq. 5.3.

Multicollinearity might affect the estimation results. The majority of the correlation coefficients between predictor variables are less than .5 (see correlation matrix in Table 5.3). Only the bivariate correlations between IJ–PJ (.62) and PJ–DJ (.67) appear high. Prior studies have frequently reported high correlations between JDs (e.g., Liao 2007). Moreover, in their meta-analysis, Gelbrich and Roschk (2011) find poor discriminant validity among the three JDs in general and contend that this is potentially due to a customer's inability to clearly distinguish between their individual characteristics. However, when testing for discriminant validity (Fornell and Larcker 1981), I find Fornell-Larcker ratios of between .29 and .87, which suggests that the Fornell-Larcker criterion is satisfied and constructs have discriminant validity (see Table C.2 in Appendix C). In addition, variance inflation factors (VIF) were examined, and all scores were between 1.5 and 2.39, which is substantially below the 10 guideline (Hair et al. 1998). Thus, I can conclude that multicollinearity does not affect the estimation results.

For a detailed description of the research design, data collection, sample description, measurements, validity, and reliability analyses, see Chapter 3.

## 5.4 Results

Overall, as the results reveal, the proposed model receives partial support. That is, 32.24% of the variation in postfailure purchase spending is explained by satisfaction and lagged sales (eq. 5.3).[35] Of this percentage, 28.22% can be ascribed to prefailure purchase spending. Thus, the inclusion of satisfaction in the model adds 4.02% of explained variance. The R-square of postfailure satisfaction is .67 (eq. 5.4).[36] Thereof, prefailure satisfaction captures 55.58% of the explained variance. Adding the three JDs

[35] Note on eq. 5.3: N = 145 after exclusion of observations with missing values.
[36] Note on eq. 5.4: N = 108 after exclusion of observations with missing values.

**Table 5.3:** Descriptive Statistics and Correlations of the Study Variables (Project II)

| | **Variable** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | Purchase spending | 1 | | | | | | | | | | |
| 2. | Purchase spending$_{t-1}$ | .65*** | 1 | | | | | | | | | |
| 3. | Satisfaction | .17** | -.04 | 1 | | | | | | | | |
| 4. | Satisfaction$_{t-1}$ | .17** | -.01 | .75*** | 1 | | | | | | | |
| 5. | Interactional Justice | .11 | -.04 | .75*** | .59*** | 1 | | | | | | |
| 6. | Procedural Justice | .08 | -.06 | .58*** | .45*** | .62*** | 1 | | | | | |
| 7. | Distributive Justice | .14 | -.09 | .44*** | .33*** | .45*** | .67*** | 1 | | | | |
| 8. | Failure severity | .10 | .02 | .05 | .13 | -.05 | .04 | -.03 | 1 | | | |
| 9. | Failure responsibility | .02 | -.02 | -.13 | -.04 | -.22** | -.18* | -.21** | .28*** | 1 | | |
| 10. | Relationship length | -.03 | -.08 | .19** | .14* | .13 | .12 | -.11 | .06 | .05 | 1 | |
| 11. | Age | -.08 | -.10 | .09 | .01 | .20** | .20** | .03 | -.10 | -.09 | .22*** | 1 |
| | M | 821.25 | 1316.52 | 4.67 | 4.92 | 5.15 | 4.38 | 3.26 | 4.89 | 4.97 | 11.85 | 48.32 |
| | SD | 748.11 | 1253.16 | 1.64 | 1.51 | 1.49 | 1.98 | 2.13 | 1.79 | 2.14 | 6.17 | 12.35 |

$^{***}p < .01$, $^{**}p < .05$, $^{*}p < .10$.

to the model increases the explained variance by 11.91%. Because all control variables remained nonsignificant, they were removed from the model. The results can be summarized as follows:

$$\mathrm{S_{i,t}} = 10.81 + 126.86^{*} \times \mathrm{SAT_{i,t}} + .91^{***} \times \mathrm{S_{i,t-1}} + \varepsilon_S, \text{ and} \tag{5.3}$$

$$\mathrm{SAT_{i,t}} = 2.42^{***} + .55^{***} \times \mathrm{IJ_{i,t}} + .17 \times \mathrm{PJ_{i,t}} + .06 \times \mathrm{DJ_{i,t}} + .49^{***} \times \mathrm{SAT_{i,t-1}} + \varepsilon_{SAT}. \tag{5.4}$$

IJ exerts the strongest impact on satisfaction and is significant ($\beta$ = .55; t = 5.24, $p$ < .001); thus, $\mathrm{H_{1a}}$ is confirmed. No significant effects emerged for PJ ($\beta$ = .17; t = 1.34, $p$ = .18) and DJ ($\beta$ = .06; t = .53, $p$ = .60); consequently, $\mathrm{H_{1b}}$ and $\mathrm{H_{1c}}$ remain unconfirmed. However, the results offer support for $\mathrm{H_2}$, as the satisfaction–purchase spending link was significant ($\alpha$ = 126.86; t = 2.53, $p$ < .05). I also controlled for the lagged effects of satisfaction ($\beta$ = .49; t = 6.66, $p$ < .001) and purchase spending ($\alpha$ = .91; t = 6.97, $p$ < .001), which, as expected, turned out to be good predictors of the respective dependent variables.

To further test the model's robustness, several alternative regression models were run. For example, ordinary least squares (OLS) regression also testing for direct effects of all JDs and lagged sales on repurchase spending. A simultaneous estimation approach (three-stage least squares) was also employed and delivered almost exactly the same results as the proposed SUR model with respect to significance and effect sizes. This finding adds to the validity and robustness of the proposed model. Because none of the alternative models show better fit indices, the proposed model seems the best representation of the data.

To examine the mediating role of satisfaction on the justice perception–purchase spending link ($H_{3a-c}$), I estimated the models following the procedures that Baron and Kenny (1986) recommend. Accordingly, four conditions must be met for mediation to be present. The first condition is satisfied if the independent variables (JDs) affect the mediator (satisfaction). As the analysis revealed nonsignificant effects of PJ and DJ on satisfaction, the first condition is not satisfied for both independent variables. Thus, mediation is not present, and $H_{3b}$ and $H_{3c}$ remain unsupported. The second condition is satisfied if the mediator affects the dependent variables (purchase spending). Regarding IJ, both the first and the second condition are met because $H_{1a}$ and $H_2$ were confirmed. The third condition is satisfied if the independent variable (IJ) affects the dependent variables (purchase spending) directly. Thus, I estimated a model with a direct path from IJ to purchase spending. A significant direct effect emerged ($\alpha$ = 135.24; t = 2.21, $p < .05$ ), thus satisfying the third condition. The fourth mediating condition is met if the direct path from the independent variable to the dependent variable becomes nonsignificant (i.e., full mediation) or reduced (partial mediation) when the mediating

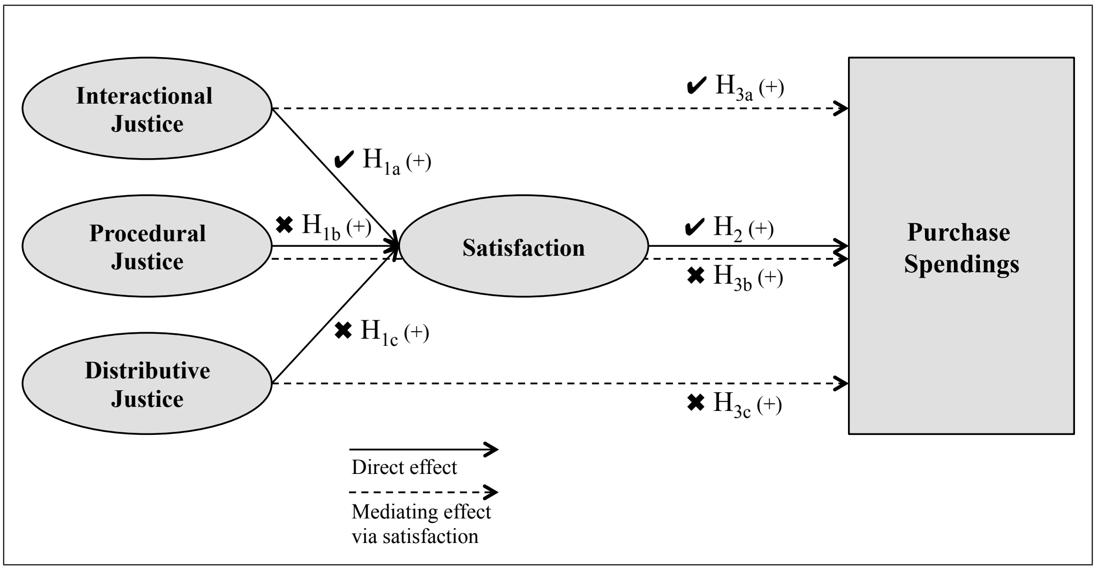


**Figure 5.2:** Summary of Results of Hypotheses Tests

variable is included in the model. After inclusion of satisfaction, the direct effect of IJ becomes nonsignificant (t= -.27, p=.78), suggesting that full mediation is present. Note that the direct effect of IJ ($\alpha$ = 135.24; t = 2.21, $p < .05$ ) is slightly larger than the total mediating effect of satisfaction ($\alpha$ = 126.86; t = 2.53, $p < .05$), suggesting that IJ itself has a stronger influence on the dependent variable than the mediator satisfaction. In summary, as one of three JDs, IJ is fully mediated by the satisfaction construct, in support of $H_{3a}$. Figure 5.2 summarizes the results of the hypotheses tests.

## 5.5 Discussion

In this project, I analyze whether and how perceived JDs affect satisfaction and actual postfailure purchase behavior. In summary, the results in eq. 5.3 and eq. 5.4 confirm three of seven hypotheses, thus lending partial support that perceived justice translates into postfailure purchase behavior. Prior research has not established this relationship. The results are discussed in more detail in the following sections that provide theoretical and managerial implications, as well as implications for future research.

### 5.5.1 Implications for Research

*Influences on satisfaction.* As the results suggest, IJ exerts the strongest influence on satisfaction. This is in line with the findings of Varela-Neira, Vazquez-Casielles, and Iglesias-Arguelles (2008) and the meta-analytic results of Gelbrich and Roschk (2011), which suggest that IJ has a stronger relative effect than the other JDs. The second-strongest effect emerges for prefailure satisfaction, which, as expected, turned out to be a good predictor of postfailure satisfaction. Thus, carryover effects are present, and postfailure outcomes substantially depend on prefailure satisfaction levels. With an R-square of .67, two-thirds of the variance in postfailure satisfaction can be explained. This is relatively high in comparison with previous research that omits lagged satisfaction. For example, Maxham and Netemeyer (2002b, 2003) report lower R-square values ranging between .40 and .48 for satisfaction. Thus, the inclusion of prefailure outcomes levels significantly improves the prediction of postrecovery outcomes. PJ has the third-largest effect; however, it is not significant. Surprisingly, DJ is nonsignificant and exhibits the weakest impact. Similarly, some research also has found that DJ has the weakest impact on satisfaction (e.g., Maxham and Netemeyer 2002b; Severt 2002) and has no significant effect (Varela-Neira, Vazquez-Casielles, and Iglesias-Arguelles 2008). However, other research reports that DJ is of utmost importance and competes with IJ in terms relative strength (e.g., Davidow 2003a; Gelbrich and Roschk 2011; Martínez-Tur et al. 2006). The reasons for the dominance of IJ and the nonsignificant effects of the other JDs may be fourfold. First, in general, in a retailing context, aspects of IJ may be of particular importance. For example, Babakus, Bienstock, and Van Scotter (2004) and De Wulf, Odekerken-Schröder, and Iacobucci (2001) note that employing highly skilled and motivated service personnel is one of the most important success factors of retailing. Because no prior comparable analysis exists in a retail setting, further research should examine whether IJ always assumes a salient role in retailing. Second, unfortunately, for this study no information on failure type was available on an individual level. A satisfying resolution of outcome failures (including some monetary loss, e.g., product malfunctions) requires higher levels of DJ than mere service or process failures (Smith, Bolton, and Wagner 1999). Similarly, Maxham and Netemeyer (2003) find that DJ exerts the strongest impact on satisfaction in a more product-related failure context of online electronic consumer goods retailing.

In addition, meta-analytic results suggest that IJ is particularly important in service industries and for nonmonetary complaints (Gelbrich and Roschk 2011). Considering the possibility that the failures examined in this study were not accompanied by strong economic losses, this could partially explain the finding that DJ has a weak and nonsignificant effect on satisfaction. Third, this study controlled for prefailure satisfaction. Previous research has neglected prefailure assessments of relationship satisfaction. This study's finding that not all JDs are significant is potentially because of the inclusion of lagged satisfaction. In a supplementary analysis, I exclude prefailure satisfaction from eq. 5.3 , which turns PJ significant at the 10% level ($\beta$ = .16; t = 1.83, $p$ = .07).[37] This result indicates that some prior studies may have found significant effects of JDs on cumulative satisfaction because of the failure to control for prefailure satisfaction, and thus they may suffer from omitted variable bias. Potentially, some share of the variance in postfailure satisfaction can be explained by both JDs and prefailure satisfaction, which seems plausible when considering the evidence that justice perceptions depend on prior experience with the firm (Severt 2002). Fourth, the sample size was relatively small after exclusion of observations with missing values. PJ and DJ might have reached signficant levels with a larger sample size.

*Influences on purchase spending.* As expected, the results show that past sales are the best predictor for future purchase spending. This is in line with conventional wisdom on habitual buying and inertia effects, which suggests that past behavior is the best predictor of future behavior (e.g., Ajzen 2001). Past measures can substantially increase the explanatory power of models and often better account for the total model variance than other predictors. As De Cannière, De Pelsmacker, and Geuens (2009, p. 88) report,

[37] DJ however remains nonsignificant ($\beta$ = .03; t = .43, $p$ = .67).

"attitudinal antecedents and intentions fail to predict behavior when combined with past behavior," whereas these were typically showing significant effects when no lagged variables were included. The finding that satisfaction exhibits a significant effect even with the lagged purchase spending measure included indicates that postfailure satisfaction is of critical importance to customer loyalty. This finding explains additional variance and substantially contributes to the explanatory power and validity of the model. The R-square of .32 can be considered as remarkably high for a model with objective purchase behavior as a dependent variable (Seiders et al. [2005] report an $R^2$ of .10 in a similar, dynamic model of purchase behavior in retailing). This adds to the notion that information on performance failures and complaint handling outcomes can significantly enhance sales forecast models (van Oest and Knox 2011).

*The mediating role of satisfaction.* Among the JDs, IJ turned out to be the most important in predicting future purchase behavior. Postfailure satisfaction mediates the effect of IJ on purchase spending, and thus it could be shown that IJ translates into future revenues. This is in line with Chebat and Slusarczyk (2005), who report a direct effect of IJ on loyalty–exit behavior. Remarkably, the direct effect of IJ on purchase spending was slightly larger than the mediated effect through satisfaction. This suggests that cumulative satisfaction does not fully capture all relevant aspects for predicting loyalty behavior. Potentially, aspects included in IJ, such as human interaction and emotional clues, have a distinct effect which goes beyond the conceptualization of the mere satisfaction construct. Another possibility is that, as Baron and Kenny (1986) note, this is due to feedback effects or measurement error.

Due to the absence of a significant effect on satisfaction, no mediation could be detected for PJ and DJ. This may be because of the previously mentioned reasons. Moreover, PJ and DJ were examined for direct effects on purchase behavior, but no signficant relationship was detected. Therefore, this study suggests that PJ and DJ do not translate into postfailure buying behavior. However, in other contexts, JDs may assume different levels of importance and exhibit significant effects. Future research could try to explore context-specific contingencies for JDs.

### 5.5.2 Implications for Practice

As Davidow (2003b, p. 244) acknowledges, “there is a lot of value in investigating overall influences, such as which dimension is the most important one or what is the impact of a dimension on a specific postcomplaint customer behavior ....” Accordingly, this study shows how companies could apply the approach to understand the drivers that are most important for influencing the postrecovery buying behavior of their complainants. Companies can increase complaint management profitability by implementing efficient recovery strategies that appropriately consider the effect of JDs on postfailure outcomes. The model results inform managers how investments in complaint management dimensions translate into future revenues, providing guidance on the allocation of resources. With a hypothetical customer base of 1.5 million and with approximately 7% experiencing a serious performance failure, the monetary impact (MI) of a (one unit) performance improvement in IJ would lead to, ceteris paribus, the following financial effect:

$$\mathrm{MI}_{\mathrm{IJ}} = .55 \times 126.86 \times 105,000 = 7{,}326{,}165\,€.$$

This does not imply that this impact is generalizable to other contexts and providers, but it illustrates the substantial effect of an improvement in perceived justice and shows how managers can advance a profitable investment decision. The MI of the remaining dimensions and lagged customer satisfaction can similarly be assessed:

$$\mathrm{MI}_{\mathrm{PJ}} = .17 \times 126.86 \times 105,000 = 2{,}264{,}451\,€,$$

$$\mathrm{MI}_{\mathrm{DJ}} = .06 \times 126.86 \times 105,000 = \quad 799{,}218\,€, \text{ and}$$

$$\mathrm{MI}_{\mathrm{SAT}_{t-1}} = .49 \times 126.86 \times 105,000 = 6{,}526{,}947\,€.$$

According to these examples, retail managers would gain a higher benefit by investing proportionately more money in improving employee interaction skills than by offering generous and expensive compensation to complainants. With a comprehensive assessment of postcomplaint purchase drivers' MI and the inclusion of cost information, trade-offs between potential investment decisions in various response dimensions could be made. Ultimately, managers would be able to derive optimal solutions and establish efficient and profitable complaint management strategies because the approach makes complaint handling more accountable.

Of note, IJ has the strongest effect on satisfaction as well as purchase behavior. This may be because retailing naturally has many employee–customer interactions, and failures are frequently of a nonmonetary nature (e.g., personnel unavailable, bad advisory service, waiting times). Thus, companies should train their staff to be polite and respectful in their communications and to extend courteous behavior to complainants because "the ability of 'frontliners' to provide strong service recovery greatly affects customer loyalty" (Brown 2000). In addition, to achieve superior employee performance, managers could incentivize employee extra-role behaviors and ensure fair employee treatment

(Bowen, Gilliland, and Folger 1999; Maxham and Netemeyer 2003). In comparison with organizational responses aimed at DJ, improvement in IJ requires relatively low investments. For example, as one of the most cost-effective company actions, Goodman (2006) reports that friendly 90-second customer-staff interactions boost loyalty by 25%. Moreover, comprehensive training of employees not only contributes to a short-term amelioration of IJ and recovery performance but also has an substantial impact on general satisfaction levels and customers' perceived service quality (e.g., Liao and Chuang 2004). This study's finding that lagged satisfaction has the second-highest impact on postfailure satisfaction also supports the notion that highly skilled and motivated service personnel pays off on multiple levels and may have an effect on satisfaction that persists over time. Thus, fair customer interactions can lead to short- and long-term cost savings—very likely even beyond the scope of complaint management—and can be implemented with relatively small investments in personnel trainings, in turn possibly generating a great impact on the bottom line.

### 5.5.3 Limitations and Further Research

As with all research, this study is constrained by limitations that at the same time suggest areas for further research. First, I examined the impact of JDs on postfailure satisfaction and customer behavior in retailing, a popular noncontractual setting. The significance and magnitude of effects may be different in other contexts. Exchanges in Internet business, for example, typically do not comprise any form of personal interaction. Therefore, IJ should play a less salient role in e-commerce. Future research could replicate and extend the findings and investigate potential differences to other industries and contractual settings. Second, no individual information on the failure type was available. A satisfying resolution of outcome failures requires higher levels of DJ than mere process failures (Smith, Bolton, and Wagner 1999) and thus can have an impact on the relative effects of JDs. Consequently, whereas this study investigates the main effects of JDs on satisfaction and purchase behavior, further research might analyze possible interaction effects between the individual JDs and also include contingent variables (e.g., failure type, magnitude, responsibility) that might moderate the studied relationships. Third, selection bias might be an issue because the sample included only customers who were enrolled in the retailers' loyalty program. Thus, the estimated effects in this study might be somewhat attenuated because loyalty program members tend to have decreased negative perceptions when experiencing poor service encounters (Bolton, Kannan, and Bramlett 2000). Hence, future research may use transactional data from sources other than loyalty programs to eliminate such potential effects.

### 5.5.4 Conclusion

Despite these limitations, this study closes a frequently pinpointed research gap. By linking perceived justice to actual postfailure purchase behavior, it shows how investments in complaint management can be traded off in accordance with monetary return. More specifically, this research contributes to current knowledge in three ways: (1) It estimated the impact of perceived justice on postfailure satisfaction and purchase behavior, and (2) it examined the mediating role of satisfaction within this functional chain while (3) accounting for potential carryover effects of prefailure satisfaction levels in the model. Notably, the results show that IJ plays a crucial role, whereas the other dimensions have no significant impact. This finding suggests that organizational responses that include elements of personal interaction are of greater relevance than processes and compensation. The applied approach enhances the understanding of the drivers of post-failure purchase behavior and helps companies evaluate complaint handling strategies and obtain guidance for resource allocation.

# 6 The Moderating Effects of Recovery, Relationship, and Marketplace Characteristics on the Failure Resolution–Purchase Behavior Link

## 6.1 Overall Background

In general, research on product and service failures assumes that successful recoveries can restore damaged customer relationships to their prefailure state. According to Hart, Heskett, and Sasser (1990, p. 148), "a good recovery can turn angry, frustrated customers into loyal ones" and may "create more goodwill than if things had gone smoothly in the first place." Several studies support this notion, claiming that after positive recoveries, postfailure outcomes equal or even exceed prefailure levels (e.g., Magnini et al. 2007; Maxham and Netemeyer 2002a; McCollough, Berry, and Yadav 2000; Smith and Bolton 1998). However, because these studies rely on evidence based on self-reported, intentional data, questions remain whether a successful recovery from failure indeed translates into loyal postfailure repurchase behavior. From a theoretical perspective, behavioral loyalty may follow different mechanisms than attitudinal loyalty. For example, customers often fail to account for intervening contingency effects when predicting their own future behavior (Seiders et al. 2005). Thus, prior frameworks that use purchase intent as a proxy for behavior have potentially omitted important variables at play in failure scenarios. From a managerial perspective, postfailure repurchase behavior is an essential component of complaint management profitability, helping make recovery initiatives more accountable and thereby drawing oftentimes-lacking top management attention to the topic (Stauss and Schoeler 2004). Overall, extant research has produced limited evidence of postfailure behavioral loyalty and its boundary conditions. In particular, three important aspects have largely been neglected.

First, a substantial body of literature details how best to resolve performance failures, how such failures can affect customer relationships, and how to mitigate their potential negative consequences (for a summary, see Davidow 2003b; de Matos, Henrique, and Rossi 2007; Gelbrich and Roschk 2011; Orsingher, Valentini, and de Angelis 2010). However, although some studies investigate postfailure behaviors such as exit (Chebat, Davidow, and Borges 2011; Chebat and Slusarczyk 2005; Gustafsson, Johnson, and Roos 2005), change in customer share (van Doorn and Verhoef 2008), and purchase behavior (Evanschitzky, Brock, and Blut 2011; Gilly 1987; Gilly and Gelb 1982; von Wangenheim and Bayón 2007), evidence of the behavioral consequences in terms of their financial impact remains scarce. Attitudinal data cannot satisfactorily answer the questions of how much to spend on a recovery and how to allocate the money. Therefore, researchers have called for a database approach to complaint management (Rust and Chung 2006) because a quantification of the effects of failure and recovery on postfailure customer purchase behavior can help trading off efforts and planning efficient and effective recovery strategies (Davidow 2003b; Parasuraman 2006).

Second, when trying to answer the question whether damaged relationships can be restored to their prefailure levels, consideration of the prior relationship state is crucial because "the best predictor of loyalty after an experience is usually loyalty before that experience" (Brockner, Tyler, and Cooper-Schneider 1992, p. 241). Accounting for the prefailure relationship state is particularly important because "prefailure conditions can influence perceptions of recovery" (DeWitt and Brady 2003, p. 204) and performance failures may "affect relationships differently depending on the initial state" (van Doorn and Verhoef 2008, p. 124). Prior studies have found a moderating effect of prefailure relationship characteristics (e.g., Evanschitzky, Brock, and Blut 2011; Grégoire and Fisher 2008), but these studies largely rely on cross-sectional data from scenario-based surveys, lab experiments, or retrospectively interrogated customer experiences. Research designs with a dynamic, unbiased prefailure assessment of relationship health in real-life settings are scarce. Only a few field studies have assessed pre- and postfailure relationship outcomes (Maxham 2001; Maxham and Netemeyer 2002a,b; van Doorn and Verhoef 2008). However, these studies are limited to analyzing self-reported, intentional outcome measures. Thus, unlike prior work, I consider the prefailure relationship state and examine its direct and moderating effect on objective postfailure purchase behavior.

Third, extant studies have primarily focused on the design of complaint handling and largely neglected boundary conditions, such as situational moderators and factors beyond company control. Gilly and Gelb (1982, p. 327) recognize that there is "no evidence that once a company response is 'satisfactory,' the degree of satisfaction affects repurchase significantly. Presumably, other market factors take precedence." Thus, this gap remains unexplored, perhaps because recovery strategies have often been examined as if performance failures occur in a vacuum. As previously noted, there is a general lack of observational field studies. In a lab experiment, customers may express their intent to quit, whereas in a real-life setting, other considerations, such as marketplace characteristics, can constrain customers to remain with a company (Chebat and Slusarczyk 2005). Recently, Homburg, Fürst, and Koschate (2010, p. 280) noted that competition-related market conditions play a major role in failure situations, and they encourage researchers "to systematically consider moderating effects" in future frameworks. Thus, I do not focus primarily on the design of complaint handling but rather consider relevant contingency factors—namely, the oftentimes omitted marketplace and relationship characteristics. Table 6.1 shows that the study aims to make unique contributions by testing formerly understudied moderating variables in a failure context; specifying an integrative, dynamic model linking survey measures to objective, longitudinal transaction data; and thereby assessing determinants of postfailure purchase behavior.

Against this background, the overarching goal of this project is to determine whether and how a damaged customer relationship can be restored to a prefailure state with regard to actual purchase behavior. Overall, this research aims to make three key contributions: (1) to develop a comprehensive and integrative model of postfailure purchase behavior, (2) to examine how the prefailure relationship state affects postfailure purchase behavior, and (3) to analyze the moderating effects of recovery, relationship, and marketplace characteristics on the link between failure resolution and postfailure purchase behavior.

The remainder of this project proceeds as follows: In the next section, I develop the conceptual model building on prior research results. Then, I elaborate on the theoretical basis and deduce the hypotheses. Next, I outline the research methodology and test my propositions using hierarchical regression analysis.[38] Finally, I present the results and

[38] For a description of the research design, empirical setting, the data, and measures employed for this study, see Chapter 3.

**Table 6.1:** Prior Field Studies Investigating Postfailure Behaviors

| **Study** | **Context** | **Longi-tudinal** | **Repeated Survey** | **Account for Prefailure Relationship State**[a] | **Actual Behavior** | **Purchase Behavior** | **Characteristics Affecting Recovery Effectiveness** | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | **Recovery** | **Relation-ship** | **Market-place** |
| Gilly and Gelb (1982) | Oil industry | × | - | - | × | × | × | - | - |
| Gilly (1987) | Oil industry | × | - | - | × | × | × | - | - |
| Gustafsson, Johnson, and Roos (2005) | Telco | - | - | - | × | - | - | × | - |
| Chebat and Slusarczyk (2005) | Banking | - | - | - | × | - | × | - | - |
| Chebat, Davidow, and Borges (2011) | Banking | - | - | - | × | - | - | - | × |
| Maxham and Netemeyer (2002a) | Banking | × | × | × | - | - | × | - | - |
| Maxham and Netemeyer (2002b) | Banking | × | × | - | - | - | × | - | - |
| van Doorn and Verhoef (2008) | Transportation | × | × | × | - | - | × | × | - |
| Grégoire, Tripp, and Legoux (2009) | Misc. | × | × | - | - | - | × | × | - |
| Evanschitzky, Brock, and Blut (2011) | Fast food | × | - | - | × | × | × | × | - |
| **This study** | Retailing | × | × | × | × | × | × | × | × |

[a]Refers to a dynamic assessment of the prefailure relationship status, not retrospectively interrogated prior experiences.

discuss the implications of the findings.

## 6.2 Conceptual Development

The conceptual model is based on previous work predicting purchase behavior (Seiders et al. 2005). This research extends current knowledge by explicitly investigating performance failure situations. Within this scope, I consider recovery, relationship, and marketplace characteristics that are theoretically relevant to this particular context and, as I propose, constitute major influencing factors of postfailure purchase behavior. Figure 6.1 depicts the conceptual model.

With regard to *recovery characteristics*, several studies advocate that a good recovery from failure can restore satisfaction and other key relationship outcomes to a prefailure level (e.g., Maxham and Netemeyer 2002a; Smith and Bolton 1998), especially when a specific set of recovery dimensions is carefully considered and accurately executed (Davidow 2003b). A successful recovery predominantly consists of the resolution of the problem and a set of further recovery characteristics, such as staff friendliness, timeliness, and apology. In the proposed model, recovery characteristics comprise failure resolution and resolution speed, which reflect the most decisive aspects of any recovery. The degree to which a satisfactory failure resolution can be provided is the key determinant of postfailure outcomes, as a substantial body of literature demonstrates. In their meta-analysis, Gelbrich and Roschk (2011) find that the link between (transactional) recovery satisfaction and loyalty (intentions) is significant. The time required to resolve a failure has also been identified as a key ingredient in effective complaint management (SOCAP 1994) and is considered one of the most critical response dimensions in the recovery process (e.g., Davidow 2003b; Tripp and Grégoire 2011; Voorhees et al. 2009).

*Relationship characteristics* play a key role in failure situations and can substantially alter postfailure outcomes (e.g., Ganesan et al. 2010; Grégoire, Tripp, and Legoux 2009; Sajtos, Brodie, and Whittome 2010). Gelbrich and Roschk (2011) note a lack of studies analyzing the moderating role of relationship aspects in a failure context. Similarly, Homburg, Fürst, and Koschate (2010, p. 281) specify that "research should certainly also consider the perceived quality of the business relationship." Thus, I include relationship and affective commitment in the model. Commitment is considered a key variable in customer relationships (e.g., Fullerton 2003; Morgan and Hunt 1994) and constitutes an essential component of the relationship quality construct (De Wulf, Odekerken-Schröder, and Iacobucci 2001; Grégoire, Tripp, and Legoux 2009). Moreover, commitment plays a decisive role in failure scenarios (Ganesan et al. 2010; Mattila 2004), in particular with regard to postfailure purchase behavior (Evanschitzky, Brock, and Blut 2011).

*Marketplace characteristics* have been shown to moderate the satisfaction–repurchase behavior relationship (Seiders et al. 2005; Voss, Godfrey, and Seiders 2010) and are theorized to represent intervening contingency factors that customers fail to consider when they predict their own future purchase intent (Seiders et al. 2005). A basic assumption of most previous studies in the research domain of performance failures is that actual behavior stems from customers' attitudes and intentions. However, whether behavioral intentions indeed translate into actual behavior may depend on additional factors. Marketplace characteristics, such as competitor attractiveness, switching costs, and locational convenience, may take effect after the formation of consumer intent yet right before the execution of actual behaviors. These factors may prevent customers from behaving in accordance with their previously developed attitudes and thus are important to account for in models linking perceptual data to objective purchase behavior. Gelbrich and Roschk (2011) note a lack of studies analyzing the moderating role of switching barriers in failure-related research. Consequently, because of their high importance in this particular context, I incorporate switching costs (Chebat, Davidow, and Borges 2011) and locational convenience (Jones, Mothersbaugh, and Beatty 2003) as marketplace characteristics.

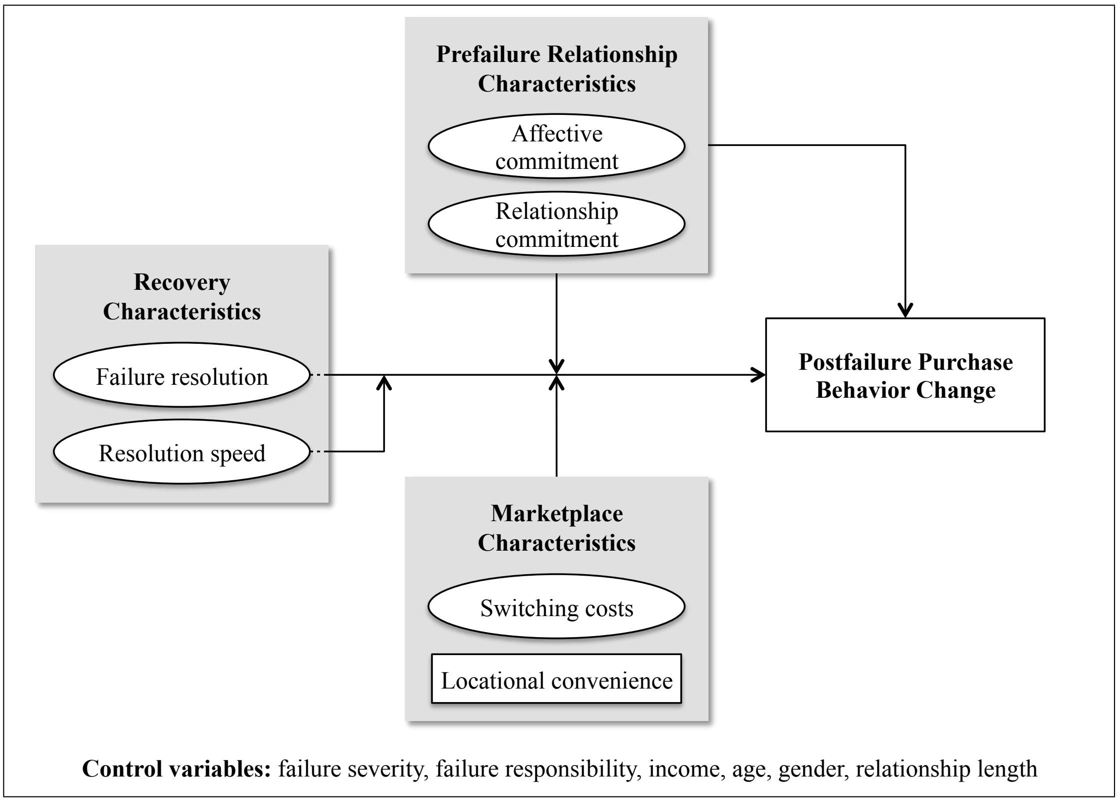


**Figure 6.1:** Hypothesized Model

The model depicted in Figure 6.1 conceptualizes failure resolution as customers' self-reported perceptions of how well the failure was resolved and purchase behavior as the objectively observed behavior after a failure. The default expectation is that failure resolution positively influences postfailure purchase behavior (Evanschitzky, Brock, and Blut 2011). Moreover, I predict that recovery, relationship, and marketplace characteristics moderate this link, and I propose that relationship characteristics of the prefailure state also have a direct effect on postfailure behavior. In addition to the hypothesized effects, the model includes controls for various situational factors that can affect postfailure purchase behavior. These comprise the severity of the failure and failure responsibility, both of which affect customer responses to failure recovery (e.g., Smith, Bolton, and Wagner 1999). Moreover, I control for the customer characteristics of age, gender (Mittal and Kamakura 2001), and income (Seiders et al. 2005), as well as relationship length (Reinartz and Kumar 2000), because these were identified as factors that can influence purchase behavior.

## 6.3 Theoretical Basis and Hypotheses

The conceptual framework presented in Figure 6.1 proposes three categories of variables that operate at different levels. Recovery characteristics explain variations in the failure resolution–repurchase relationship due to individual failure- and recovery-specific differences, relational characteristics reflect customers' investments in building relationships with a company, and marketplace characteristics account for variations related to market-level competition. I propose and subsequently test up to two moderating variables for each category, for which I expect an interaction effect, while controlling for main effects.

### Recovery Characteristics

Recovery characteristics encompass dimensions of organizational responses to complaints that affect postcomplaint customer behavior (Davidow 2003b). I examine failure resolution as a direct effect and resolution speed as a moderating factor. Because both variables are considered highly critical dimensions of the recovery process, they are likely to be among the most significant recovery-level influences.

*Failure resolution.* The degree to which a favorable problem resolution can be provided is the key determinant to recovery success and postfailure outcomes. Among other theories, consistency theory (Festinger 1957) can serve as a theoretical basis: Customers with a dissatisfying failure resolution should display attitude-consistent behavior and reduce or stop purchasing from the respective company, whereas satisfied customers who received an adequate company response during a failure episode should act consistently and return support to the company by maintaining future purchase levels. Prior research has frequently studied satisfaction with complaint handling and oftentimes found that it is positively related to various postfailure outcomes. For example, Bitner, Booms, and Tetreault (1990) find that customers exhibit positive reactions to encounters in which performance failures were followed by effective recoveries. Tax, Brown, and Chandrashekaran (1998) show a significant relationship of satisfaction with complaint handling to commitment and trust, and Maxham and Netemeyer (2002b) show that satisfaction with recovery is significantly related to purchase intent. Moreover, in their meta-analysis, Gelbrich and Roschk (2011) confirm the significant link between (transactional) postfailure satisfaction and Evanschitzky, Brock, and Blut (2011) and Gilly and Gelb (1982) provide evidence for the complaint satisfaction–purchase behavior link. Accordingly, the default expectation within the conceptual model is that failure resolution positively influences postfailure purchase behavior.

$H_1$ : *Failure resolution has a positive effect on postfailure purchase behavior.*

*Resolution speed.* Resolution speed reflects the perceived time to resolve a failure. The longer it takes to solve a problem, the more the attempted recovery turns into a second failure (a double deviation), creating increased frustration for the customer. Prior research suggests that a performance failure can mark a trigger point that initiates a cognitive-updating process leading to relationship reevaluation and consideration of alternatives (Smith and Bolton 1998; van Doorn and Verhoef 2008). With increased waiting time, a customer's feelings and perceptions of the failure episode may increasingly turn negative and give way to intense rumination (Strizhakova, Tsarenko, and Ruth 2012) about what would happen if the failure cannot be resolved; such thoughts may in turn evoke negative visions about future interactions. Furthermore, a desire to avoid further emotional and economic losses may arise (Grégoire, Tripp, and Legoux 2009), and the concrete idea of using alternative providers in the future may become mentally set. Consequently, this anticipation of reduced future purchase behavior can translate into actual change of purchase activity in cases of low resolution speed.

$H_2$ : *Resolution speed moderates (enhances) the positive effect of failure resolution on postfailure purchase behavior.*

**Relationship Characteristics**

Regarding the role of relationship perceptions in a performance failure situation, two opposing views are prominent in the literature. One the one hand, research has long advocated that strong customer relationships are among the key assets of a firm (e.g., Heskett, Sasser, and Schlesinger 1997) and mitigate the effects of a poor recovery on outcomes (Mattila 2001; Tax, Brown, and Chandrashekaran 1998). On the other hand, there is growing evidence that relationship strength may amplify customers' negative responses (Aaker, Fournier, and Brasel 2004; Ganesan et al. 2010; Grégoire and Fisher 2008; Grégoire, Tripp, and Legoux 2009; Johnson, Matear, and Thomson 2011). Thus, I propose that the role of prior relationship experience in failure situations is dependent on the type of bonds that are dominant in a relationship (Mattila 2004). For some customers, the relationship might be more of an economic or rational nature; others may perceive a stronger emotional attachment toward a certain provider (Gundlach, Achrol, and Mentzer 1995). In the same vein, this conceptual lens accommodates two distinct aspects of a typical failure situation: When a performance failure occurs, customers may suffer losses on an economic and emotional level (Smith, Bolton, and Wagner 1999). I use relationship commitment and affective commitment to capture these two dimensions. Commitment reflects a strong desire to maintain a relationship and is the most prominent perception representing its strength (Gustafsson, Johnson, and Roos 2005; Moorman, Deshpandé, and Zaltman 1993; Morgan and Hunt 1994). In addition, research suggests that prior experience with a company has an effect on postrecovery outcomes (Tax, Brown, and Chandrashekaran 1998) and supports the notion that "positive antecedent states can help mitigate the negative effects of service failure" (DeWitt and Brady 2003, p. 201). Thus, I propose a dynamic model in which the roles of prefailure relationship commitment and prefailure affective commitment represent the two distinct aspects of economic and emotional goodwill from a relationship asset perspective.

*Relationship commitment.* In general, commitment reflects a consumer's enduring desire to continue a valued relationship with a supplier and his or her willingness to apply considerable efforts to maintaining it (e.g., Moorman, Deshpandé, and Zaltman 1993). De Wulf, Odekerken-Schröder, and Iacobucci (2001) emphasize customers' willingness to make efforts to sustain the relationship as a necessary condition. Moreover, the conceptualization represents the more rational form of commitment and is largely based on economic considerations (Gustafsson, Johnson, and Roos 2005). It captures a willingness to make short-term sacrifices to realize long-term relationship benefits (Anderson and Weitz 1992)—that is, a general readiness "to go the extra mile" to remain a customer of a certain supplier (De Wulf, Odekerken-Schröder, and Iacobucci 2001). Accordingly, relationship commitment is likely to play an important role when performance failures occur. In such a situation, customers suffer from economic and emotional losses, such as time, effort, and money. Consequently, customers with high relationship commitment should be more forgiving because they are more willing to make these sacrifices to maintain the relationship.

$H_{3a}$ : *Prefailure relationship commitment has a positive effect on postfailure purchase behavior.*

$H_{3b}$ : *Prefailure relationship commitment moderates (enhances) the positive effect of failure resolution on postfailure purchase behavior.*

*Affective commitment.* Affective commitment is the more emotional construct and comprises psychological attachment, identification, and value congruence (Gundlach, Achrol, and Mentzer 1995). In an organizational context, Brockner, Tyler, and Cooper-Schneider (1992, p. 241) find that layoff survivors' "most negative reactions were exhibited by those who previously felt highly committed." Similarly, previous research on performance failures has proposed a so-called love-turns-into-hate effect (Grégoire, Tripp, and Legoux 2009), which should be particularly prevalent in relationships dominated by emotional bonds (Mattila 2004). In such relationships, customers' feelings of trust violation are stronger, and from this felt betrayal, they tend to retaliate against the company. Ganesan et al. (2010) propose a similar rationale and find that affective commitment negatively affects postfailure switching intentions during major transgressions. Evanschitzky, Brock, and Blut (2011, p. 420) obtain contrasting results with regard to a moderating effect of affective commitment that enhances the complaint satisfaction–postrecovery behavior link, but they admit that in other industries with high switching costs and "a failure of relatively high severity, the buffering effect may turn into an amplifying effect." Because my conceptualization of performance failures comprises severe transgressions, I propose an adverse effect, such that affectively committed customers have less favorable perceptions of recovery, are less forgiving and retaliate by buying less after occurrence of a performance failure.

$H_{4a}$ : *Prefailure affective commitment has a negative effect on postfailure purchase behavior.*

$H_{4b}$ : *Prefailure affective commitment moderates (mitigates) the positive effect of failure resolution on postfailure purchase behavior.*

**Marketplace Characteristics**

Customer purchase decisions are substantially determined by marketplace characteristics, such as switching costs, the offered supplier convenience, and availability of alternative suppliers (e.g., Jones, Mothersbaugh, and Beatty 2000; Smith and Bolton 1998). Gelbrich and Roschk (2011) ascertain that switching barriers are likely to be important moderators in failure scenarios. Similarly, Valenzuela, Pearson, and Epworth (2005) state that switching barriers represent important market conditions that need to be accounted for in research investigating effectiveness of recovery strategies. They also argue that "these conditions might affect positively or negatively ... customers' post-complaint behaviour" (p. 245). Furthermore, Estelami (2000) finds that competitive intensity has a direct influence on complaint resolution. Therefore, I include switching costs and locational convenience as contingency factors, which I hypothesize to moderate the relationship between failure resolution and postrecovery purchase behavior.

*Switching costs.* Burnham, Frels, and Mahajan (2003, p. 110) define switching costs as "onetime costs that customers associate with the process of switching from one provider to another." The most traditional understanding of switching costs is the time, money, and effort involved when changing providers (Jones, Mothersbaugh, and Beatty 2000). In a broader sense, switching costs do not need to be incurred immediately on switching and are not necessarily pure "economic" costs (Burnham, Frels, and Mahajan 2003). Rather, they may be perceived as impediments occurring along the switching process, such as search costs, learning costs, transaction costs, cognitive effort, loss of loyalty benefits, abandonment of routines and habits, and financial, social, and psychological risks (Fornell 1992; Lam, Shankar, and Murthy 2004). Conceptually, switching costs also reflect a buyer's dependence on a vendor often resulting from a lack of viable alternative providers (Klemperer 1995). Prior studies have produced mixed findings regarding the moderating effect of switching costs on the satisfaction–loyalty intentions link. For example, Jones, Mothersbaugh, and Beatty (2000) and Woisetschläger, Lentz, and Evanschitzky (2011) find a negative moderating effect, whereas Burnham, Frels, and Mahajan (2003) and Lam, Shankar, and Murthy (2004) find no significant interaction. However, intentional measures are a weak indicator of actual behavior. A single study examines the moderating impact of switching costs on the complaint satisfaction–exit behavior link and finds partial support (Chebat, Davidow, and Borges 2011). To my knowledge, no prior research has examined the moderating role of switching costs in a failure scenario with actual purchase behavior as an outcome variable, despite several calls for research in this area. For example, in their meta-analysis, de Matos, Henrique, and Rossi (2007) suggest that future studies should investigate postrecovery effects for customers with high versus low switching costs within a given industry. Similarly, Estelami (2000) conjectures that switching barriers affect complaint-handling outcomes and calls for future research efforts.

In general, a customer's loyalty will constantly be challenged when there are numerous providers competing with offerings similar in price and quality. Customers can switch vendors easily in markets with high competitive intensity because switching costs are typically low (Farrell and Klemperer 2007; Fornell 1992). More or less, people follow variety-seeking motives (McAlister and Pessemier 1982; Sánchez-García et al. 2012), and "when switching costs are low, consumers feel freer to experiment other providers even if they are satisfied" (Chebat, Davidow, and Borges 2011, p. 824). Customers may especially be tempted to do so when firms give them a reason to rethink their current business relationship. A performance failure can be such a reason and has been theorized to mark a trigger point that initiates a cognitive process of relationship reevaluation and consideration of alternatives (van Doorn and Verhoef 2008). Switching costs "make it costly for the customer to switch to another supplier" (Fornell 1992, p. 10). They may even constrain dissatisfied customers to remain with a company if they outweigh perceived switching benefits (Jones, Mothersbaugh, and Beatty 2000). Conversely, satisfied customers may decide to switch despite successful recoveries when switching costs are low because they may, for example, remember the hassle, still hold a grudge (Andreassen 2001; Grégoire, Tripp, and Legoux 2009), or want to try something new. A satisfying recovery can still entail reduced repatronage under low switching costs, whereas a dissatisfying failure resolution not necessarily causes a customer to reduce or stop purchasing under high switching costs. Thus, in a performance failure scenario, high (low) switching costs may enhance (mitigate) the effect of failure resolution on postfailure loyalty; that is, high (low) switching costs may buffer (amplify) potential negative consequences, such as reduced repurchase spending.

$H_5$ : *Switching costs moderate (enhance) the positive effect of failure resolution on postfailure purchase behavior.*

*Locational convenience.* Overall, prior research has conceptualized convenience as a five-dimensional, second-order construct reflecting a consumer's perceived time and effort costs of purchasing (Seiders et al. 2007). Locational convenience refers to the accessibility dimension of the broader construct and involves "providing a service to a consumer at a place that minimizes the overall travel cost" (Jones, Mothersbaugh, and Beatty 2003, p. 703). Typically, research has conceptualized it as a fixed cost reflecting the distance a customer must travel between, for example, his or her home address and a retail store (Bell, Ho, and Tang 1998). Locational convenience is conceptually distinct from switching costs (Seiders et al. 2005) in that it comprises the ongoing travel costs associated with using a provider in an established repeat purchase relationship; switching costs are onetime costs that are associated with switching from one provider to another (Burnham, Frels, and Mahajan 2003). The model includes locational convenience because it is one of the long-established and most powerful drivers of patronage behavior in retailing (e.g., Brooks, Kaufmann, and Lichtenstein 2008; Craig, Ghosh, and McLafferty 1984). Furthermore, accessibility is considered a prerequisite to the other convenience dimensions; without it "all other forms of convenience are irrelevant" (Seiders, Berry, and Gresham 2000, p. 81). Grewal, Levy, and Kumar (2009, p. 1) contend that though location "repeatedly gets cited as central to a retailer's success," research has paid limited attention to this strategic topic in recent years. Fox, Postrel, and McLaughlin (2007) find that store proximity (measured in travel time) is a key predictor of consumer spending and, thus, retailer revenues. Therefore, Grewal, Levy, and Kumar call for investigations of the role of critical location variables, such as proximity to customers.

Prior work partially confirms a negative moderating effect of locational convenience on the general satisfaction–loyalty intention relationship (Jones, Mothersbaugh, and Beatty 2003; Wu 2011), but no previous research has investigated its role as a potential moderator of recovery effectiveness in terms of postrecovery purchase behavior. Only a few viable alternatives are available in the marketplace when frequented providers have comparatively high locational convenience (e.g., high accessibility, short travel distance). The closer a customer lives to a retailer's store, the farther he or she will need to travel to purchase at a competitor's store. Because agglomeration of different types of stores is more beneficial to retailers than near stores that offer similar product categories (Miller, Reardon, and McCorkle 1999), it is less likely that a customer living near a specialty store will have a similar, close competing retail store. Thus, the customer will experience additional search and travel costs resulting from geographic dispersion of competitors' store locations. Models in retailer location research predicting consumers' store choice are largely based on cost minimization rationales (Brooks, Kaufmann, and Lichtenstein 2008), which can also serve to theoretically explain the moderating role of locational convenience in a failure context. Customers are most likely to visit the store with the lowest total shopping cost (Bell, Ho, and Tang 1998). A customer's preferred provider typically offers low total shopping costs (including high locational convenience), whereas alternative suppliers are less attractive in this regard. In a performance failure situation, this preference structure may change as a customer faces additional costs caused by the failure. Additional costs can comprise economic and emotional losses resulting from, for example, waiting time, travel costs, frustration, and anger. Failure recoveries may help minimize additional costs (losses), but most likely the customer's perceived total shopping cost associated with the transgressing provider will increase during a failure episode. Consequently, the customer's decision to switch to a competitor (or stay with the original provider) depends on whether postfailure shopping costs turn out to be lower (higher) for alternative providers than for the original supplier. Thus, I hypothesize that high locational convenience facilitates the effectiveness of failure resolution efforts. It strengthens the impact of successful recoveries on postfailure purchase behavior because the sum of standard costs and additional failure-incurred economic and emotional costs are kept minimal. For inert customers, high locational convenience may prevent them from defecting even after a dissatisfying recovery. Conversely, when locational convenience is low, consumers may be prone to reduce patronage despite a successfully recovered failure.

> $H_6$ : *Locational convenience moderates (enhances) the positive effect of failure resolution on postfailure purchase behavior.*

## 6.4 Methodology

As before, I use retail data[39] to examine the hypotheses. In this study, I investigate the change in purchase behavior after a performance failure and therefore calculate before-and-after differences, that is, the delta of 12 months of pre- and postfailure purchase behavior.[40] This measure serves as the final dependent variable. A more detailed description of the independent variables, including scales, validity and reliability results, appear in Tables C.3 and D.1 in the Appendix. Table 6.2 presents the means, standard deviations, and bivariate correlations among all study variables.

I use a hierarchical linear regression for hypothesis testing. The hierarchical approach is particularly suitable when investigating multiplicative terms and potentially correlated predictor variables (Bagozzi 1984; Cohen et al. 2003). Following Cohen et al. (2003), I standardize all independent variables to avoid nonessential problems with multicollinearity when building interaction terms. I present four hierarchical models to demonstrate the stability of individual coefficients and overall model fit. Model 1 includes the control variables, and Model 2 contains the main effects of failure resolution, resolution speed, switching costs, locational convenience, and prefailure affective and

[39] Chapter 3 provides a detailed description of the research design, data collection, measurements, and samples.

[40] The dependent variable is operationalized as the difference in 12 months of post- and prefailure purchase behavior ($t_1 - t_0$). For a more detailed description, see sec. 3.3.1.

**Table 6.2:** Descriptive Statistics and Correlations of the Study Variables (Project III)

| | **Variable** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | Δ Revenue | 1 | | | | | | | | | | |
| 2. | Affective commitment | $-.28^{***}$ | 1 | | | | | | | | | |
| 3. | Relationship commitment | $-.23^{**}$ | $.72^{***}$ | 1 | | | | | | | | |
| 4. | Switching costs | $-.21^{**}$ | $.52^{***}$ | $.39^{***}$ | 1 | | | | | | | |
| 5. | Locational convenience | $-.14^{*}$ | -.09 | .03 | -.06 | 1 | | | | | | |
| 6. | Resolution speed | $-.15^{*}$ | $.31^{***}$ | $.30^{***}$ | $.32^{***}$ | .07 | 1 | | | | | |
| 7. | Failure resolution | $-.26^{***}$ | $.20^{**}$ | $.17^{**}$ | $.24^{***}$ | .02 | $.50^{***}$ | 1 | | | | |
| 8. | Failure severity | -.11 | .08 | .13 | .07 | -.01 | -.00 | $-.19^{**}$ | 1 | | | |
| 9. | Failure responsibility | -.05 | -.05 | -.04 | .06 | -.11 | $-.21^{**}$ | -.14 | $.28^{***}$ | 1 | | |
| 10. | Relationship length | -.03 | $.19^{**}$ | $.14^{*}$ | .06 | -.10 | .12 | .10 | .06 | .05 | 1 | |
| 11. | Age | .03 | .13 | .08 | .04 | $-.14^{*}$ | $.24^{**}$ | .10 | -.10 | -.09 | $.22^{***}$ | 1 |
| | M | -495.27 | 3.61 | 4.52 | 3.69 | 11.13 | 4.32 | 3.88 | 4.89 | 4.97 | 11.85 | 48.32 |
| | SD | 958.38 | 1.86 | 1.66 | 1.91 | 8.51 | 2.24 | 2.24 | 1.79 | 2.14 | 6.17 | 12.35 |

$^{***}p < .01$, $^{**}p < .05$, $^{*}p < .10$.

relationship commitment. Model 3 adds the interaction terms between failure resolution and the proposed recovery, relationship, and marketplace characteristics. Although I do not put forth explicit hypotheses for three-way interactions, I include significant terms of the prior model in Model 4 to examine additional interrelationships. Model 4 exhibits the highest explanatory power, which suggests that three-way interactions are important. Table 6.3 details the results of the regression analysis.

Although the control variables were not significant in any of the models (all *p*s > .09), I maintained them in the analysis to enhance model robustness. Overall, except from Model 1, which includes only the controls, all models are significant (Model 2: $p < .05$; Model 3 and Model 4: $p < .001$). After inserting main effects, Model 2 explains 9% ($\Delta R^2 = .08$, $p < .01$) of the variance in postfailure purchase behavior.[41] In Model 3, R-square substantially increases to .17 ($\Delta R^2 = .08$, $p < .01$) after inclusion of recovery, relationship, and marketplace moderators, demonstrating that these factors play a key role in failure scenarios and are essential for predicting postfailure behavioral loyalty. Accounting for three-way interactions, Model 4 exhibits an even better fit, explaining 19% ($\Delta R^2 = .02$, $p > .05$) of the variance in the dependent variable.

To assess monetary effects and relative importance, Table 6.3 presents nonstandardized (b) and standardized ($\beta$) coefficients along with t-values. A comparison of effects across models indicates stable results in terms of significance levels and effect sizes, except for failure resolution in Model 2 ($\beta = -.22$, $p < .05$) and Model 3 ($\beta = .24$, $p > .05$). A common reason for unstable coefficients is multicollinearity, which was ruled out in this study (VIF < 10). Another cause can be a factor of particular relevance, which Model

41 Throughout the project, I refer to the more conservative adjusted R-square. Prior research recommends this when many predictor variables are used with a relatively small sample size because the adjusted R-square calculates explained variance accounting for the number of predictors and observations (e.g., Cohen et al. 2003).

**Table 6.3:** Regression Estimates Δ Revenue

| | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | b | $\beta$ | t-Value | b | $\beta$ | t-Value | b | $\beta$ | t-Value | b | $\beta$ | t-Value |
| **Direct Effects** | | | | | | | | | | | | | |
| Failure resolution | $H_1$ | | | | **-97.30** | **-.22**$^{**}$ | **-2.40** | 105.67 | .24 | 1.13 | 96.36 | .22 | .97 |
| Resolution speed | | | | | 14.37 | .01 | .14 | 23.45 | .02 | .24 | -.74 | -.00 | -.01 |
| Relationship commitment ($t_0$) | $H_{3a}$ | | | | 2.38 | .00 | .02 | 13.43 | .02 | .12 | 30.90 | .03 | .28 |
| Affective commitment ($t_0$) | $H_{4a}$ | | | | -218.18 | -.23$^{*}$ | -1.82 | **-262.79** | **-.28**$^{**}$ | **-2.14** | **-262.15** | **-.28**$^{**}$ | **-2.15** |
| Switching costs | | | | | 25.54 | .03 | .28 | 92.37 | .10 | 1.01 | 158.80 | .17 | 1.62 |
| Locational convenience | | | | | -135.37 | -.14$^{*}$ | -1.74 | 166.82 | .18 | 1.24 | 130.32 | .14 | .97 |
| **Moderating Effects** | | | | | | | | | | | | | |
| Failure resolution × resolution speed | $H_2$ | | | | | | | **302.94** | **.26**$^{***}$ | **3.05** | **345.17** | **.30**$^{***}$ | **3.44** |
| Failure resolution × relationship commitment ($t_0$) | $H_{3b}$ | | | | | | | 2.53 | .00 | .02 | 11.62 | .01 | .10 |
| Failure resolution × affective commitment ($t_0$) | $H_{4b}$ | | | | | | | 33.91 | .03 | .28 | -3.93 | -.00 | -.03 |
| Affective commitment ($t_0$) × relationship commitment ($t_0$) | | | | | | | | 131.37 | .14$^{*}$ | 1.62 | 140.86 | .15$^{*}$ | 1.75 |
| Failure resolution × switching costs | $H_5$ | | | | | | | **247.01** | **.26**$^{**}$ | **2.57** | **224.45** | **.23**$^{**}$ | **2.30** |
| Failure resolution × locational convenience | $H_6$ | | | | | | | **-98.14** | **-.60**$^{**}$ | **-2.52** | **-91.21** | **-.56**$^{**}$ | **-2.34** |
| Failure resolution × affective c. ($t_0$) × relationship c. ($t_0$) | | | | | | | | | | | -86.25 | -.11 | -.99 |
| Failure resolution × resolution speed × switching costs | | | | | | | | | | | **200.40** | **.21**$^{**}$ | **2.06** |
| **Control Variables** | | | | | | | | | | | | | |
| Income | | 77.14 | .07 | .83 | 26.58 | .03 | .29 | -5.83 | -.01 | -.07 | -13.89 | -.01 | -.16 |
| Age | | -7.43 | -.01 | -.08 | 36.49 | .04 | .40 | 37.13 | .04 | .42 | 27.80 | .03 | .32 |
| Gender | | 148.71 | .07 | .76 | 100.74 | .05 | .53 | 88.65 | .04 | .49 | 74.24 | .03 | .42 |
| Relationship length | | -70.04 | -.07 | -.85 | -35.73 | -.04 | -.45 | -9.26 | -.01 | -.12 | -9.33 | -.01 | -.12 |
| Failure severity | | -102.45 | -.09 | -1.02 | -118.81 | -.11 | -1.22 | -131.06 | -.12 | -1.39 | -104.09 | -.09 | -1.10 |
| Failure responsibility | | -16.25 | -.02 | -.17 | -62.64 | -.06 | -.65 | .50 | .00 | .01 | -29.06 | -.03 | -.30 |
| Constant | | -773.44$^{***}$ | | -2.73 | -206.89 | | -.64 | -311.72 | | .988 | -306.05 | | -.94 |
| Model F value | | .56 | | | 2.12$^{**}$ | | | 2.67$^{***}$ | | | 2.68$^{***}$ | | |
| $R^2$ (adjusted) | | .01 | | | .09 | | | .17 | | | .19 | | |
| $\Delta R^2$ (adjusted) | | | | | .08$^{***}$ | | | .08$^{***}$ | | | .02 | | |

$^{***}p < .01$, $^{**}p < .05$, $^{*}p < .10$; n.s. = not significant. Notes: N=155, nonstandardized regression coefficient (b) values in €

2 did not consider (Kennedy 2005), as "often a counterintuitive, significant estimate results from the omission of a key variable" (Wooldridge 2002b, p. 134). Because I follow a contingency approach and propose that the effectiveness of failure resolution is dependent on moderating factors, this appearance is essentially in line with the basic theoretical argument. Although the main effect is no longer significant in Model 3, failure resolution has a significant effect in conjunction with three of the proposed moderating factors. Moreover, the coefficient of the main effect now offers directional support. Thus, it is concluded that the potential issue of omitted variable bias was successfully removed from Model 2 after inclusion of the moderating variables in Model 3. Because other results are similar across models, I refer to the best-fitting model (Model 4) to present the empirical results.

## 6.5 Results

*Recovery characteristics.* As previously mentioned, the default expectation and main effect of failure resolution is nonsignificant but offers directional support ($\beta$ = .22, $p$ > .05). Thus, the results do not provide support for $H_1$, suggesting that recovery efforts per se do not affect postfailure purchase behavior. Regarding the moderating role of resolution speed ($H_2$), the failure resolution $\times$ resolution speed interaction was significant ($\beta$ = .30, $p$ < .001). This indicates that the relationship between failure resolution and postfailure purchase behavior is significantly different depending on the time required to resolve the performance failure. To provide greater insight, I plot this interaction on the basis of median splits.[42] Figure 6.2, Panel A, depicts the relationship between failure

[42] It is a well-known phenomenon that in noncontractual settings, cohorts of customers exhibit negative trends in purchase behavior over time (e.g., Reinartz and Kumar 2000). This is potentially due to satiation, that is, a decreasing individual-level demand in certain product categories (Voss, Godfrey,

resolution and postfailure purchase behavior for customers who experienced high/low resolution speed in the categories of high/low degrees of failure resolution. For the category of high satisfaction with failure resolution, customers do not change their purchase behavior with high resolution speed ($+\Delta 11$ €; .8%). With low resolution speed, for both categories—high/low degrees of failure resolution—customers substantially reduce purchase volume by $-\Delta 500$ €/$-\Delta 469$ € ($-37.9\%/-35.6\%$). With unsuccessful failure resolution and high resolution speed, customers reduce their purchase activity the most ($-\Delta 755$ €; $-57.4\%$), which is probably due to a halfhearted, unsuccessful attempt to resolve the problem quickly without serious effort, resulting in a double deviation and increased frustration for the customer.

*Relationship characteristics.* Regarding the main effect, the results offer directional support but no significant effect for prefailure relationship commitment ($H_{3a}$; $\beta = .03$, $p > .05$). Similarly, the relationship commitment $\times$ failure resolution interaction ($H_{3b}$) is not supported ($\beta = .01$, $p > .05$). Apparently, prior relationship commitment neither carries over to the postfailure state (loyal purchase behavior) nor leads to more favorable customer perceptions of the failure resolution and thereby enhances its link to postfailure purchase behavior. Thus, the results reveal that relationship commitment is not effective in building a protective relationship layer that can buffer negative effects from performance failures. Regarding the proposed effects of prefailure affective commitment, the model results are supportive of the main effect ($H_{4a}$; $\beta = -.28$, $p$

and Seiders 2010). Similarly, in the database, a slight negative trend in global purchase levels occurred. The average reduction of annual purchase spending in the control (nonfailure) group was 205 €. For reasons of better comparability and interpretabilty, I therefore trend-adjusted the presented graphical results accordingly for the failure group in Figure 6.2 by anchoring the origin of the scale at +205 €. This way, the exhibited origin (0) represents the average purchase level of customers who had not experienced a failure. Figure 6.2 displays the percentage change of purchase behavior in terms of annual spending. The mean prefailure annual purchase volume is 1316,52 €.

$< .05$) but not of the interaction effect ($H_{4b}$; $\beta = -.00$, $p > .05$). This suggests that customers high in prefailure affective commitment retaliate by substantially reducing postfailure purchase activity ($-\Delta 262$ €; $-19.9\%$). Because the affective commitment $\times$ failure resolution interaction is nonsignificant, the model indicates that this retaliatory mechanism is present regardless of any attempted and potentially successful failure resolution. This adverse effect of affective commitment is in line with evidence from prior research. However, for the first time it renders evidence of a behavioral manifestation of the love-turns-into-hate effect in terms of purchase levels. I also examined additional moderating effects because prior research suggests an interaction of affective commitment with calculative commitment (Ganesan et al. 2010). Thus, although I did not put forth specific hypotheses, I also tested for a combined effect of both forms of commitment and an additional three-way interaction of both commitment forms with failure resolution. However, neither the affective commitment $\times$ relationship commitment interaction ($\beta = .15$, $p > .05$)[43] nor the affective commitment $\times$ relationship commitment $\times$ failure resolution interaction ($\beta = -.11$, $p > .05$) reveals a significant result.

*Marketplace characteristics.* For both switching costs and locational convenience, the results are consistent with expectations. The data provide support for $H_5$; switching costs significantly moderate the failure resolution–purchase behavior link ($\beta = .23$, $p < .05$). Figure 6.2, Panel B, depicts this interaction. Regarding the moderating role of locational convenience, $H_6$ receives support ($\beta = -.56$, $p < .05$); Figure 6.2, Panel C, plots this interaction. For both moderators—switching costs and locational convenience—the results follow a similar pattern: Customers who received a favorable failure resolution essentially maintain their previous purchase levels under high switch-

43 However, the interaction term is significant at the 10% level ($p = .091$) and might reach the 5% threshold if a larger sample were available.

ing costs (high locational convenience), with only an insignificant reduction of purchase volume of $-3.0\%/-\Delta 40$ € ($-4.8\%/-\Delta 63$ €). However, under low switching costs (low locational convenience), a notable decrease of $-13.4\%/-\Delta 176$ € ($-11.8\%/-\Delta 156$ €) emerges. Customers who received a dissatisfying failure resolution reduce repatronage by $-27.4\%/-\Delta 361$ € ($-26.1\%/-\Delta 344$ €) when facing high switching costs (high locational convenience) and repurchase substantially less under low switching costs (low locational convenience), with a delta of $-53.5\%/-\Delta 704$ € ($-60.6\%/-\Delta 798$ €). These results suggest that customers maintain purchasing levels when the recovery is successful and using alternatives is not easy. Under low switching costs (low locational convenience), a performance failure may mark a trigger point that sparks customers to consider and try alternative providers as they reduce purchase spending, even when failure resolution is successful. Unsuccessful recoveries lead to severe consequences under low switching costs (low locational convenience), with a reduction of prior purchase volumes by more than half. Though not specifically hypothesized, I examined an additional three-way interaction of failure resolution $\times$ resolution speed $\times$ switching costs and find that it is significant ($\beta = .21$, $p < .05$). Figure 6.2, Panel D, exhibits the interaction plot. High switching costs substantially decrease the slope of the interactions. Conversely, low switching costs intensify potential losses from reduced postfailure purchase behavior, as indicated by the steeper slope of interactions. This pattern of results is similar to that found for the two-way interactions: When failure resolution is favorable and resolution speed is high, purchase levels remain unaltered, but any deviation from this configuration leads to a reduction in repatronage, which becomes even more severe if failure resolution is unsuccessful, speed is slow, and switching costs are low. Figure 6.3 summarizes the results of the hypotheses tests.

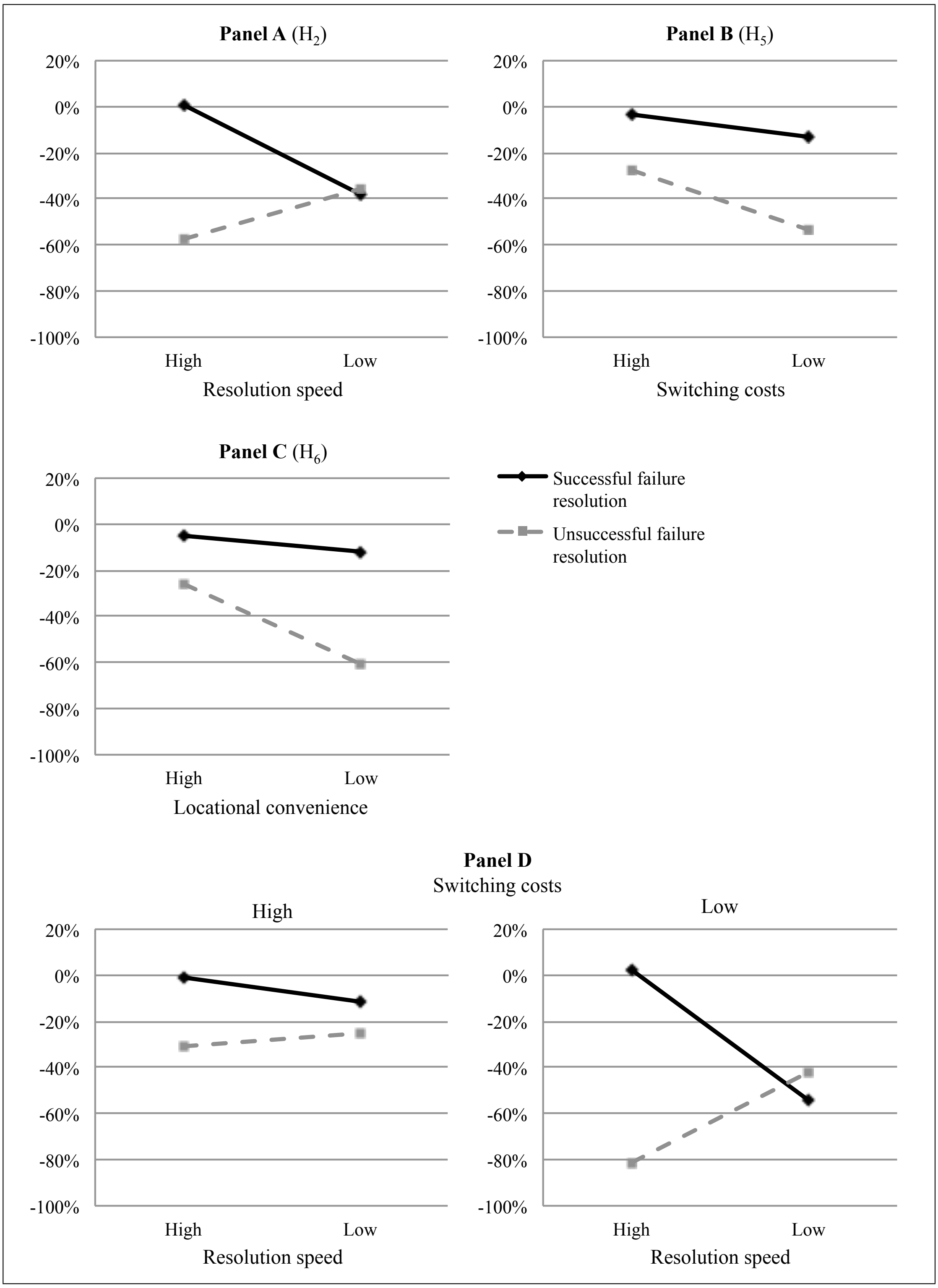


**Figure 6.2:** Significant Interaction Plots

- ☐ $H_1$: *Failure resolution has a positive effect on postfailure purchase behavior.*
- ☑ $H_2$: *Resolution speed moderates (enhances) the positive effect of failure resolution on postfailure purchase behavior.*
- ☐ $H_{3a}$: *Prefailure relationship commitment has a positive effect on postfailure purchase behavior.*
- ☐ $H_{3b}$: *Prefailure relationship commitment moderates (enhances) the positive effect of failure resolution on postfailure purchase behavior.*
- ☑ $H_{4a}$: *Prefailure affective commitment has a negative effect on postfailure purchase behavior.*
- ☐ $H_{4b}$: *Prefailure affective commitment moderates (mitigates) the positive effect of failure resolution on postfailure purchase behavior.*
- ☑ $H_5$: *Switching costs moderate (enhance) the positive effect of failure resolution on postfailure purchase behavior.*
- ☑ $H_6$: *Locational convenience moderates (enhances) the positive effect of failure resolution on postfailure purchase behavior.*

**Figure 6.3:** Summary of Results of Hypotheses Tests

## 6.6 Discussion

The overarching goal of this project was to determine whether and how a damaged customer relationship can be restored to its prefailure state in terms of actual purchase behavior. This research extends current knowledge by providing an integrative framework and dynamically assessing the complex interrelationships of failure, recovery, relationship, and marketplace characteristics, with observed postfailure purchase behavior as an outcome variable. Considering the general difficulty in uncovering moderating effects (e.g., Aiken and West 1991) the overall level of support for the hypotheses is surprising. For the recovery characteristics, both failure resolution and resolution speed affect postfailure purchase behavior. Moreover, the results provide support for the notion that the prefailure relationship state (i.e., affective commitment) is an important determinant of postrecovery outcomes. Regarding marketplace characteristics, both proposed moderating effects (i.e., switching costs and locational convenience) were also significant. Overall, the model explains 19% of the variance in the dependent variable, which is rather high for a model predicting actual purchase behavior with predominantly attitudinal data. For example, Seiders et al. (2005) report an adjusted R-square of .10. As the modeling literature indicates (van Oest and Knox 2011), performance failures apparently offer substantial explanatory power for prediction of future purchase behavior. The remainder of the discussion of implications for research and practice centers on two focal questions: (1) How does the prefailure relationship state affect postfailure purchase behavior? and (2) Do the recovery and marketplace characteristics moderate the link between failure resolution and postfailure purchase behavior?

### The Role of Prefailure Relationship Characteristics

Two schools of thought exist regarding the role of positive prior experience and strong customer relationships with a provider. For failure situations, many researchers have found evidence for a "protection" effect that buffers negative consequences resulting from performance failures (e.g., Evanschitzky, Brock, and Blut 2011; Hess, Ganesan, and Klein 2003; Tax, Brown, and Chandrashekaran 1998). Conversely, growing evidence suggests that strong relationships amplify negative responses (e.g., Ganesan et al. 2010; Grégoire and Fisher 2008; Grégoire, Tripp, and Legoux 2009; Johnson, Matear, and Thomson 2011; Mattila 2004). Considering the unfavorable findings regarding the adverse effect of prefailure affective commitment ($H_{4a}$), I add to the second position and extend prior knowledge by showing that this effect translates into actual postfailure purchase behavior. This finding provides behavioral evidence that there is a "dark side" of strong customer relationships (Anderson and Jap 2005). Moreover, the finding demonstrates that the prefailure relationship status matters: regardless of the failure resolution, customers with high prefailure affective commitment will retaliate and reduce repatronage behavior. I did not find a significant effect of relationship commitment, which would suggest that customers high in prefailure relationship commitment are not more forgiving and that relationship does not buffer negative consequences of performance failures in terms of a change in purchase behavior. I proposed that customer relationships can be characterized and dominated by emotional and economic bonds and that, depending on the intensity of one or the other aspect, an amplifying or buffering effect of negative failure consequences would emerge. By examining relationship and affective commitment as relationship characteristics, I tried to accommodate these two distinct aspects accordingly. However, relationship commitment was not significant, and thus no support emerges for the economic mechanism theorized ($H_{3a}$). Moreover, both relationship characteristics did not moderate the failure resolution–purchase behavior link ($H_{3b}$ and $H_{4b}$). Evanschitzky, Brock, and Blut (2011) find a buffering, moderating effect of affective commitment on purchase behavior. This contrasts this study's finding of an amplifying main effect that I ascribe to context differences. The authors analyzed complaints to a fast-food delivery chain—a low-involvement context with failures of low severity. In contrast, in this study I used a medium-involvement retail context and considered only serious performance failures. Further research should try to disentangle these conflicting results by considering such contextual factors.

For managers, the finding that previously affectively committed customers substantially retaliate against the company by reducing purchase volume suggests that marketing efforts aimed to increase emotional loyalty of customers can have a clear downside. An analysis of the marginal effect on postfailure purchase behavior reveals that, all else being equal, a one-point increase in prefailure affective commitment leads to a revenue decrease of 262 € (19.9%) the subsequent year for the retailer. Thus, strong emotional attachment can backfire in performance failure situations and should be taken into account when evaluating potential investments in customer relationship initiatives. Moreover, managers need to be aware that some customers are highly sensitive to emotional clues and violations of trust. Therefore, particular cautiousness to this emotional side is required in any customer interaction and should be addressed, for example, by training service personnel at customer touch points and establishing adequate behaviors in complaint-handling guidelines.

### The Moderating Role of Recovery and Marketplace Characteristics

With regard to the moderating effects, the general results of the analyses render a positive view on failure and recovery; I find support for the notion that a damaged relationship can be restored to its prefailure level in terms of purchase activity. However, in contrast with conventional wisdom from prior studies, high-quality recovery efforts alone are not always effective in doing so. Rather, other intervening variables not always under company control also play a role. Most intriguing is the evidence that recovery efforts per se do not effectively influence postfailure purchase behavior. Many studies show that recovery satisfaction has a significant, positive main effect on postfailure purchase intent and overall satisfaction. However, this apparently does not hold for observed purchase behavior as a dependent variable. Potentially, customers' attitudes can be favorable and reach prefailure levels after a convincing recovery, but when they are actually ready to purchase again, customers might recall the hassles of that conflict situation and—in that particular situation—likely give other providers a chance, especially if the effort required to do so is low. In line with this argument, the results suggest that recovery is effective with regard to purchase behavior only in conjunction with moderating factors. For example, a critical aspect is for the failure to be resolved quickly. No substantial reduction in purchase volume was detected in the case of a quick and successful recovery. Otherwise, customers have time to reevaluate their relationship and may begin to seriously consider shifting future purchase spending to competitors. As prior studies suggest, a performance failure may trigger this cognitive process, wake up customers from their business-as-usual routine, and potentially lead to the concrete idea and action of using alternative providers in the future. In such a situation, the findings support the notion that marketplace characteristics play a significant role. Under low switching costs and low locational convenience, investments in recovery are at high risk. Under these circumstances, patronage behavior will be reduced substantially when the recovery takes a long time, even when the problem is solved successfully. The worst case is a quick but unsuccessful attempt to resolve the failure (a double deviation) and low switching costs. Annual purchase spending is reduced by $-81.2\%/-\Delta 1069$ € (Figure 6.2, Panel D), which is close to relationship termination.

Thus, managers should ensure a fast problem resolution at all times and empower customer-care employees accordingly to facilitate quick decision making. Moreover, practitioners should understand that recovery effectiveness is contingent on marketplace characteristics, and therefore investments do not pay off per se. Accordingly, an adaptive approach to complaint handling is necessary. Overall, it would be worthwhile for managers to estimate the marginal (and monetary) effect of critical contingency factors for their most typical performance failure scenarios. Doing so would enable them to answer the questions of how much to spend on a recovery and how to allocate the money. Guidelines describing adequate complaint-handling policies for various configurations of recovery, relationship, and marketplace characteristics could be developed and recoveries—most adequate and effective in influencing postfailure purchase behavior—could be performed accordingly. For example, practitioners could account for the number of viable alternatives complainers have in terms of switching costs or locational convenience, when deciding on how much to invest in an individual failure resolution. However, such policies should be designed with caution, giving attention to other, more indirect outcomes as well, because these can also have severe financial consequences. For example, by no means should managers completely dispense recovery efforts in unwinnable cases in order to avoid, for example, negative word of mouth or online public complaining.

### Limitations and Conclusion

This study provides substantial insight into the behavioral consequences of performance failures, but it is also subject to several limitations that might be overcome in further research. First, the study was conducted in a noncontractual setting in the retailing sector. Future studies might replicate the findings and investigate potential differences in other industries and contractual settings. Second, because of data limitations, I could not assess complaint management profitability. In general, it is difficult to acquire information about company cost structures, and even more difficult is the task of allocating costs appropriately to respective customers and specific complaint-handling processes. To assess a return on complaint management, future studies would need to acquire this information for the benefit of investigating efficient recovery strategies. Third, the results provide evidence of the existence of a dark side of strong relationships by demonstrating that customers with strong emotional bonds in the prefailure state substantially reduce their purchase activity; as such, this again calls for a reevaluation of the benefits of customer relationship management. Future studies could shed light on which forms of relationship management efforts yield competitive advantage, particularly in failure situations. Finally, recovery efforts were not effective per se; rather, they were contingent on recovery, relationship, and marketplace characteristics. This finding implies that resolution speed, switching costs, and locational convenience moderate the effectiveness of recovery efforts. Further research should systematically explore additional variables that potentially moderate the influence of recovery efforts on observed loyalty behaviors. In summary, the study results render a positive view by demonstrating that postrecovery purchase activity can potentially be restored to its prefailure level. However, given the strong influence of marketplace characteristics, a similar study in the rapidly expanding e-commerce sector might lead to less positive outcomes because of low switching barriers. This context may represent an ultimate challenge for the design of successful complaint management strategies.

# 7 General Discussion and Conclusion

This dissertation pursues the overarching goal of studying postfailure purchase behavior and its determinants. Postfailure loyalty is the central outcome variable of the research field; however, prior work has largely neglected to investigate actual repurchase behavior as an important loyalty outcome. From a theoretical perspective, behavioral loyalty may follow different mechanisms than intentional loyalty. From a managerial perspective, postfailure repurchase behavior is a key element of complaint management profitability, helping make investments in complaint handling and recovery initiatives more accountable. The thesis examines the impact of failure, recovery, and contingency factors on postfailure purchase behavior and quantifies their relative (monetary) effects. As such, it provides valuable theoretical insights and managerial implications.

This chapter presents a general discussion of the major findings and proceeds as follows: First, I summarize the key results of the individual projects. From this, I discuss the overall implications for theory, research, and management. Finally, I close the dissertation with a general conclusion and outlook.

## 7.1 Summary of the Key Results

In the endeavor to investigate postfailure purchase behavior and its determinants, I address three important research gaps in three empirical projects, which are summarized in the following subsections.

*Project I: The Causal Effect of Performance Failure on Relationship Outcomes*

The first project focuses on the research question of how performance failures affect relationship outcomes. Overall, *Project I* aims to contribute to prior research by (1) comprehensively assessing the average relationship damage of performance failures on attitudinal and behavioral outcomes, (2) clearly establishing causality, and (3) estimating the financial impact in terms of postfailure purchase behavior. Building on equity theory and Hirschman's theory of exit, voice, and loyalty, a negative causal effect of performance failure on six relationship outcomes—that is, satisfaction, repurchase intent, word-of-mouth intent, share of wallet, average transaction value, and annual customer purchase spending—is hypothesized and tested using a matching methodology combined with difference-in-differences estimation. This technique is particularly suited for causal inference as it enables an analysis and comparison of transaction behavior and attitudes across failure and nonfailure customer groups before and after a performance failure for a substantial period and on a detailed level. The results suggest a negative effect of performance failure on satisfaction, word-of-mouth intent, average value per transaction, and annual customer purchase spending. A projection of financial effects shows that performance failure has a strong negative impact on customer equity.

*Project II: The Effects of Perceived Justice on Postfailure Purchase Behavior*

The second project concentrates on the research question of how perceived justice—that is, perceptions of interactional, procedural, and distributive justice—affects postfailure behavioral loyalty. *Project II* aims to contribute to prior research by (1) investigating the effect of perceived justice dimensions on postfailure satisfaction and purchase behavior, (2) analyzing whether satisfaction mediates the effect of justice perceptions on purchase behavior, and (3) accounting for prefailure levels of satisfaction, which are examined for potential carryover effects. Building on justice theory, hypotheses are derived and tested in a dynamic, multiple equation model with seemingly unrelated regression estimation. The results show that interactional justice plays a crucial role as it affects both postfailure satisfaction and purchase behavior. Moreover, satisfaction fully mediates the link between interactional justice and purchase behavior. In addition, carryover effects are present as prefailure outcomes turn out to be a good predictor of postfailure outcomes. No significant effects emerge for procedural and distributive justice. This suggests that elements of personal interaction in organizational response to failures are of greater relevance for postfailure loyalty than processes and compensation. The results highlight the importance of consumers' perceived justice with complaint handling as well as their responsiveness to different justice dimensions and thereby enhance understanding the drivers of postfailure purchase behavior.

*Project III: The Moderating Effects of Recovery, Relationship, and Marketplace Characteristics on the Failure Resolution–Purchase Behavior Link*

The third project strives to answer the research questions of whether and how a damaged customer relationship can be restored to its prefailure state in terms of actual purchase behavior. Overall, *Project III* aims to make three key contributions: (1) to develop a dynamic, integrative model of postfailure purchase behavior, (2) to examine how the prefailure relationship state affects postfailure purchase behavior, and (3) to analyze the moderating effects of recovery, relationship, and marketplace characteristics on the link between failure resolution and postfailure purchase behavior. Building on the theories of relationship marketing and switching costs, a conceptual model, which suggests a contingency approach to postfailure purchase behavior, is developed and subsequently tested with hierarchical regression analysis. The results indicate that postfailure purchase behavior is influenced by failure resolution, resolution speed, switching costs, locational convenience, and prefailure affective commitment. Overall, successful and speedy failure resolution can effectively restore purchase activity to its prefailure level. However, in cases of low switching costs and low locational convenience, investments in recovery are at risk because even successful recoveries can lead to a reduction in purchase spending, particularly when delayed. Moreover, customers with high prior affective commitment significantly reduce their repurchase spending regardless of recovery success, which provides behavioral evidence of the dark side of strong customer relationships. The results reveal the relative (monetary) impact of different configurations of situational factors on recovery strategies.

## 7.2 General Discussion

### 7.2.1 Implications for Theory and Research

Overall, this thesis aims to contribute to an enhanced understanding of postfailure outcomes and its determinants. More specifically, the focal outcome variable of this dissertation's empirical projects is behavioral loyalty—operationalized as *actual purchase behavior*—, which was studied in conjunction with its most important influencing factors. The results contribute to theory by incorporating, analyzing, and confirming determinants of postfailure loyalty from five different categories: (1) organizational response/recovery characteristics, (2) perceived justice of the recovery, (3) the prefailure relationship state, (4) relationship characteristics, and (5) marketplace characteristics. Figure 7.1 depicts which categories of outcome determinants were investigated by the individual projects and how these are positioned within the general research framework that was introduced in section 2.3 (Figure 2.1). The following discussion of general implications for theory and (further) research will first be conducted along these five categories, then more general implications will be highlighted.

*Organizational response/recovery characteristics.* Overall, the thesis renders a positive view on failure and recovery because it was found that damaged customer relationships can be restored to their prefailure state in terms of actual purchase behavior. However, a successful failure resolution does not always maintain postfailure loyalty. In *Project III*, no significant main effect emerged for the relationship between failure resolution and postfailure purchase behavior. This suggests that recoveries are not effective per se; rather—as the further results show—their effectiveness is dependent on contingency

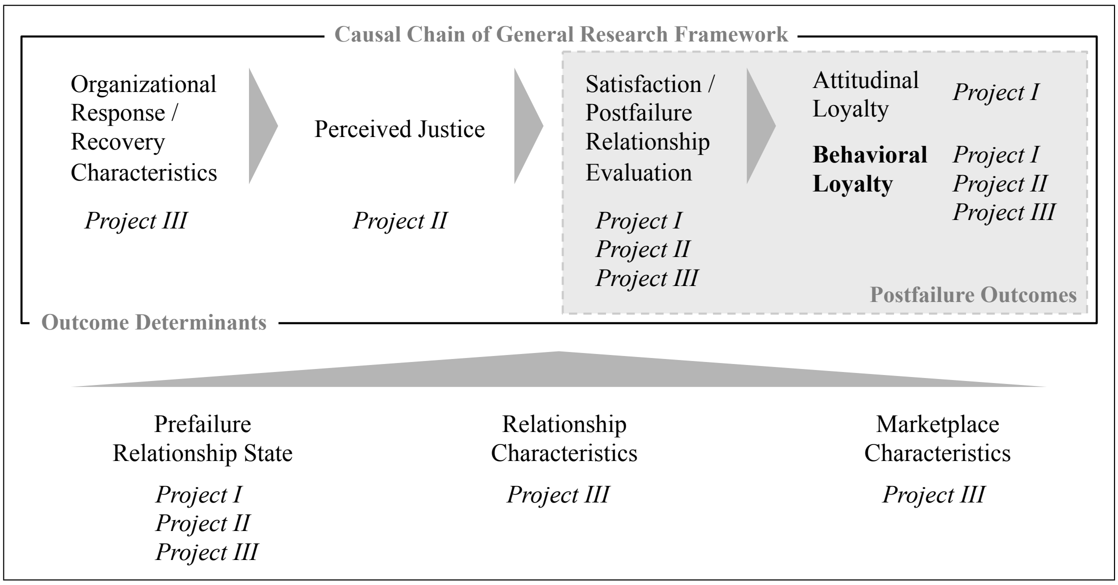


**Figure 7.1:** Summary of Individual Project Contributions to General Research Framework

factors. The findings highlight the critical moderating role of the recovery characteristic resolution speed. It was theorized that the longer the duration of the failure resolution, the more likely it is that complainants start considering switching their business to competitive retailers. When the transgressing provider takes too long to come up with an appropriate remedy, customers may already have decided to switch to another supplier at the time when the solution arrives and thus, the provider misses the chance to restore customer loyalty. Theoretically, this finding extends current knowledge by demonstrating that a long waiting time for failure resolution apparently not only yields anger and regret (Voorhees et al. 2009), but also significantly obstructs a recovery's effectiveness in terms of postfailure purchase behavior. When failure resolution is successful and resolution speed is high, purchase levels remain unaltered, but in case of low resolution speed, repatronage may be reduced despite a successful recovery.

It has to be acknowledged that resolution speed was subjectively measured. Thus, it is not clear what the average waiting time was in absolute terms. That is, it could not be analyzed which duration can be considered acceptable for consumers before recovery effectiveness is negatively affected. Further research could shed light into this and use a database approach which objectively measures time to resolve a failure. In general, research could pursue such an approach when investigating organizational responses and try determining the critical thresholds that mark the turning point at which favorable customer perceptions become unfavorable and eventually lead to dissatisfaction and defection. For example, research questions that future works may address could be: How long can waiting times be until satisfaction ratings and postfailure loyalty drop significantly? How much time should be devoted to listening to customers for conveying the impression that their complaint is taken seriously and that considerable efforts to resolve the failure will be made? And how generous should the remuneration be? Naturally, objectively measured variables are less decisive for customer behavior than a complainant's subjective fairness assessments of the failure resolution. However, such research could help better understand what it needs to provide recoveries that sustain customer loyalty. In doing so, researchers could develop analytical models that trade off efforts and inform the design of optimal recovery strategies and companies could better determine the cost of keeping a customer happy after failure and incorporate this in individual CLV assessments.

*Perceived justice.* Within the thesis, it was investigated how the frequently employed justice framework can help explain postfailure loyalty in terms of purchase behavior *(Project II)*. The findings suggest that only interactional justice translates into postfailure loyalty. No significant effects emerged for procedural justice and interactional justice.

This may be because the retailing context is in general particularly sensitive to interactional justice. Apparently, elements of personal interaction are of greater relevance than processes or compensation in complaint- or failure-related interactions. However, this result was obtained by assessing main effects only. Moderating factors may alter the effects of justice dimensions and should be considered by future research. For example, prior work shows that the failure type (process vs. outcome) modulates the effects of individual justice dimensions (Smith, Bolton, and Wagner 1999); thus, failure characteristics should be examined as moderators of the perceived justice–behavioral loyalty link. In addition, further research could examine whether the finding holds true in other contexts as well or whether other justice dimensions become paramount.

In general, prior research has long acknowledged the importance of employee-customer interactions (e.g., Brown and Lam 2008; Crosby, Evans, and Cowles 1990; Harris, Baron, and Parker 2000; Hartline and Ferrell 1996; Solomon et al. 1985) and in particular identified the crucial role of social interactions in retailing (De Wulf, Odekerken-Schröder, and Iacobucci 2001). However, there is no evidence of how high-quality interactions can affect actual loyalty behaviors and financial outcomes. Thus, this thesis extends current knowledge by demonstrating that high interactional justice can have a substantial, positive monetary effect. This is a critical finding considering that companies increasingly replace personal interaction by IT- and software-enabled communication systems. Such strategies can be risky because firms would give up on a critical success factor when introducing automated communication and self-service offerings on a large scale. It should be a priority for future research to investigate potential negative consequences that may result from a lack of personal interaction. More specifically, research efforts could analyze whether the benefits of such strategies outweigh their cost and whether there are ways to compensate for the absence of personal and emotional clues in interactions.

*The prefailure relationship state.* Overall, all three projects of the thesis follow a longitudinal approach and study postfailure loyalty while controlling for the prefailure relationship state. Throughout the thesis I specify and test dynamic models incorporating pre- and postfailure relationship perceptions and longitudinal purchase data. Thereby, it was possible to account for the prefailure relationship state on a behavioral and attitudinal level when analyzing postfailure loyalty outcomes and their determinants. *Project I* revealed how a performance failure caused a change in transaction behavior and attitudes by comparing the pre- and postfailure levels of these outcomes. In *Project II* it was shown that carryover effects are present for prefailure satisfaction and purchase spending to postfailure satisfaction and purchase spending. In *Project III* it was shown that the prefailure relationship state matters in terms of how affectively committed customers were before the failure. In view of these results, it is surprising that most studies previously conducted in the research domain do not account for that. The majority of prior work is cross-sectional, some studies retrospectively interrogate prior experiences, and only few authors try to account for prefailure characteristics and carryover effects using a longitudinal research design. In summary, the findings suggest that the prefailure relationship state is an important determinant of postfailure outcomes; thus, future research could try to conduct more longitudinal studies for the benefit of an unbiased assessment of prefailure relationship health. As such, this can help better understand the dynamics in a scenario of failure and recovery and contribute to knowledge of how optimal postfailure outcomes can be achieved.

*Relationship characteristics.* Two schools of thought exist regarding the role of positive prior relationship experience with a provider. For failure situations, many researchers have found evidence for a "protection" effect that buffers negative consequences resulting from performance failures. Conversely, growing evidence suggests that strong relationships amplify negative responses. Considering the unfavorable findings regarding the adverse effect of prefailure affective commitment, I add to the second position and extend prior works by showing that this effect translates into actual postfailure purchase behavior. This extends current knowledge by providing behavioral evidence that there is a "dark side" of strong customer relationships (Anderson and Jap 2005). I proposed that customer relationships can be characterized and dominated by emotional and economic bonds and that, depending on the intensity of one or the other aspect, an amplifying or buffering effect of negative failure consequences would emerge. By examining relationship and affective commitment as relationship characteristics, I tried to accommodate these two distinct aspects accordingly. However, as the results suggest, no support emerges for the economic mechanism theorized, suggesting that customers high in prefailure relationship commitment are not more forgiving and that relationship commitment does not buffer negative consequences of performance failures in terms of behavioral loyalty. In summary, the findings demonstrate that emotional attachment and prefailure relationship quality matters: regardless of the failure resolution, customers with high prefailure affective commitment will retaliate and reduce repatronage behavior. Other research, that studies similar mechanisms, suggests that such an amplifying effect may occur in cases of severe transgression (Evanschitzky, Brock, and Blut 2011; Ganesan et al. 2010). Since the failures studied in this thesis were classified to be severe, the results are in line with this general rationale; however, much work needs to be done in future research to explain conflicting results (see Table 2.1) and disentangle the decisive factors that make proactive strategies amplify or buffer negative failure consequences.

*Marketplace characteristics.* The thesis results suggest a strong influence of marketplace characteristics and report that high swichting costs and locational convenience can enhance recovery effectiveness and buffer negative failure consequences in terms of postfailure loyalty. Conversely, under low switching costs and locational convenience negative effects may be more severe, if the failure resolution is not successful or delayed. This extends current knowledge by revealing that recovery effectiveness is contingent on external factors which are not fully under company control. The findings enhance the understanding of the boundary conditions of failure and recovery, and reveal a more holistic picture of postfailure processes. With regard to further research, it could be studied how companies can artificially create switching barriers that are capable of buffering negative consequences of performance failure. For example, loyalty programs are considered a proactive strategy that aims at increasing perceptions of switching costs (e.g., Wirtz, Mattila, and Lwin 2007); thus, it could be studied whether loyalty programs prove beneficial in enhancing recovery effectiveness with regard to postfailure purchase behavior. Similarly, further moderating factors in the failure/recovery process could be investigated. In this regard, a challenge would be to find a consensus on which set of moderators affects which link in the causal chain of the general research framework. Undoubtedly, some moderators may influence more than one of the central relationships, which makes it even more challenging and complex for research to identify critical configurations of determinants that yield positive postfailure outcomes.

Overall, the results of the thesis have some general research implications. It is remarkable that emotion-related constructs—that is affective commitment and interactional justice—have a strong effect on actual purchase behavior. In *Project II*, the direct effect of interactional justice on purchase behavior appears even slightly larger than the effect of satisfaction. Apparently, emotions play a critical role in the field of failure, recovery, and complaint management. Extant research has provided little evidence for emotions to be related to postfailure loyalty/exit-behavior (Chebat and Slusarczyk 2005). Thus, research should consider emotional factors more frequently in its models and study more explicitly which links are affected within the general research framework.

In general, marketing research has embraced satisfaction as one of its most important outcome measures that is frequently investigated as a precursor to customer loyalty and "usually regarded as the central mediator of postpurchase behavior" (Tax, Brown, and Chandrashekaran 1998, p. 641). However, although research has established the link between satisfaction and actual purchase behavior, the process is usually not straightforward and typically the predictive power of satisfaction is relatively weak. In view of this dissertation's findings, it can be a fruitful avenue for further research to explore how emotion-related constructs can serve to predict actual customer behavior and contribute to explain variance in dependent variables beyond the capability of satisfaction. In general, marketing decision models could incorporate more emotional predictor variables and not solely rely on rational information like price, perceived value or utility. In the last decade, behavioral economists increasingly provided evidence that human behavior and decision-making is more strongly dominated by emotions than researchers have long assumed (Rick and Loewenstein 2008). Therefore, emotional predictor variables should particularly be considered in emotion-loaded situations like failure episodes and could also be more frequently represented in other fields of behavioral marketing research.

In addition, the findings of the thesis add to the rationale of Seiders et al. (2005), who show that marketplace characteristics are an important determinant for actual customer loyalty. The results provide theoretical insights and can help explain why the mere satisfaction–purchase behavior link was frequently found nonsignificant by researchers. A basic assumption of most previous studies in the research domain of performance failures is that actual behavior stems from customers' attitudes and intentions. However, whether behavioral intentions indeed translate into actual behavior may depend on additional factors not always under company control. Marketplace characteristics, such as competitor attractiveness, switching costs, and locational convenience, may take effect after the formation of consumer intent yet right before the execution of actual behaviors. These factors may prevent customers from behaving in accordance with their previously developed attitudes and thus are important to account for in models linking perceptual data to objective purchase behavior. A large majority of marketing studies neglects such external factors and does not consider them in their models. If accounting for competition-related factors is not feasible, it might help to employ measures that capture satisfaction or similar constructs relative to how competitive offerings would be evaluated by consumers when predicting customer loyalty. After all, future research could more frequently employ objectively measured behavioral data when studying customer loyalty.

### 7.2.2 Implications for Management

In summary, the results of this dissertation suggest that companies should pay particular attention to providing high interactional justice, a quick recovery solution, and should account for relationship and marketplace characteristics. From this, it seems commendable to pursue complaint management with a twofold approach: (1) the execution of a *baseline strategy* that applies at all times, and (2) a set of *enhanced strategies*, which follow a *value-based approach* and an *adaptive approach* that takes into account situational contingencies.

*Baseline strategy.* The thesis results suggest that customers who experienced a performance failure are particularly sensitive to emotional clues. Interactional justice and affective commitment have a strong effect on postfailure loyalty which underscores the important role of interpersonal and emotional aspects in complaint interactions and has important implications for managerial practice. Companies could consider it a paramount goal of complaint handling to ensure high interactional justice at all times and in any case and could implement a baseline strategy that aims to reach this objective. Pursuing interactional justice as a primary strategic goal can make sense because of three reasons: First, the thesis results suggest a nonsignficant effect of distributive justice but a salient role of interactional justice with a substantial impact on postfailure loyalty. Anecdotal evidence from practitioners supports this finding and contends that during personal interactions "mistreatment and incompetence often result in five times more damage to loyalty than do monetary concerns" (Goodman 2006, p. 29). Furthermore, research evidence suggests that generously compensated customers can still be dissatisfied when service personnel is impolite (Blodgett, Hill, and Tax 1997). Thus, great monetary or material amendments may be wasted when they are not delivered in a friendly and apologizing manner. Second, excellently trained customer service agents may pay off in the long-run and on multiple levels because 92% of customers form their attitudes about a company based on the quality of call center interactions that they experienced (Aksin, Armony, and Mehrotra 2007). Furthermore, McKinsey recently reported that “by focusing more thoughtfully on the human side of customer service, ... companies are lowering costs by 10 percent or more while improving customer satisfaction scores by up to 30 percent” (DeVine, Lal, and Zea 2012, p. 2). Interactional justice can be provided with “no-cost actions” (Smith, Bolton, and Wagner 1999, p. 369) and investments in personnel training are relatively cost effective, in particular because a highly skilled customer-contact staff is beneficial not only for complaint management. Third, even when the company fails to fully deliver the required outcomes for the customer to be completely satisfied, an empathic employee behavior combined with sincere apologies and explanations should be capable of reducing a complainant’s anger to the point that it prevents strong negative word of mouth or other retaliatory behaviors.

*Enhanced strategies.* In addition to these basic efforts aiming at establishing high interactional justice, managers could pursue a set of more enhanced strategies that account for situational factors. That is, companies could follow an *adaptive approach* to complaint handling which meets the specific demands of a given failure situation. As such, the adaptive approach takes on a differentiated view and helps determine adequate actions that are likely to yield positive outcomes for both the customer and the company.

For this purpose, it is inevitable for firms to use technology to effectively execute such enhanced recovery strategies. In recent years, “consumers are gaining more power than firms because they are quicker to adopt disruptive technologies" (Hagel, Brown, and Kulasooriya 2011, p. 13) but at the same time "the customer service landscape is changing as social media and new mobilephone technologies give companies unprecedented access to data on customer interactions" (DeVine, Lal, and Zea 2012, p. 2). Thus, companies should take advantage of the increased information availability and use customer knowledge as well as IT- and software-tools to keep up with consumers and to provide customers a convincing, high-quality service that sustains their loyalty. Firms who use technology can enhance their success with failure recoveries and complaint handling in several ways: First, as the thesis results demonstrate, it is crucial to provide a quick failure resolution to maintain customer loyalty. If company representatives have extensive customer knowledge and information on products, services, and prior company actions readily available in their databases, they will be able to respond to complaints immediately and resolve failures at increased speed. Second, customer knowledge can help to identify certain customer types. A segmentation of complainants according to their expected recovery responsiveness and postfailure loyalty could be established. This thesis has identified a number of contingency factors (e.g., relationship and marketplace characteristics) which could be incorporated as segmentation variables. Ultimately, such an approach can help determine the likelihood of whether a positive recovery outcome can be achieved in terms of behavioral loyalty. On an ongoing basis, different configurations of such contingency factors should be recognized, refined and verified. Then, specific recovery actions could be tailored to fit these situational factors and, after testing these strategies, success rates and learnings should be documented and stored in the database for the benefit of knowledge building and better management of future complaints. Third, on this basis, managers could also implement decision-support systems that determine the propensity for the customer to stay loyal after failure resolution. Similarly to churn scores, probability scores could be provided that incorporate effects of outcome determinants and indicate how likely it is that a complaint case can be successfully resolved and keep a customer loyal. Thereby, frontline employees can be informed whether to initiate the baseline strategy or provide increased recovery efforts.

The adaptive approach can be complemented by a *value-based approach* that helps determine the resources that should be allocated to resolution of a particular failure. Overall, the thesis shows how a performance failure can affect CLV and CE in terms of purchase behavior. Considering these effects can support decision-making and the determination of adequate investments in recovery, such as compensation or special engagements that surprise or delight customers to make it up to them. “Considering how much it costs to lose a customer, few recovery efforts are too extreme” (Hart, Heskett, and Sasser 1990, p. 151). Complaint managers could make use of database information, develop and implement customer costing (Kaplan 2012) and establish CLV projections to guide resource allocation and make profitable investment decisions. Moreover, the value-based approach to complaint handling can help face the newly acquired consumer power in the marketplace, which makes some consumers prone to opportunistic behaviors such as demanding unreasonable high compensation after failures (Wirtz and McColl-Kennedy 2010; Wirtz 2011). An individual CLV assessment helps better evaluate such claims and decide on the extent of an appropriate redress. In addition, the value-based approach can be particularly valuable when applied in conjunction with the adaptive approach because it could inform the design of optimal recoveries and lead to more effective resource allocation and increased productivity of the customer service function. For example, in some complaint cases it might be very unlikely that customers will not switch despite successful failure resolution. In such a situation, it could be considered to withhold extra resources and investments in recovery and stick to the baseline strategy. After all, “a good understanding of how ... complaining customers are treated is not only an ethical question, it is also a matter of profitable management” (Chebat and Slusarczyk 2005, p. 664).

The finding that customers who frequently complain can be up to six times more profitable than noncomplaining customers (Bain & Company 2012) illustrates the potential for complaint management to become a highly profitable company activity. Once companies start to measure gains and losses that result from failure, recovery, and altered postfailure purchase behavior, recovery activities become more accountable and the ROC can be calculated. The findings revealed by this thesis provide insights into the monetary consequences of performance failure. Such a quantification of financial effects is not only a starting point to but also “the only way of getting top management’s attention” (Hart, Heskett, and Sasser 1990, p. 150).

## 7.3 Conclusion and Outlook

Complaint management has become more critical than ever. A fundamental shift in consumer markets is taking place that—driven by technological developments—leads to greater consumer power and increased brand disloyalty. High market transparency, reduced information asymmetries, and low switching costs enable aggrieved customers to exert more pressure on transgressing providers by imposing sanctions via exit and voice. The substantial repercussions performance failures can cause today make the management of failures and complaints a crucial task that demands serious managerial attention.

The findings of this dissertation contribute to these new challenges in three ways: First, the result that interactional justice has a strong effect on repatronage indicates that relatively simple and cost-effective reactive strategies, such as courteous and polite employee behavior, can yield substantial positive financial outcomes. In view of the general trend that technology increasingly replaces personal customer-firm interaction, this finding indicates that complaint-related interactions should be cautiously designed and ideally comprise person-to-person communication. Second, as a proactive strategy, the establishment of strong customer relationships may not always be effective in protecting firms from negative failure consequences. On the contrary, strong emotional attachment of the customer may backfire in the case of a severe performance failure. Consequently, research and management should carefully reexamine the benefits of relationship marketing in the modern marketplace. Third, in line with this, marketplace characteristics were identified as important determinants of postfailure customer loyalty. Low switching barriers can lead to customer churn despite high levels of satisfaction with failure resolution, and in turn—in the case of high switching barriers—customers may stay with a supplier even though they had dissatisfying experiences. Although not fully under company control, managers should try to account for these contingency factors when prioritizing the allocation of limited resources. Finally, an assessment of postfailure purchase behavior can help reveal the return on investment of complaint handling. By understanding the link between the various outcome determinants, loyalty, and profits, managers can obtain guidance for adequate resource allocation and thereby increase the productivity of customer service functions.

Although this thesis suggests an adaptive and value-based approach to complaint handling, firms should avoid implementing recovery strategies solely on the grounds of projected financial return. As discussed, performance failures can also negatively affect attitudinal outcomes and lead to more indirect detrimental consequences. The financial impact of negative word of mouth and online publicity, though difficult to quantify, can be severe. Thus, to avoid a negative company reputation and remain competitive in the modern marketplace, firms should ensure fair policies and serve their customers with sincere intentions and behaviors to the best of their ability and available resources.

# References

Aaker, J., S. Fournier, and S. A. Brasel (2004), "When Good Brands Do Bad," *Journal of Consumer Research*, 31 (1), 1–16.

Adams, J. S. (1965), "Inequity in Social Exchange," in *Advances in Experimental Social Psychology*, L. Berkowitz, ed., New York, NY: Academic Press, 267–99.

Aiken, L. S. and S. G. West (1991), *Multiple Regression: Testing and Interpreting Interactions*, London: Sage Publications.

Ajzen, I. (2001), "Nature and Operation of Attitudes," *Annual Review of Psychology*, 52 (1), 27–58.

Aksin, Z., M. Armony, and V. Mehrotra (2007), "The Modern Call Center: A Multi-Disciplinary Perspective on Operations Management Research," *Production & Operations Management*, 16 (6), 665–88.

Ambrose, M., R. L. Hess, and S. Ganesan (2007), "The Relationship between Justice and Attitudes: An Examination of Justice Effects on Event and System-Related Attitudes," *Organizational Behavior and Human Decision Processes*, 103 (1), 21–36.

Anderson, E. W. and M. W. Sullivan (1993), "The Antecedents and Consequences of Customer Satisfaction for Firms," *Marketing Science*, 12 (2), 125–43.

Anderson, E. and S. D. Jap (2005), "The Dark Side of Close Relationships," *MIT Sloan Management Review*, 46 (3), 75–82.

Anderson, E. and B. Weitz (1992), "The Use of Pledges to Build and Sustain Commitment in Distribution Channels," *Journal of Marketing Research*, 29 (1), 18–34.

Andreassen, T. W. (2001), "From Disgust to Delight: Do Customers Hold a Grudge?" *Journal of Service Research*, 4 (1), 39–49.

Angrist, J. D. and J.-S. Pischke (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, NJ: Princeton University Press.

Armstrong, J. and T. Overton (1977), "Estimating Nonresponse Bias in Mail Surveys," *Journal of Marketing Research*, 14 (3), 396–402.

Ashenfelter, O. and D. Card (1985), "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs," *The Review of Economics and Statistics*, 67 (4), 648–60.

Babakus, E., C. Bienstock, and J. Van Scotter (2004), "Linking Perceived Quality and Customer Satisfaction to Store Traffic and Revenue Growth," *Decision Sciences*, 35 (4), 713–37.

Bagozzi, R. P. and Y. Yi (1988), "On the Evaluation of Structural Equation Models," *Journal of the Academy of Marketing Science*, 16 (1), 74–94.

Bagozzi, R. (1984), "Expectancy-Value Attitude Models an Analysis of Critical Measurement Issues," *International Journal of Research in Marketing*, 1 (4), 295–310.

Bai, H. (2011), "Using Propensity Score Analysis for Making Causal Claims in Research Articles," *Educational Psychology Review*, 23 (2), 273–78.

Bain & Company (2012), "Converting Complainers into More Profitable Customers," Available from: `http://www.bain.com/about/client-results/converting-complainers-into-more-profitable-customers.aspx` [accessed April 30, 2012].

Baron, R. M. and D. A. Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (6), 1173–82.

Bell, D. R., T.-H. Ho, and C. S. Tang (1998), "Determining Where to Shop: Fixed and Variable Costs of Shopping," *Journal of Marketing Research*, 35 (3), 352–69.

Bergkvist, L. and J. R. Rossiter (2007), "The Predictive Validity of Multiple-Item Versus Single-Item Measures of the Same Constructs," *Journal of Marketing Research*, 44 (2), 175–84.

Berry, L. L. (1983), *Relationship Marketing*, Chicago, IL: American Marketing Association.

——— (2002), "Relationship Marketing of Services Perspectives from 1983 and 2000," *Journal of Relationship Marketing*, 1 (1), 59–77.

Bettencourt, L. A. (1997), "Customer Voluntary Performance: Customers as Partners in Service Delivery," *Journal of Retailing*, 73 (3), 383–406.

Bitner, M. J., B. H. Booms, and M. S. Tetreault (1990), "The Service Encounter: Diagnosing Favorable and Unfavorable Incidents," *Journal of Marketing*, 54 (1), 71–84.

Bitner, M., B. Booms, and L. Mohr (1994), "Critical Service Encounters: The Employee's Viewpoint," *Journal of Marketing*, 58 (4), 95–106.

Blodgett, J. G., D. H. Granbois, and R. G. Walters (1993), “The Effects of Perceived Justice on Complainants’ Negative Word-of-Mouth Behavior and Repatronage Intentions,” *Journal of Retailing*, 69 (4), 399–428.

Blodgett, J. G., D. J. Hill, and S. S. Tax (1997), “The Effects of Distributive, Procedural, and Interactional Justice on Postcomplaint Behavior,” *Journal of Retailing*, 73 (2), 185–210.

Boehm, M. (2008), “Determining the Impact of Internet Channel Use on a Customer’s Lifetime,” *Journal of Interactive Marketing*, 22 (3), 2–22.

Bolton, R. N. (1998), “A Dynamic Model of the Duration of the Customer’s Relationship with a Continuous Service Provider: The Role of Satisfaction,” *Marketing Science*, 17 (1), 45–65.

Bolton, R. N., R. K. Kannan, and M. D. Bramlett (2000), “Implications of Loyalty Program Membership and Service Experiences for Customer Retention and Value,” *Journal of the Academy of Marketing Science*, 28 (1), 95–108.

Bolton, R. N., K. N. Lemon, and P. C. Verhoef (2004), “The Theoretical Underpinnings of Customer Asset Management: A Framework and Propositions for Future Research,” *Journal of the Academy of Marketing Science*, 32 (3), 271–92.

Boshoff, C. (1997), “An Experimental Study of Service Recovery Options,” *International Journal of Service Industry Management*, 19 (1), 34–51.

Bowen, D. E., S. W. Gilliland, and R. Folger (1999), “HRM and Service Fairness: How Being Fair with Employees Spills over to Customers,” *Organizational Dynamics*, 27 (3), 7–23.

Brady, M. K., J. J. Cronin Jr, G. L. Fox, and M. L. Roehm (2008), "Strategies to Offset Performance Failures: The Role of Brand Equity," *Journal of Retailing*, 84 (2), 151–64.

Brockner, J., T. R. Tyler, and R. Cooper-Schneider (1992), "The Influence of Prior Commitment to an Institution on Reactions to Perceived Unfairness: The Higher They Are, the Harder They Fall," *Administrative Science Quarterly*, 37 (2), 241–61.

Bronnenberg, B. J., J.-P. Dub, and C. F. Mela (2010), "Do Digital Video Recorders Influence Sales?" *Journal of Marketing Research*, 47 (6), 998–1010.

Brooks, C. M., P. J. Kaufmann, and D. R. Lichtenstein (2008), "Trip Chaining Behavior in Multi-Destination Shopping Trips: A Field Experiment and Laboratory Replication," *Journal of Retailing*, 84 (1), 29–38.

Brown, S. W. (2000), "Practicing Best-in-Class Service Recovery," *Marketing Management*, 9 (2), 8–9.

Brown, S. P. and S. K. Lam (2008), "A Meta-Analysis of Relationships Linking Employee Satisfaction to Customer Responses," *Journal of Retailing*, 84 (3), 243–55.

Browne, M. W. and R. Cudeck (1992), "Alternative Ways of Assessing Model Fit," *Sociological Methods & Research*, 21 (2), 230–58.

Burnham, T. A., J. K. Frels, and V. Mahajan (2003), "Consumer Switching Costs: A Typology, Antecedents, and Consequences," *Journal of the Academy of Marketing Science*, 31 (2), 109–26.

Caliendo, M. and S. Kopeinig (2008), "Some Practical Guidance for the Implementation of Propensity Score Matching," *Journal of Economic Surveys*, 22 (1), 31–72.

Campbell, D. and F. Frei (2010), "Cost Structure, Customer Profitability, and Retention Implications of Self-Service Distribution Channels: Evidence from Customer Behavior in an Online Banking Channel," *Management Science*, 56 (1), 4–24.

Carlzon, J. (1989), *Moments of Truth*, Cambridge, MA: Harper Business.

Chandon, P., V. G. Morwitz, and W. J. Reinartz (2005), "Do Intentions Really Predict Behavior? Self-Generated Validity Effects in Survey Research," *Journal of Marketing*, 69 (2), 1–14.

Chebat, J.-C., M. Davidow, and A. Borges (2011), "More on the Role of Switching Costs in Service Markets: A Research Note," *Journal of Business Research*, 64 (8), 823–29.

Chebat, J.-C. and W. Slusarczyk (2005), "How Emotions Mediate the Effects of Perceived Justice on Loyalty in Service Recovery Situations: An Empirical Study," *Journal of Business Research*, 58 (5), 664–73.

Churchill, G. A. (1979), "A Paradigm for Developing Better Measures of Marketing Constructs." *Journal of Marketing Research*, 16 (1), 64–73.

Clemmer, E. and B. Schneider (1996), "Fair Service," in *Advances in Services Marketing and Management*, T. A. Swartz, D. E. Bowen, and S. W. Brown, eds., Greenwich, CT: JAI Press, 13–229.

Cochran, W. G. (1968), "The Effectiveness of Adjustment by Subclassification in Removing Bias in Observational Studies," *Biometrics*, 24 (2), 295–313.

Cochran, W. G. and D. B. Rubin (1973), "Controlling Bias in Observational Studies: A Review," *Sankhyā: The Indian Journal of Statistics*, 35 (4), 417–46.

Cohen, J., P. Cohen, S. G. West, and L. S. Aiken (2003), *Applied Multiple Regression/Correlation Analysis for Behavioral Sciences*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Colquitt, J. A., D. E. Conlon, M. J. Wesson, C. O. L. H. Porter, and K. Y. Ng (2001), "Justice at the Millennium: A Meta-Analytic Review of 25 Years of Organizational Justice Research," *Journal of Applied Psychology*, 86 (3), 425–45.

Craig, C. S., A. Ghosh, and S. McLafferty (1984), "Models of the Retail Location Process: A Review," *Journal of Retailing*, 60 (1), 5–36.

Crosby, L. A., K. A. Evans, and D. Cowles (1990), "Relationship Quality in Services Selling: An Interpersonal Influence Perspective," *Journal of Marketing*, 54 (3), 68–81.

Davidow, M. (2000), "The Bottom Line Impact of Organizational Responses to Customer Complaints," *Journal of Hospitality & Tourism Research*, 24 (4), 473–90.

——— (2003a), "Have You Heard the Word? The Effect of Word of Mouth on Perceived Justice, Satisfaction and Repurchase Intentions Following Complaint Handling," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 16, 67–81.

——— (2003b), "Organizational Responses to Customer Complaints: What Works and What Doesn't," *Journal of Service Research*, 5 (3), 225–50.

Day, R. L., K. Grabicke, T. Schaetzle, and F. Staubach (1981), "The Hidden Agenda of Consumer Complaining," *Journal of Retailing*, 57 (3), 86–106.

De Cannière, M. H., P. De Pelsmacker, and M. Geuens (2009), "Relationship Quality and the Theory of Planned Behavior Models of Behavioral Intentions and Purchase Behavior," *Journal of Business Research*, 62 (1), 82–92.

De Wulf, K., G. Odekerken-Schröder, and D. Iacobucci (2001), "Investments in Consumer Relationships: A Cross-Country and Cross-Industry Exploration," *Journal of Marketing*, 65 (4), 33–50.

DeVine, J., S. Lal, and M. Zea (2012), "The Human Factor in Service Design," *McKinsey Quarterly*, 2012 (1), 118–23.

DeWitt, T., D. Nguyen, and R. Marshall (2008), "Exploring Customer Loyalty Following Service Recovery," *Journal of Service Research*, 10 (3), 269–81.

DeWitt, T. and M. K. Brady (2003), "Rethinking Service Recovery Strategies," *Journal of Service Research*, 6 (2), 193–207.

van Doorn, J. (2011), "Comment: Customer Engagement," *Journal of Service Research*, 14 (3), 280–82.

van Doorn, J. and P. C. Verhoef (2008), "Critical Incidents and the Impact of Satisfaction on Customer Share," *Journal of Marketing*, 72 (4), 123–42.

Dwyer, F. R., P. H. Schurr, and S. Oh (1987), "Developing Buyer-Seller Relationships," *Journal of Marketing*, 51 (2), 11–27.

Estelami, H. (2000), "Competitive and Procedural Determinants of Delight and Disappointment in Consumer Complaint Outcomes," *Journal of Service Research*, 2 (3), 285–300.

Evanschitzky, H., C. Brock, and M. Blut (2011), "Will You Tolerate This? The Impact of

Affective Commitment on Complaint Intention and Postrecovery Behavior," *Journal of Service Research*, 14 (4), 410–25.

Farrell, J. and P. Klemperer (2007), "Coordination and Lock-in: Competition with Switching Costs and Network Effects," in *Handbook of Industrial Organization*, M. Armstrong and R. Porter, eds., Amsterdam: Elsevier, 1967–2072.

Festinger, L. (1957), *A Theory of Cognitive Dissonance*, Palo Alto, CA: Stanford University Press.

Forbes, L. P., S. W. Kelley, and K. D. Hoffmann (2006), "Typologies of E-Commerce Retail Failures and Recovery Strategies," *Journal of Services Marketing*, 19 (5), 220–92.

Fornell, C. (1992), "A National Customer Satisfaction Barometer: The Swedish Experience," *Journal of Marketing*, 56 (1), 6–21.

Fornell, C. and D. F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (1), 39–50.

Fornell, C. and B. Wernerfelt (1987), "Defensive Marketing Strategy by Customer Complaint Management: A Theoretical Analysis," *Journal of Marketing Research*, 24 (4), 337–46.

——— (1988), "A Model for Complaint Management," *Marketing Science*, 7 (3), 287–98.

Fox, E., S. Postrel, and A. McLaughlin (2007), "The Impact of Retail Location on

Retailer Revenues: An Empirical Investigation," working paper, Cox School of Business, Southern Methodist University.

Francken, D. A. (1983), "Postpurchase Consumer Evaluations, Complaint Actions and Repurchase Behavior," *Journal of Economic Psychology*, 4 (3), 273–90.

Fullerton, G. (2003), "When Does Commitment Lead to Loyalty?" *Journal of Service Research*, 5 (4), 333–44.

Funches, V., M. Markley, and L. Davis (2009), "Reprisal, Retribution and Requital: Investigating Customer Retaliation," *Journal of Business Research*, 62 (2), 231–38.

Ganesan, S., S. P. Brown, B. J. Mariadoss, and H. Ho (2010), "Buffering and Amplifying Effects of Relationship Commitment in Business-to-Business Relationships," *Journal of Marketing Research*, 47 (2), 361–73.

Gelbrich, K. and H. Roschk (2011), "A Meta-Analysis of Organizational Complaint Handling and Customer Responses," *Journal of Service Research*, 14 (1), 24–43.

Gensler, S., P. Leeflang, and B. Skiera (2012), "Impact of Online Channel Use on Customer Revenues and Costs to Serve: Considering Product Portfolios and Self-Selection," *International Journal of Research in Marketing*, (forthcoming).

Gerbing, D. W. and J. C. Anderson (1988), "An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment," *Journal of Marketing Research*, 25 (2), 186–92.

Gilly, M. C. (1987), "Postcomplaint Processes: From Organizational Response to Repurchase Behavior," *Journal of Consumer Affairs*, 21 (2), 293–313.

Gilly, M. C. and B. D. Gelb (1982), "Post-Purchase Consumer Processes and the Complaining Consumer," *Journal of Consumer Research*, 9 (3), 323–28.

Goodman, J. (2006), "Manage Complaints to Enhance Loyalty," *Quality Progress*, 30 (2), 28–34.

Goodwin, C. and I. Ross (1989), "Salient Dimensions of Perceived Fairness in Resolution of Service Complaints," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 2, 87–92.

——— (1992), "Consumer Responses to Service Failures: Influence of Procedural and Interactional Fairness Perceptions," *Journal of Business Research*, 25 (2), 149–63.

Grégoire, Y. and R. Fisher (2006), "The Effects of Relationship Quality on Customer Retaliation," *Marketing Letters*, 17 (1), 31–46.

——— (2008), "Customer Betrayal and Retaliation: When Your Best Customers Become Your Worst Enemies," *Journal of the Academy of Marketing Science*, 36 (2), 247–61.

Grégoire, Y., T. M. Tripp, and R. Legoux (2009), "When Customer Love Turns into Lasting Hate: The Effects of Relationship Strength and Time on Customer Revenge and Avoidance," *Journal of Marketing*, 73 (6), 18–32.

Grewal, D., M. Levy, and V. Kumar (2009), "Customer Experience Management in Retailing: An Organizing Framework," *Journal of Retailing*, 85 (1), 1–14.

Grönroos, C. (1988), "Service Quality: The Six Criteria of Good Perceived Service Quality," *Review of Business*, 9 (3), 10–3.

——— (1994), "From Marketing Mix to Relationship Marketing: Towards a Paradigm Shift in Marketing," *Management Decision*, 32 (2), 4–20.

Gummesson, E. (1997), "Relationship Marketing as a Paradigm Shift: Some Conclusions from the 30R Approach," *Management Decision*, 35 (4), 267–72.

Gundlach, G. T., R. S. Achrol, and J. T. Mentzer (1995), "The Structure of Commitment in Exchange," *Journal of Marketing*, 59 (1), 78–92.

Gupta, S., D. R. Lehmann, and J. A. Stuart (2004), "Valuing Customers," *Journal of Marketing Research*, 41 (1), 7–18.

Gursoy, D., E. H. Ekiz, and C. G. Chi (2007), "Impacts of Organizational Responses on Complainants' Justice Perceptions and Post-Purchase Behaviors," *Journal of Quality Assurance in Hospitality & Tourism*, 8 (1), 1 – 25.

Gustafsson, A., M. D. Johnson, and I. Roos (2005), "The Effects of Customer Satisfaction, Relationship Commitment Dimensions, and Triggers on Customer Retention," *Journal of Marketing*, 69 (4), 210–18.

Hagel, J., J. S. Brown, and D. Kulasooriya (2011), "The 2011 Shift Index: Measuring the Forces of Long-Term Change," Available from: `http://www.deloitte.com/assets/Dcom-UnitedStates/Local%20Assets/Documents/us_tmt_2011shiftindex_111011.pdf` [accessed April 30, 2012].

Hair, J. F., R. E. Anderson, R. L. Tatham, and W. C. Black (1998), *Multivariate Data Analysis*, Upper Saddle River, NJ: Prentice Hall.

Harris, K., S. Baron, and C. Parker (2000), “Understanding the Consumer Experience : It’s ’Good To Talk’,” *Journal of Marketing Management*, 16 (1-3), 111–27.

Hart, C. W. L., J. L. Heskett, and W. E. Sasser (1990), “The Profitable Art of Service Recovery,” *Harvard Business Review*, 68 (4), 148–56.

Hartline, M. D. and O. C. Ferrell (1996), “The Management of Customer-Contact Service Employees: An Empirical Investigation,” *Journal of Marketing*, 60 (4), 52–70.

Harvard Business Review (2011), *Harvard Business Review on Increasing Customer Loyalty*, Boston, MA: Harvard Business Press.

Heckman, J., H. Ichimura, J. Smith, and P. Todd (1998), “Characterizing Selection Bias Using Experimental Data,” *Econometrica*, 66 (5), 1017–98.

Heckman, J. J. (1976), “The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models,” *Annals of Economic and Social Measurement*, 5 (4), 475–92.

——— (1979), “Sample Selection Bias as a Specification Error,” *Econometrica*, 47 (1), 153–61.

Heckman, J. J., H. Ichimura, and P. E. Todd (1997), “Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme,” *Review of Economic Studies*, 64 (221), 605–54.

Helberg, C. (1996), “Pitfalls of Data Analysis,” *Practical Assessment, Research and Evaluation*, 5 (5), 12–7.

Hennig-Thurau, T., K. P. Gwinner, and D. D. Gremler (2002), “Understanding Relationship Marketing Outcomes: An Integration of Relational Benefits and Relationship Quality," *Journal of Service Research*, 4 (3), 230–47.

Heskett, J. L., W. E. Sasser, and L. A. Schlesinger (1997), *The Service Profit Chain: How Leading Companies Link Profit to Loyalty, Satisfaction, and Value*, New York: The Free Press.

Hess, R. L., S. Ganesan, and N. M. Klein (2003), "Service Failure and Recovery: The Impact of Relationship Factors on Customer Satisfaction," *Journal of the Academy of Marketing Science*, 31 (2), 127–45.

——— (2007), "Interactional Service Failures in a Pseudorelationship: The Role of Organizational Attributions," *Journal of Retailing*, 83 (1), 79–95.

Hirschman, A. O. (1970), *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*, Cambridge, MA: Harvard University Press.

Hocutt, M. A., M. R. Bowers, and T. D. Donavan (2006), "The Art of Service Recovery: Fact or Fiction?" *Journal of Services Marketing*, 20 (3), 199–207.

Hocutt, M. A. and G. Chakraborty (1997), "The Impact of Perceived Justice on Customer Satisfaction and Intention to Complain in a Service Recovery," *Advances in Consumer Research*, 24 (1), 457–63.

Hoffman, K., S. Kelley, and H. Rotalsky (1995), "Tracking Service Failures and Employee Recovery Efforts," *Journal of Services Marketing*, 9 (2), 49–61.

Hogan, J., D. Lehmann, M. Merino, R. Srivastava, J. Thomas, and P. Verhoef (2002), "Linking Customer Assets to Financial Performance," *Journal of Service Research*, 5 (1), 26–38.

Hogan, J., K. Lemon, and R. Rust (2002), “Customer Equity Management,” *Journal of Service Research*, 5 (1), 4–12.

Holloway, B. B. and S. E. Beatty (2003), “Service Failure in Online Retailing,” *Journal of Service Research*, 6 (1), 92–105.

Homburg, C. and A. Fürst (2005), “How Organizational Complaint Handling Drives Customer Loyalty: An Analysis of the Mechanistic and the Organic Approach,” *Journal of Marketing*, 69 (3), 95–114.

——— (2007), “See No Evil, Hear No Evil, Speak No Evil: A Study of Defensive Organizational Behavior Towards Customer Complaints,” *Journal of the Academy of Marketing Science*, 35 (4), 523–36.

Homburg, C., A. Fürst, and N. Koschate (2010), “On the Importance of Complaint Handling Design: A Multi-Level Analysis of the Impact in Specific Complaint Situations,” *Journal of the Academy of Marketing Science*, 38 (3), 265–87.

Homburg, C. and A. Giering (1996), “Konzeptualisierung und Operationalisierung komplexer Konstrukte: Ein Leitfaden für die Marketingforschung [Conceptualization and Operationalization of Complex Constructs: A Guide for Marketing Research],” *Marketing ZfP*, 1, 5–24.

Hu, L. and P. M. Bentler (1999), “Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives,” *Structural Equation Modeling: A Multidisciplinary Journal*, 6 (1), 1–55.

Hunt, S. and R. Morgan (1995), “The Comparative Advantage Theory of Competition,” *The Journal of Marketing*, 59 (2), 1–15.

Jayachandran, S., S. Sharma, P. Kaufman, and P. Raman (2005), "The Role of Relational Information Processes and Technology Use in Customer Relationship Management," *Journal of Marketing*, 69 (4), 177–92.

Johnson, A. R., M. Matear, and M. Thomson (2011), "A Coal in the Heart: Self-Relevance as a Post-Exit Predictor of Consumer Anti-Brand Actions," *Journal of Consumer Research*, 38 (1), 108–25.

Johnson, M., E. Anderson, and C. Fornell (1995), "Rational and Adaptive Performance Expectations in a Customer Satisfaction Framework," *Journal of Consumer Research*, 21 (4), 695–707.

Johnston, R. (2001), "Linking Complaint Management to Profit," *International Journal of Service Industry Management*, 12 (1), 60–9.

Jones, M. A., D. L. Mothersbaugh, and S. E. Beatty (2000), "Switching Barriers and Repurchase Intentions in Services," *Journal of Retailing*, 76 (2), 259–74.

——— (2003), "The Effects of Locational Convenience on Customer Repurchase Intentions across Service Types," *Journal of Services Marketing*, 17 (7), 701–12.

Jones, T., P. A. Dacin, and S. F. Taylor (2011), "Relational Damage and Relationship Repair," *Journal of Service Research*, 14 (3), 318–39.

Kale, S. (2004), "CRM Failure and the Seven Deadly Sins," *Marketing Management*, 13 (5), 42–6.

Kaplan, R. S. (2012), "When to Drop an Unprofitable Customer," *Harvard Business Review*, 90 (4), 137–41.

Kau, A.-K. and E. W.-Y. Loh (2006), "The Effects of Service Recovery on Consumer

Satisfaction: A Comparison between Complainants and Non-Complainants," *Journal of Services Marketing*, 20 (2), 101–11.

Keaveney, S. M. (1995), "Customer Switching Behavior in Service Industries: An Exploratory Study," *Journal of Marketing*, 59 (2), 71–82.

Kelley, S. and M. Davis (1994), "Antecedents to Customer Expectations for Service Recovery," *Journal of the Academy of Marketing Science*, 22 (1), 52–61.

Kelley, S. W., K. D. Hoffman, and M. A. Davis (1993), "A Typology of Retail Failures and Recoveries," *Journal of Retailing*, 69 (4), 429–52.

Kennedy, P. E. (2003), *A Guide to Econometrics*, Cambridge, MA: MIT Press.

Kennedy, P. (2005), "Oh No! I Got the Wrong Sign! What Should I Do?" *Journal of Economic Education*, 36 (1), 77–92.

Kim, T., W. Kim, and H. Kim (2009), "The Effects of Perceived Justice on Recovery Satisfaction, Trust, Word-of-Mouth, and Revisit Intention in Upscale Hotels," *Tourism Management*, 30 (1), 51–62.

Klein, J. and N. Dawar (2004), "Corporate Social Responsibility and Consumers' Attributions and Brand Evaluations in a Product-Harm Crisis," *International Journal of Research in Marketing*, 21 (3), 203–17.

Klemperer, P. (1995), "Competition When Consumers Have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade," *Review of Economic Studies*, 62 (4), 515–39.

LaBarbera, P. A. and D. Mazursky (1983), "A Longitudinal Assessment of Consumer

Satisfaction/Dissatisfaction: The Dynamic Aspect of the Cognitive Process," *Journal of Marketing Research*, 20 (4), 393–404.

Lam, S. Y., V. Shankar, and M. K. E. B. Murthy (2004), "Customer Value, Satisfaction, Loyalty, and Switching Costs: An Illustration from a Business-to-Business Service Context," *Journal of the Academy of Marketing Science*, 32 (3), 293–311.

Lewis, R. C. (1983), "Consumers Complain - What Happens When Business Responds?" in *International Fare in Consumer Satisfaction and Complaining*, R. L. Day and H. K. Hunt, eds., Bloomington, IN: Bureau of Business Research, 88–94.

Liao, H. and A. Chuang (2004), "A Multilevel Investigation of Factors Influencing Employee Service Performance and Customer Outcomes," *Academy of Management Journal*, 47 (1), 41–58.

Liao, H. (2007), "Do it Right this Time: The Role of Employee Service Recovery Performance in Customer-Perceived Justice and Customer Loyalty after Service Failures," *Journal of Applied Psychology*, 92 (2), 475–89.

Lidén, S. and P. Skålén (2003), "The Effect of Service Guarantees on Service Recovery," *International Journal of Service Industry Management*, 14 (1), 36–58.

Magnini, V. P., J. B. Ford, E. B. Markowski, and E. D. Honeycutt (2007), "The Service Recovery Paradox: Justifiable Theory or Smoldering Myth?" *Journal of Services Marketing*, 21 (3), 284–98.

Martínez-Tur, V., J. M. Peir, J. O. S. Ramos, and C. Moliner (2006), "Justice Perceptions as Predictors of Customer Satisfaction: The Impact of Distributive, Procedural, and Interactional Justice," *Journal of Applied Social Psychology*, 36 (1), 100–19.

de Matos, C. A., J. L. Henrique, and C. A. V. Rossi (2007), “Service Recovery Paradox: A Meta-Analysis,” *Journal of Service Research*, 10 (1), 60–77.

Mattila, A. S. (2001), “The Impact of Relationship Type on Customer Loyalty in a Context of Service Failures,” *Journal of Service Research*, 4 (2), 91–101.

——— (2004), “The Impact of Service Failures on Customer Loyalty: The Moderating Role of Affective Commitment,” *International Journal of Service Industry Management*, 15 (2), 134–49.

Mattila, A. (2006), “The Power of Explanations in Mitigating the Ill-Effects of Service Failures,” *Journal of Services Marketing*, 20 (7), 422–28.

Maxham, I., James G. (2001), “Service Recovery’s Influence on Consumer Satisfaction, Positive Word-of-Mouth, and Purchase Intentions,” *Journal of Business Research*, 54 (1), 11–24.

Maxham, I., James G. and R. G. Netemeyer (2002a), “A Longitudinal Study of Complaining Customers’ Evaluations of Multiple Service Failures and Recovery Efforts,” *Journal of Marketing*, 66 (4), 57–71.

——— (2002b), “Modeling Customer Perceptions of Complaint Handling over Time: The Effects of Perceived Justice on Satisfaction and Intent,” *Journal of Retailing*, 78 (4), 239–52.

——— (2003), “Firms Reap What They Sow: The Effects of Shared Values and Perceived Organizational Justice on Customers’ Evaluations of Complaint Handling,” *Journal of Marketing*, 67 (1), 46–62.

McAlister, D. and R. Erffmeyer (2003), “A Content Analysis of Outcomes and Responsibilities for Consumer Complaints to Third-Party Organizations," *Journal of Business Research*, 56 (4), 341–51.

McAlister, L. and E. Pessemier (1982), "Variety Seeking Behavior: An Interdisciplinary Review," *Journal of Consumer Research*, 9 (3), 311–22.

McCollough, M. A., L. L. Berry, and M. S. Yadav (2000), "An Empirical Investigation of Customer Satisfaction after Service Failure and Recovery," *Journal of Service Research*, 3 (2), 121–37.

Michel, S. and M. L. Meuter (2008), "The Service Recovery Paradox: True but Overrated?" *International Journal of Service Industry Management*, 19 (4), 441–75.

Miller, C. E., J. Reardon, and D. E. McCorkle (1999), "The Effects of Competition on Retail Structure: An Examination of Intratype, Intertype, and Intercategory Competition," *Journal of Marketing*, 63 (4), 107–20.

Miller, J., C. Craighead, and K. Karwan (2000), "Service Recovery: A Framework and Empirical Investigation," *Journal of Operations Management*, 18 (4), 387–400.

Mindlin, A. (2009), "No Fury like a Customer Mistreated," *The New York Times*. Available from: `http://www.nytimes.com/2009/10/05/business/05drill.html` [accessed April 30, 2012].

Mithas, S., M. Krishnan, and C. Fornell (2005), "Why Do Customer Relationship Management Applications Affect Customer Satisfaction?" *Journal of Marketing*, 69 (4), 201–09.

Mittal, V. and W. A. Kamakura (2001), "Satisfaction, Repurchase Intent, and Repurchase Behavior: Investigating the Moderating Effect of Customer Characteristics," *Journal of Marketing Research*, 38 (1), 131–42.

Moorman, C., R. Deshpandé, and G. Zaltman (1993), "Factors Affecting Trust in Market Research Relationships," *Journal of Marketing*, 57 (1), 81–101.

Morgan, R. M. and S. D. Hunt (1994), "The Commitment-Trust Theory of Relationship Marketing," *Journal of Marketing*, 58 (3), 20–38.

Morimoto, L., E. White, and P. Newcomb (2003), "Selection Bias in the Assessment of Gene-Environment Interaction in Case-Control Studies," *American Journal of Epidemiology*, 158 (3), 259–63.

Morwitz, V. G. and D. Schmittlein (1992), "Using Segmentation to Improve Sales Forecasts Based on Purchase Intent: Which 'Intenders' Actually Buy?" *Journal of Marketing Research*, 29 (4), 391–405.

Mulhern, F. (1999), "Customer Profitability Analysis: Measurement, Concentration, and Research Directions," *Journal of Interactive Marketing*, 13 (1), 25–40.

Netemeyer, R. G., W. O. Bearden, and S. Sharma (2003), *Scaling Procedures: Issues and Applications*, Thousand Oaks, CA: Sage Publications.

Nunnally, J. C. (1978), *Psychometric Theory*, New York, NY: McGraw-Hill.

van Oest, R. and G. Knox (2011), "Extending the BG/NBD: A Simple Model of Purchases and Complaints," *International Journal of Research in Marketing*, 28 (1), 30–7.

Oliver, R. L. (1996), *Satisfaction: A Behavioral Perspective on the Consumer*, Boston, MA: Irwin/McGraw-Hill.

Oliver, R. (1980), “A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions,” *Journal of Marketing Research*, 17 (4), 460–69.

——— (1981), “Measurement and Evaluation of Satisfaction Processes in Retail Settings,” *Journal of Retailing*, 57 (3), 25–48.

Olsen, L. L. and M. D. Johnson (2003), “Service Equity, Satisfaction, and Loyalty: From Transaction-Specific to Cumulative Evaluations,” *Journal of Service Research*, 5 (3), 184–95.

Orsingher, C., S. Valentini, and M. de Angelis (2010), “A Meta-Analysis of Satisfaction with Complaint Handling in Services,” *Journal of the Academy of Marketing Science*, 38 (2), 169–86.

Palmatier, R. W., R. P. Dant, D. Grewal, and K. R. Evans (2006), “Factors Influencing the Effectiveness of Relationship Marketing: A Meta-Analysis,” *Journal of Marketing*, 70 (4), 136–53.

Parasuraman, A. (2006), “Modeling Opportunities in Service Recovery and Customer-Managed Interactions,” *Marketing Science*, 25 (6), 590–93.

Parsons, L. S. (2001), “Reducing Bias in a Propensity Score Matched-Pair Sample Using Greedy Matching Techniques,” *Proceedings of the Twenty-Sixth Annual SAS Users Group International Conference*, SAS Institute, Paper 214–26.

Parvatiyar, A., J. Sheth, and F. Whittington (1992), “Paradigm Shift in Interfirm Marketing Relationships: Emerging Research Issues,” working paper, Goizueta Business School, Emory University.

Patterson, P. G., E. Cowley, and K. Prasongsukarn (2006), “Service Failure Recovery:

The Moderating Impact of Individual-Level Cultural Value Orientation on Perceptions of Justice," *International Journal of Research in Marketing*, 23 (3), 263–77.

Payne, A. and P. Frow (2005), "A Strategic Framework for Customer Relationship Management," *Journal of Marketing*, 69 (4), 167–76.

Porter, M. (2001), "Strategy and the Internet," *Harvard Business Review*, 79 (3), 62–79.

Priluck, R. (2003), "Relationship Marketing Can Mitigate Product and Service Failures," *Journal of Services Marketing*, 17 (1), 37–52.

Reichheld, F. F. (2003), *Loyalty Rules: How Today's Leaders Build Lasting Relationships.*, Boston, MA: Harvard Business School Press.

Reichheld, F. F. and W. E. Sasser (1990), "Zero Defections: Quality Comes to Services," *Harvard Business Review*, 68 (5), 105–11.

Reinartz, W., M. Krafft, and W. D. Hoyer (2004), "The Customer Relationship Management Process: Its Measurement and Impact on Performance," *Journal of Marketing Research*, 41 (3), 293–305.

Reinartz, W. J. and V. Kumar (2000), "On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing," *Journal of Marketing*, 64 (4), 17–35.

Rezabakhsh, B., D. Bornemann, U. Hansen, and U. Schrader (2006), "Consumer Power: A Comparison of the Old Economy and the Internet Economy," *Journal of Consumer Policy*, 29 (1), 3–36.

Rick, S. and G. Loewenstein (2008), "The Role of Emotion in Economic Behavior," in *Handbook of Emotions*, M. Lewis, J. M. Haviland-Jones, and L. Feldman Barrett, eds., New York, NY: The Guilford Press.

del Río-Lanza, A. B., R. Vázquez-Casielles, and A. M. Díaz-Martín (2009), "Satisfaction with Service Recovery: Perceived Justice and Emotional Responses," *Journal of Business Research*, 62 (8), 775–81.

Roehm, M. L. and M. K. Brady (2007), "Consumer Responses to Performance Failures by High-Equity Brands," *Journal of Consumer Research*, 34 (4), 537–45.

Rosenbaum, P. R. and D. B. Rubin (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70 (1), 41–55.

——— (1985), "Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score," *American Statistician*, 39 (1), 33–8.

——— (1984), "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score," *Journal of the American Statistical Association*, 79 (387), 516–24.

Rossiter, J. R. (2002), "The C-OAR-SE Procedure for Scale Development in Marketing," *International Journal of Research in Marketing*, 19 (4), 305–35.

Rubin, D. B. (1973), "Matching to Remove Bias in Observational Studies," *Biometrics*, 29 (1), 159–83.

——— (1977), "Assignment to Treatment Group on the Basis of a Covariate," *Journal of Educational Statistics*, 2 (1), 1–26.

Rust, R. T., V. A. Zeithaml, and K. N. Lemon (2000), *Driving Customer Equity: How Customer Lifetime Value is Reshaping Corporate Strategy*, New York, NY: The Free Press.

Rust, R. T. and T. S. Chung (2006), “Marketing Models of Service and Relationships,” *Marketing Science*, 25 (6), 560–80.

Rust, R. T., K. N. Lemon, and V. A. Zeithaml (2004), “Return on Marketing: Using Customer Equity to Focus Marketing Strategy,” *Journal of Marketing*, 68 (1), 109–27.

Sajtos, L., R. J. Brodie, and J. Whittome (2010), “Impact of Service Failure: The Protective Layer of Customer Relationships,” *Journal of Service Research*, 13 (2), 216–29.

Sánchez-García, I., R. Pieters, M. Zeelenberg, and E. Bigné (2012), “When Satisfied Consumers Do Not Return: Variety Seeking’s Effect on Short- and Long-Term Intentions,” *Psychology & Marketing*, 29 (1), 15–24.

SAS Institute (2000), *SAS Onlinedoc Version 8*, Cary, NC: SAS Institute Inc.

Schoefer, K. and A. Diamantopoulos (2008), “The Role of Emotions in Translating Perceptions of (In)Justice into Postcomplaint Behavioral Responses,” *Journal of Service Research*, 11 (1), 91–103.

Seiders, K. and L. L. Berry (1998), “Service Fairness: What it is and Why it Matters,” *Academy of Management Executive*, 12 (2), 8–20.

Seiders, K., L. L. Berry, and L. G. Gresham (2000), “Attention, Retailers! How Convenient is Your Convenience Strategy?” *Sloan Management Review*, 41 (3), 79–89.

Seiders, K., G. B. Voss, A. L. Godfrey, and D. Grewal (2007), “SERVCON: Development and Validation of a Multidimensional Service Convenience Scale,” *Journal of the Academy of Marketing Science*, 35 (1), 144–56.

Seiders, K., G. B. Voss, D. Grewal, and A. L. Godfrey (2005), “Do Satisfied Customers

Buy More? Examining Moderating Influences in a Retailing Context," *Journal of Marketing*, 69 (4), 26–43.

Severt, D. E. (2002), *The Customer's Path to Loyalty: A Partial Test of the Relationships of Prior Experience, Justice, and Customer Satisfaction.*, Ph.D. dissertation: Virginia Polytechnic Institute at Blacksburg.

Silverman, B. W. (1986), *Density Estimation: For Statistics and Data Analysis*, London: Chapman and Hall.

Singh, J. (1990), "Voice, Exit, and Negative Word-of-Mouth Behaviors: An Investigation across Three Service Categories," *Journal of the Academy of Marketing Science*, 18 (1), 1–15.

Smith, A. K. and R. N. Bolton (1998), "An Experimental Investigation of Customer Reactions to Service Failure and Recovery Encounters: Paradox or Peril?" *Journal of Service Research*, 1 (1), 65–81.

——— (2002), "The Effect of Customers' Emotional Responses to Service Failures on Their Recovery Effort Evaluations and Satisfaction Judgments," *Journal of the Academy of Marketing Science*, 30 (1), 5–23.

Smith, A. K., R. N. Bolton, and J. Wagner (1999), "A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery," *Journal of Marketing Research*, 36 (3), 356–72.

Smith, J. A. and P. E. Todd (2005), "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics*, 125 (1-2), 305–53.

SOCAP (1994), *Corporate Guide to Effective Complaint Management*, Alexandria, VA: SOCAP.

Solomon, M. R., C. Surprenant, J. A. Czepiel, and E. G. Gutman (1985), "A Role Theory Perspective on Dyadic Interactions: The Service Encounter," *Journal of Marketing*, 49 (1), 99–111.

Srinivasan, R. and C. Moorman (2005), "Strategic Firm Commitments and Rewards for Customer Relationship Management in Online Retailing," *Journal of Marketing*, 69 (4), 193–200.

Srivastava, R., T. Shervani, and L. Fahey (1998), "Market-Based Assets and Shareholder Value: A Framework for Analysis," *The Journal of Marketing*, 62 (1), 2–18.

Stauss, B. and A. Schoeler (2004), "Complaint Management Profitability: What Do Complaint Managers Know?" *Managing Service Quality*, 14 (2/3), 147–56.

Stauss, B. and W. Seidel (2004), *Complaint Management: The Heart of CRM*, Mason, OH: Thompson/South-Western.

Strativity (2007), "Keeping Customers," *Marketing Management*, 16 (3), 6.

Strizhakova, Y., Y. Tsarenko, and J. A. Ruth (2012), "I'm Mad and I Can't Get That Service Failure Off My Mind: Coping and Rumination as Mediators of Anger Effects on Customer Intentions," *Journal of Service Research*, (forthcoming).

Tax, S. S., S. W. Brown, and M. Chandrashekaran (1998), "Customer Evaluations of Service Complaint Experiences: Implications for Relationship Marketing," *Journal of Marketing*, 62 (2), 60–76.

Technical Assistance Research Program (1986), *Consumer Complaint Handling in*

*America: An Update Study (Part II)*, Washington, DC: Technical Assistance Research Program Institute and the United States Office of Consumer Affairs.

Tripp, T. M. and Y. Grégoire (2011), "When Unhappy Customers Strike Back on the Internet," *MIT Sloan Management Review*, 52 (3), 37–44.

Valenzuela, F., D. Pearson, and R. Epworth (2005), "Influence of Switching Barriers on Service Recovery Evaluation," *Journal of Services Research*, (Special Issue), 239–57.

Varela-Neira, C., R. Vazquez-Casielles, and V. Iglesias-Arguelles (2008), "The Influence of Emotions on Customer's Cognitive Evaluations and Satisfaction in a Service Failure and Recovery Context," *Service Industries Journal*, 28 (4), 497–512.

Venkatesan, R. and V. Kumar (2004), "A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy," *Journal of Marketing*, 68 (4), 106–25.

Verhoef, P. C. (2003), "Understanding the Effect of Customer Relationship Management Efforts on Customer Retention and Customer Share Development," *Journal of Marketing*, 67 (4), 30–45.

Vogel, V., H. Evanschitzky, and B. Ramaseshan (2008), "Customer Equity Drivers and Future Sales," *Journal of Marketing*, 72 (6), 98–108.

Voorhees, C. M., J. Baker, B. L. Bourdeau, E. D. Brocato, and J. Cronin, J. Joseph (2009), "It Depends: Moderating the Relationships among Perceived Waiting Time, Anger, and Regret," *Journal of Service Research*, 12 (2), 138–55.

Voss, G. B., A. Godfrey, and K. Seiders (2010), "How Complementarity and Substitution Alter the Customer Satisfaction-Repurchase Link," *Journal of Marketing*, 74 (6), 111–27.

von Wangenheim, F. and T. Bayón (2007), "Behavioral Consequences of Overbooking Service Capacity," *Journal of Marketing*, 71 (4), 36–47.

Ward, J. C. and A. L. Ostrom (2006), "Complaining to the Masses: The Role of Protest Framing in Customer-Created Complaint Web Sites," *Journal of Consumer Research*, 33 (2), 220–30.

Weun, S., S. E. Beatty, and M. A. Jones (2004), "The Impact of Service Failure Severity on Service Recovery Evaluations and Post-Recovery Relationships," *Journal of Services Marketing*, 18 (2), 133–46.

Wirtz, J. (2011), "How to Deal with Customer Shakedowns," *Harvard Business Review*, 89 (4), 24.

Wirtz, J., A. S. Mattila, and M. O. Lwin (2007), "How Effective Are Loyalty Reward Programs in Driving Share of Wallet?" *Journal of Service Research*, 9 (4), 327–34.

Wirtz, J. and J. McColl-Kennedy (2010), "Opportunistic Customer Claiming During Service Recovery," *Journal of the Academy of Marketing Science*, 38 (5), 654–75.

Woisetschläger, D. M., P. Lentz, and H. Evanschitzky (2011), "How Habits, Social Ties, and Economic Switching Barriers Affect Customer Loyalty in Contractual Service Settings," *Journal of Business Research*, 64 (8), 800–08.

Wooldridge, J. M. (2002a), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press.

——— (2002b), *Introductory Econometrics: A Modern Approach*, Mason, OH: South-Western.

Wu, L.-W. (2011), "Beyond Satisfaction: The Relative Importance of Locational Convenience, Interpersonal Relationships, and Commitment across Service Types," *Managing Service Quality*, 21 (3), 240–63.

Xue, M., L. M. Hitt, and P.-y. Chen (2011), "Determinants and Outcomes of Internet Banking Adoption," *Management Science*, 57 (2), 291–307.

Zineldin, M. (2006), "The Royalty of Loyalty: CRM, Quality and Retention," *Journal of Consumer Marketing*, 23 (7), 430–37.

# Appendix


**A Descriptive Sample Statistics** **187**

**B Evaluation Criteria for Latent Variables** **189**

**C Multi-Item Survey Measures** **191**

**D Single-Item Survey Measures** **195**

**E Results of the Logistic Regression** **196**

# A Descriptive Sample Statistics

**Table A.1:** Descriptive Sample Statistics - Part I

| **Characteristic** | **Random Sample (N=24015)** | **Sample after 1st Survey (N=5688)** | **Sample after 2nd Survey (N=2318)** |
|---|---|---|---|
| Age | NA | 49.28<br>(13.65) | 50.72<br>(13.76) |
| Gender<br>(% Women) | NA | 28.92 | 27.09 |
| Marital status (%) | | | |
| - single | 16.13 | 14.95 | 13.85 |
| - unmarried couple | 15.42 | 14.66 | 13.72 |
| - married | 68.45 | 70.40 | 72.43 |
| # of people in household | 2.68<br>(1.35) | 2.66<br>(1.30) | 2.65<br>(1.31) |
| # of children | 1.05<br>(1.38) | 1.01<br>(1.18) | 1.00<br>(1.17) |
| Monthly household income (%) | 9.61[a] | 8.63[a] | 7.97[a] |
| 0-1500 € | 26.08 | 24.38 | 22.89 |
| 1501-2000 € | 26.28 | 27.07 | 28.27 |
| 2001-2500 € | 17.05 | 17.97 | 19.25 |
| 2501-3750 € | 13.28 | 14.42 | 14.60 |
| 3751 € and more | 7.70 | 7.52 | 7.04 |
| Relationship length (years) | NA | 11.33<br>(6.23) | 11.70<br>(6.35) |
| Time since enrollment in loyalty program (months)[b] | 24.92<br>(9.88) | 26.13<br>(9.69) | 26.53<br>(9.75) |

**Table A.2:** Descriptive Sample Statistics - Part II

| Characteristic | Random Sample (N=24015) | Sample after 1st Survey (N=5688) | Sample after 2nd Survey (N=2318) |
|---|---|---|---|
| Revenue / year (€) | 1157.66[c] (1233.13) | 1219.04 (1229.41) | 1202.85 (1214.67) |
| Revenue / month (€) | 96.47[c] (102.76) | 101.59 (102.45) | 100.24 (101.22) |
| # of transactions / year | 26.62[c] (23.67) | 28.69 (23.71) | 28.67 (23.91) |
| # of transactions / month | 2.22[c] (1.97) | 2.39 (1.98) | 2.39 (1.99) |
| Value of monthly transactions (€) | 44.87[c] (42.82) | 44.15 (39.94) | 43.19 (37.84) |
| Interpurchase time (months) | 2.38[c] (1.73) | 2.22 (1.69) | 2.21 (1.68) |
| Store distance (kilometers) | 11.11[c] (9.03) | 11.07 (8.88) | 10.82 (8.46) |

[a]Percentage of customers who did not specify their income level. [b]Refers to the point in time when the 1st survey was conducted. The loyalty program was launched four years earlier. [c]Significant differences to survey samples could be detected.
Notes: Mean values and standard deviations are in parentheses, if not indicated otherwise;
NA = not applicable.

# B Evaluation Criteria for Latent Variables

**Table B.1:** Evaluation of Overall Model Fit

| | Fit Index | Cutoff Value | Source |
|---|---|---|---|
| Absolute model fit | Chi-Square ($\chi^2$) Test | - Significance of $\chi^2$-Test<br>- $\chi^2$/df $\leq 3$ | Homburg and Giering (1996) |
| | RMSEA | $\leq .06$<br>$\leq .05$ (close model fit) | Hu and Bentler (1999)<br>Browne and Cudeck (1992) |
| | SRMR | $\leq .08$ | Hu and Bentler (1999) |
| Model fit in comparison with null model | CFI | $\geq .95$ | Hu and Bentler (1999) |
| | TLI | $\geq .95$ | Hu and Bentler (1999) |

**Table B.2:** Evaluation Criteria of Latent Constructs

| | **Criterion** | **Cutoff Value** | **Source** |
|---|---|---|---|
| *First generation* | Cronbach's $\alpha$ | $\geq .79$ | Nunnally (1978) |
| | Item-to-total correlation | Items with low item-to-total correlations that decrease Cronbach's $\alpha$ should be removed from the scale. | Churchill (1979) |
| | Explained variance in EFA | $\geq 50\,\%$ | Netemeyer, Bearden, and Sharma (2003) |
| *Second generation* | Composite reliability | $\geq .60$ | Bagozzi and Yi (1988) |
| | Average variance extracted (AVE) | $\geq .50$ | Bagozzi and Yi (1988) |
| | Indicator reliability | $\geq .40$ | Homburg and Giering (1996) |
| | Fornell-Larcker criterion | Squared correlations must be less than than AVE. | Fornell and Larcker (1981) |

Note: The distinction between *first generation* and *second generation* criteria follows Homburg and Giering (1996).

# C Multi-Item Survey Measures

The EFA and CFA were conducted projectwise; only the constructs contained in the model of the respective project were included. The EFA was executed using SPSS, and the CFA was run in SAS with the Calis procedure (see ch. 19 in SAS Institute [2000]).

The global fit measures for the CFA of *Project II* indicate an acceptable model fit:

- $\chi^2$/degrees of freedom = 2.32, $p < .01$
- Root mean square error of approximation (RMSEA) = .05
- Standardized root mean square residual (SRMR) = .02
- Comparative fit index (CFI) = .99
- Tucker-Lewis index (TLI) = .99

The constructs employed in *Project II* appear in Table C.1. The *satisfaction* scale was taken from Bettencourt (1997) and De Wulf, Odekerken-Schröder, and Iacobucci (2001). The measure of *interactional justice* was adapted from Homburg and Fürst (2005) and Tax, Brown, and Chandrashekaran (1998), and the *procedural justice* scale comes from Blodgett, Hill, and Tax (1997) and Smith and Bolton (1998).

**Table C.1:** Multi-Item Survey Measures (*Project II)*

| Construct | Item | M | SD | I-t-t | IR | % VE | CA | CR |
|---|---|---|---|---|---|---|---|---|
| Satisfaction | I am satisfied with the relationship I have with X. | 4.53 | 1.72 | .70 | .63 | 79.92 | .87 | .88 |
| | Based on all my experiences with X, I am very satisfied. | 4.63 | 1.77 | .78 | .71 | | | |
| | Compared to other stores, I am very satisfied with X. | 4.86 | 1.76 | .79 | .77 | | | |
| Interactional justice | The employees were very keen to solve my problem. | 5.21 | 1.74 | .80 | .74 | 86.88 | .88 | .87 |
| | Personnel was unhesitating to react to my complaint. | 4.96 | 1.83 | .82 | .86 | | | |
| | I felt treated very friendly by the employees. | 5.30 | 1.45 | .68 | .47 | | | |
| Procedural justice | X facilitates customer complaints. | 4.40 | 2.11 | .66 | .49 | 89.56 | .86 | .87 |
| | My complaint was handled in a timely manner. | 4.32 | 2.24 | .78 | .83 | | | |
| | X put an adequate amount of time into resolving my problem. | 4.48 | 2.20 | .79 | .75 | | | |

Notes: M = mean, SD = standard deviation, I-t-t = Item-to-total correlation, IR = indicator reliability, VE = variance explained, CA = Cronbach's alpha, CR = composite reliability.

**Table C.2:** Discriminant Validity (*Project II*)

| | Construct | AVE | 1 | 2 | 3 |
|---|---|---|---|---|---|
| 1. | Satisfaction | .71 | 1 | | |
| 2. | Interactional justice | .70 | .56 | 1 | |
| 3. | Procedural justice | .69 | .33 | .38 | 1 |

Notes: The squared correlations between constructs appear below the diagonal; AVE = average variance extracted.

The global fit measures for the CFA of *Project III* indicate a good model fit:

- $\chi^2$/degrees of freedom = 1.99, $p < .001$
- Root mean square error of approximation (RMSEA) = .06
- Standardized root mean square residual (SRMR) = .03
- Comparative fit index (CFI) = .99
- Tucker-Lewis index (TLI) = .99

The constructs employed in *Project III* appear in Table C.3. The *affective commitment* scale was adapted from Fullerton (2003), the *relationship commitment* measure was taken from De Wulf, Odekerken-Schröder, and Iacobucci (2001), and the *switching costs* construct comes from Jones, Mothersbaugh, and Beatty (2000).

**Table C.3:** Multi-Item Survey Measures (*Project III*)

| Construct | Item | M | SD | I-t-t | IR | % VE | CA | CR |
|---|---|---|---|---|---|---|---|---|
| Affective commitment | I feel emotionally attached to X. | 4.12 | 1.92 | .81 | .72 | 88.33 | .93 | .95 |
| | I feel like a part of a family as a customer of X. | 3.29 | 2.02 | .88 | .88 | | | |
| | I feel a strong sense of belonging to X. | 3.42 | 1.99 | .91 | .95 | | | |
| Relationship commitment | Even if this store would be more difficult to reach, I would still keep buying there. | 3.82 | 1.90 | .81 | .79 | 90.69 | .90 | .87 |
| | I am willing "to go the extra mile" to remain a customer of this store. | 3.76 | 1.97 | .81 | .86 | | | |
| Switching costs | It takes me a great deal of time and effort to get used to a new store. | 3.75 | 2.12 | .78 | .68 | 73.43 | .90 | .91 |
| | In general it would be a hassle changing retail stores. | 3.74 | 2.00 | .85 | .89 | | | |
| | For me, the costs in time, money, and effort to switch the store are high. | 3.59 | 2.17 | .77 | .65 | | | |

Notes: M = mean, SD = standard deviation, I-t-t = Item-to-total correlation, IR = indicator reliability, VE = variance explained, CA = Cronbach's alpha, CR = composite reliability.

**Table C.4:** Discriminant Validity (*Project III*)

| | Construct | AVE | 1 | 2 | 3 |
|---|---|---|---|---|---|
| 1. | Affective commitment | .85 | 1 | | |
| 2. | Relationship commitment | .77 | .52 | 1 | |
| 3. | Switching costs | .77 | .27 | .15 | 1 |

Notes: The squared correlations between constructs appear below the diagonal; AVE = average variance extracted.

# D Single-Item Survey Measures

The *single-item measures* were adapted from Maxham and Netemeyer (2002a,b), Smith, Bolton, and Wagner (1999) and Tax, Brown, and Chandrashekaran (1998).

**Table D.1:** Single-Item Survey Measures

| **Measure** | **Item** | **M (SD)** | |
|---|---|---|---|
| | | $\mathbf{t_0}$ | $\mathbf{t_1}$ |
| Failure resolution | The failure was resolved to my fullest satisfaction. | NA | 3.88 (2.24) |
| Resolution speed | My complaint was processed quickly. | NA | 4.32 (2.24) |
| Failure severity | The failure I encountered was severe. | NA | 4.89 (1.79) |
| Distributive justice | I received an adequate compensation for the inconveniences associated with the failure from X. | NA | 2.97 (2.38) |
| Failure responsibility | The problem that I encountered was all X's fault. | NA | 4.97 (2.14) |
| Repurchase intent | I will purchase at X again. | 5.58 (1.36) | 5.67 (1.34) |
| Word-of-mouth intent | I will recommend X. | 5.60 (1.39) | 5.55 (1.48) |
| Share of wallet (in %) | What percentage of your total category expenditures do you spend at this retailer? | 69.06 (23.80) | 69.87 (27.66) |
| Relationship length | For how long have you been a customer of X? | 11.70 (6.35) | NA |
| Age | What is your age? | 50.72 (13.76) | NA |
| Gender | What is your gender? | 27.09[a] | NA |

[a]Percentage of female respondents. Notes: NA = not applicable.

# E Results of the Logistic Regression

**Table E.1:** Results of the Logistic Regression

| **Independent Variables** | **Performance Failure** | |
|---|---|---|
| Constant | -.27 | (.61) |
| | | |
| Revenue ($t_0$) | .78 | (.70) |
| Share of wallet ($t_0$) | .19 | (.47) |
| Relationship length ($t_0$) | .31 | (.99) |
| Satisfaction ($t_0$) | -.18$^*$ | (.10) |
| Repurchase intent ($t_0$) | -.09 | (.09) |
| Word-of-mouth intent ($t_0$) | -.17$^*$ | (.10) |
| Age | -.01$^*$ | (.70) |
| Gender | .72 | (.23) |

$^{***}p < .01$, $^{**}p < .05$, $^{*}p < .10$; n.s. = not significant.
Notes: The coefficients for gender, relationship length, and share of wallet were multiplied by 100. The coefficient of revenue was multiplied by 10,000.