The Long-Term Impact of Service Failure and Recovery

Tim Norvell¹, Piyush Kumar², and Mayukh Dass³

Abstract
This article examines customers’ short-term attitudinal and long-term behavioral responses to service failures and recovery efforts. Our data from a tracking study of casual dining restaurants customers indicate that those who did not experience any failure were more satisfied than those who experienced successful recovery following a failure. The satisfactory recovery group, in turn, was more satisfied than customers who either did not complain or were not successfully recovered following their complaints. Importantly, the pattern of brand patronage over the medium and long run differed substantially from the short-term variation in satisfaction levels across the four customer groups. In the medium term, the brand visitation frequency for those who never experienced failure was similar to those of customers who were successfully recovered. The visitation frequencies of customers who did not complain or were not successfully recovered were lower. However, over the long run, the visitation pattern changed substantially, and those who never experienced failure had higher brand patronage frequency than all the three remaining groups that behaved relatively similarly. These results suggest that customers make a distinction between the qualities of the core service and the recovery effort. Although successful recovery temporarily compensates for core failure, its positive influence dissipates over time. In the longer term, customers’ complaining behavior and the firm’s recovery efforts matter less and customers’ brand patronage depends largely on whether or not they experienced core service failure. Nevertheless, firms can recover their investments in service recovery because of increased brand patronage in the medium term.

Keywords
service quality, service recovery, satisfaction, complaining behavior, customer retention, word-of-mouth, service recovery paradox, casual dining, restaurants

The service industry in the United States accounts for more than 80% of its gross domestic product (GDP). The sector has been a key driver of GDP and job growth, accounting for more than US$9.81T and 89.7 million jobs over the past decade (U.S. Bureau of Economic Analysis, 2010). The increasing significance of the industry has attracted academic interest especially in the investigation and measurement of service performance and quality (Cronin & Taylor, 1992; Lovelock & Wirtz, 2011). However, the quality of service experiences often depends on the performance of customer-facing personnel who are often underpaid and undertrained (Bitner, Booms, & Tetreault, 1990). Consequently, many service experiences result in failures and require supplemental effort on the part of service firms to recover customers from such adverse episodes. Because of the inevitability of failure in many services, the study of failure and recovery has emerged as an important stream of research (Andreassen, 2000; Albrecht, Walsh, & Beatty, 2017; Menguc, Auh, Yeniaras, & Katsikeas, 2017).

A service firm’s recovery efforts are contingent upon not only adverse customer experiences (Gustafsson, 2009) but also on customers’ complaining behavior (Kim, Wang, & Mattila, 2010). To initiate recovery efforts, dissatisfied customers need to bring their unfavorable experiences to the firm’s attention because firms do not typically dedicate recovery efforts to noncomplainers. However, those who do complain may not necessarily be satisfied with the firm’s recovery efforts. Therefore, customers can be broadly classified into four groups (Oliver, 1987): satisfied noncomplainers, dissatisfied noncomplainers, satisfied complainers, and dissatisfied complainers.

Although service failures tend to increase customer switching (Roos, 1999), successful recovery can mitigate customer defection and increase retention rates (McCullough, Berry, & Yadav, 2000). In fact, the service recovery paradox (McCullough & Bharadwaj, 1992) suggests that customers
who experience satisfactory recovery may be more loyal than those who never experience failure to begin with. However, the quality of recovery efforts is often mediocre and about one half of customers report dissatisfaction with problem resolution by firms (Berry & Parasuraman, 1991). This may be because of firms’ underinvestment in recovery efforts because a large proportion of dissatisfied customers do not complain (Singh, 1990; Voorhees, Brady, & Horowitz, 2006).

Previous research on service failure and recovery tends to focus either on the service recovery paradox or on customers’ complaining behavior. Although both streams relate to failure, there is limited research at their intersection. As a result, the impact of the four failure-recovery outcomes (Oliver, 1987) on customers’ behavior remains unknown. In this article, we address this gap across the two streams of work and examine the following research questions:

**Research Question 1:** How do service failure, complaining behavior, and recovery efforts affect customers’ short-term attitudinal responses across the four failure-recovery customer groups?

**Research Question 2:** How does brand patronage behavior across the four groups evolve over time?

We use data from a large-scale longitudinal study among customers of casual dining restaurants to address these questions. In contrast to previous research in the area (e.g., DeWitt, Nguyen, & Marshall, 2008), we examine both the attitudinal and behavioral responses from complainers and noncomplainers. We also assess the financial impact of successfully recovered customers as well as dissatisfied noncomplainers.

Based on our analyses of customers’ attitudinal and behavioral responses, we report four key findings that contribute to multiple streams within the postpurchase literature. First, we find that, contrary to the service recovery paradox, customers who experience a successful recovery effort are less satisfied than those who do not experience failure. Second, noncomplainers, who comprise a much larger percentage of the customer population than previously reported, are less satisfied than those in the satisfactory recovery group, but more satisfied than the “double deviation” group comprising of those who complained but experienced poor recovery. Third, we find that in the medium term, customers can be divided into two groups based on their actual brand patronage behavior. Those who never experienced failure and those who were successfully recovered form one group that has higher brand patronage frequency than the second group comprising of noncomplainers and those who were not successfully recovered. Finally, we find that in the long run, the pattern of brand patronage behavior changes substantially, and those who experienced no failure stand apart from the three remaining groups whose behavior becomes largely indistinguishable. Nevertheless, we show that investing in service recovery may yet be warranted because of a favorable medium term impact on revenues.

Overall, our results provide some caveats to the service recovery paradox (McCollough & Bharadwaj, 1992) as well as the four-way classification of customers based on service failure, complaining behavior, and recovery experience (Oliver, 1987). They suggest that customers perhaps maintain a distinction between the quality of a firm’s core service and its conditional recovery effort. Although, in the short run, the quality of the recovery effort may compensate for a failure of the core service, customers’ long-run behavior depends largely on whether or not the firm fails at its core service.

**Related Literature**

It is important for firms to understand the impact of service failure because it is a key determinant of customer loyalty and switching behavior (Holloway & Beatty, 2003; Roos, 1999). Small increases in customer retention can have a magnified impact on profitability (Reichheld & Sasser, 1990). Given a potential connection between service recovery efforts and customer retention, the relationships among failure, recovery, and subsequent customer responses have been debated and also identified as an area needing additional research (De Matos, Henrique, & Alberto Vargas Rossi, 2007; McCollough et al., 2000; Tax, Brown, & Chandrashekaran, 1998).

Much of the debate on the consequences of service recovery centers on what is called the service recovery paradox, that is, the favorable response of customers who were successfully recovered following a failure relative to those who did not experience failure (McCollough & Bharadwaj, 1992). Several researchers have found empirical evidence in support of this paradox (e.g., Smith & Bolton, 2002) while others have failed to find supportive evidence (Andreassen, 2001; McCollough et al., 2000). The explanations offered to reconcile these differences tend to focus on the effects of moderators such as failure severity, failure attribution, and past experience (De Matos et al., 2007). However, some of the inferences regarding the impact of failing and recovering may have been compromised because of an absence of any information about noncomplainers who may have experienced failure but never experienced any recovery (McCollough et al., 2000).

**Theory and Hypotheses**

Service recovery efforts are contingent on the success versus failure of the core service experience (Singh, 1991) whose quality is generally viewed through the lens of the
expectation-disconfirmation paradigm (Oliver, 1993). In cases of failure, overall satisfaction is based on the combined evaluation of the initial experience and the service recovery efforts (Parasuraman, Berry, & Zeithaml, 1991). Although the relative importance of the two components of service remains a matter of debate, Halstead and Page (1992) suggest that the core service evaluation primarily drives satisfaction, and that the recovery effort only mitigates the adverse effect of a service failure. Therefore, we hypothesize that

**Hypothesis 1a (H1a):** Customers who do not experience a service failure will be more satisfied than those who experience a service failure but a satisfactory recovery.

Some customers who experience a service failure may complain to the firm to resolve their issues and obtain some restitution. If the corresponding recovery is satisfactory, the negative impact of the initial failure will be mitigated. However, if the service recovery effort itself is not satisfactory, customers will encounter two dissatisfying experiences within one service episode. This double deviation effect (Bitner et al., 1990) will intensify their overall dissatisfaction. Therefore, we hypothesize that

**Hypothesis 1b (H1b):** Customers who receive adequate service recovery will be more satisfied than those who receive inadequate service recovery.

Satisfaction with a service experience has downstream consequences on future customer behavior. For example, research on postpurchase behavior provides conceptual arguments (e.g., Oliver, 2014) and empirical data (Coil, Keiningham, Aksoy, & Hsu, 2007; Mittal & Kamakura, 2001; Voss, Godfrey, & Seiders, 2010) in support of a positive relationship between satisfaction and repurchase intentions. Research on service recovery also shows a positive link between satisfaction with both the core service as well as the recovery effort repurchase intent (Halstead & Page, 1992; Susskind & Viccari, 2011; Voorhees et al., 2006). These findings suggest that the ordinal ranking of purchase intent across the four groups identified in H1a/H1b will be similar to that for overall satisfaction. Therefore, we hypothesize that

**Hypothesis 2a (H2a):** Customers who do not experience a service failure will exhibit higher repurchase intent than those who experience a service failure with satisfactory recovery.

**Hypothesis 2b (H2b):** Customers who receive adequate service recovery will exhibit higher repurchase intent than those who receive inadequate service recovery.

Another important consequence of the quality of a service experience is word-of-mouth behavior, which tends to be related to satisfaction (Bejou & Palmer, 1998; Maxham & Netemeyer, 2002; Swanson & Kelley, 2001), and plays a role in acquiring new customers (Mittal, Kumar, & Tsios, 1999). Although there are several ways of measuring such behavior, the likelihood to recommend a service is a useful measure that tends to be correlated with future firm sales (Reichheld, 2003). This relationship exists in part because customers’ likelihood to recommend tends to be related to their own past experience and their future purchase behavior (Zeithaml, Berry, & Parasuraman, 1996). We therefore expect that the likelihood to recommend across the four groups of customers will follow the same pattern as their satisfaction levels and hypothesize that

**Hypothesis 3a (H3a):** Customers who do not experience a service failure will be more likely to recommend a brand than those who experience a service failure with adequate recovery.

**Hypothesis 3b (H3b):** Customers who receive adequate service recovery will be more likely to recommend a brand than those who experience inadequate service recovery.

**Noncomplainers**

A majority of dissatisfied customers do not complain (Singh, 1990; Stephens & Gwinner, 1988). However, very little research exists that compares noncomplainers with the other three customer groups in Oliver’s (1987) service experience classification (Voorhees et al., 2006). To that extent, our understanding of complaining behavior is somewhat limited (Sharma, Marshall, Alan Reday, & Na, 2010; Singh & Wilkes, 1996). Service firms remain concerned about the large percentage of noncomplainers because these customers represent missed opportunities to recover from failures. Furthermore, dissatisfied noncomplainers can still spread negative word-of-mouth which can hurt the firm’s reputation among other customers (Davidow, 2003). Finally, the firm misses vital feedback that may enable it to solve problems for future customers (Fornell & Wernerfelt, 1987). Because noncomplainers do not provide the firm with the opportunity to recover from its service failure, they are likely to remain dissatisfied with the brand. Their satisfaction is likely to be lower than that of customers who experienced a service failure, complained, and were provided with an adequate recovery. Therefore, we hypothesize that

**Hypothesis 4a (H4a):** Noncomplainers will have lower satisfaction than those who complain and receive adequate service recovery.
However, some customers who complain after experiencing a service failure may receive inadequate recovery. Because these customers have expended time and effort to complain, the unsatisfactory recovery will magnify the adversity of their experience (Johnston & Fern, 1999) and lower satisfaction further than what it would have been if they had not complained. Therefore, we hypothesize that

**Hypothesis 4b (H4b):** Customer who complain about a service failure and receive inadequate recovery will experience lower satisfaction that those who experience failure but do not complain.

Consistent with the extant research and prior hypotheses, the ordinal effects in customer satisfaction will translate into similar effects on repurchase intent.

**Hypothesis 5a (H5a):** Noncomplainers will exhibit lower repurchase intent than those who experience a service failure but receive adequate recovery.

**Hypothesis 5b (H5b):** Noncomplainers will exhibit higher repurchase intent than those who complain about a service failure and receive inadequate recovery.

Similarly, the ordinal effects that are stated in previous hypotheses will apply to the likelihood to recommend.

**Hypothesis 6a (H6a):** Noncomplainers will be less likely to recommend the brand than those who experience a service failure but had a satisfactory recovery.

**Hypothesis 6b (H6b):** Noncomplainers will be more likely to recommend the brand than those who experience a service failure and had an unsatisfactory recovery.

**Impact on Future Behavior**

Most studies on service recovery focus on satisfaction and repurchase intentions as the key dependent variables (De Matos et al., 2007). A few use word-of-mouth (e.g., Hocutt, Bowers, & Todd Donavan, 2006; Kau & Loh, 2006; Maxham & Netemeyer, 2002), corporate image (Andreassen, 2001; Kwortnik, 2006), trust (Kau & Loh, 2006), quality (McCollough, 1995), complaint intentions (Hocutt, Chakraborty, & Mowen, 1997), switching intentions, or willingness to pay (Zeithaml et al., 1996) as key dependent variables. However, because these studies track intentions, and not actual long-term purchase behavior, they do not directly address questions pertaining to the true cost of service failure or the financial benefits of successful recovery. In contrast, we explicitly predict that the pattern of behavioral response across the four failure-recovery groups will correspond to that on the attitudinal variables discussed in H1 and H2, and formally, we hypothesize that

**Hypothesis 7a (H7a):** Customers who do not experience a service failure will visit the brand more in the future than those who experience a service failure but have an adequate recovery.

**Hypothesis 7b (H7b):** Customers who experience a service failure but have an adequate recovery will visit the brand more in the future than those who experience a service failure but did not complain.

**Hypothesis 7c (H7c):** Customers who experience a service failure but do not complain will visit the brand more than those who experience a service failure and an inadequate recovery.

**Study**

We tested our hypotheses using data from an online study among a panel of 8,800 customers of casual dining restaurants. The respondents were recruited from a nationally representative panel of individuals who had visited a casual dining restaurant during the previous 30 days. A professional market research firm maintained the panel and compensated the respondents for their participation. The respondents provided data about 10 major brands on a range of attitudinal and behavioral measures, including visitation frequency, satisfaction, intention to revisit, and the likelihood to recommend.

We obtained specific information pertaining to service failure and recovery by asking participants detailed information about their most recent dining experience at one of the 10 restaurant chains, including whether they had experienced a major problem. If participants identified a service failure, they also reported whether or not they complained to the firm, and the satisfaction with the firm’s recovery efforts. We used information on failure, complaining behavior, and the recovery effort to classify respondents into four groups (Oliver, 1987).

We then tracked the behavior of a subset of 1,282 respondents in a longitudinal study that involved data collection every 3 months for the following 3 years. In each wave of data collection, we asked the respondents which brands of casual dining restaurants they had visited during the past 30 days and additional questions about their most recent visit. We used this information to identify whether a customer had experienced a second service failure.

**Measures**

We assigned customers to four categories based on their most recent experience during the first wave of the study. We refer to the brand identified in the first wave as the focal brand for a specific participant. We use data on the focal brand for each participant in the subsequent longitudinal analyses. Those who stated that they did not encounter a service failure in their most recent experience were placed
in the control group and were designated as customers to whom the focal brand was successful in delivering its core service. Those who experienced a service failure were then asked if they had complained to anyone and the brand they visited was noted. Out of the total sample in the first wave, 8.1% or 711 customers experienced a service failure during their most recent visit.

We then queried those who stated that they had voiced a complaint to the firm’s employee about the recovery process. Of those who experienced a failure, 306 customers did not voice their complaint to an employee. We then classified those who did complain on the basis of their satisfaction with the recovery efforts. We classified participants who provided a top-two-box rating (very satisfied or somewhat satisfied) on a 5-point scale (1 = very dissatisfied, 5 = very satisfied) as “satisfied with their service recovery.” The remaining complainers were classified as “dissatisfied with their service recovery,” also known as the double deviation group. In all, 188 participants in our sample were satisfied with the recovery and 217 were not. Therefore, about 46% of the complainers received a satisfactory recovery, which is consistent with prior studies (Berry & Parasuraman, 1991; Best & Andreasen, 1977).

Following the postpurchase literature, we used overall satisfaction, repurchase intent, and the propensity to recommend as measures of the customer postpurchase attitudinal response to their respective focal brands (Maxham & Netemeyer, 2002; Reichheld, 2003; Zeithaml et al., 1996). We measured each of these constructs using 5-point single-item scales. The anchors for repurchase intent and willingness to recommend were 1 for very unlikely and 5 for very likely. Although the use of single-item measures has been debated, such measures have been used in numerous large-scale surveys (e.g., Bolton & Drew, 1991; Mittal et al., 1999; Mittal, Ross, & Baldasare, 1998). Furthermore, because our data were collected as a part of a large, multiwave brand perception study, minimizing respondent fatigue (Lehmann, McAlister, & Staelin, 2011) and dropout rate was paramount. Multi-item scales increase survey length which can decrease the response rate leading to lower reliability (LaBarbera & Mazursky, 1983).

We then used a subset of the original sample in a longitudinal tracking study of brand visitation behavior over the next 12 quarters. The visitation measure was a binary variable based on a question about whether the respondent had visited a specific brand in the past 30 days. For our analysis, we focused on the visitation to their respective focal brands identified in the first wave of data collection. In other words, for every respondent, we focused on the behavior with respect to just one brand across all waves. A total of 1,282 customers participated in the longitudinal study. Of these, 966 were in the control group, 78 were in the satisfactory recovery group, 139 were in the noncomplainer group, and 99 were in the double deviation group.

**Results**

**Post experience attitudinal response.** To test our hypotheses, we first classified customers into one of the following four categories based on their experience, complaining behavior, and satisfaction with the recovery effort for the focal brand in the first wave of data collection:

1. Customers who did not experience a service failure (“no failure” or “control” group)
2. Customers who experienced a service failure, voiced a complaint, and were satisfied with the resolution to their problem (“satisfactory recovery” group)
3. Customers who experienced a service failure but did not voice a complaint (“noncomplainer” group)
4. Customers who experienced a service failure, voiced a complaint but were not satisfied with the resolution to their problem (“double deviation” group)

We used a multivariate general linear model to evaluate the differences among the four customer groups across the three measures of attitudinal responses. The model yielded a significant main effect of the customer group, Wilks’s $\lambda =$ .828, $F(9, 21380) = 191.467, p < .001$, and significant effects for each of the three dependent variables: satisfaction, $F(3, 8787) = 448.045, p < .001$; revisit intent, $F(3, 8787) = 382.991, p < .001$; and the likelihood to recommend, $F(3, 8787) = 408.828, p < .001$.

**Service recovery paradox.** We used planned comparisons to test our hypotheses pertaining to the attitudinal responses. We find that the control group had higher satisfaction ($M = 4.43$) than those who experienced a failure and had satisfactory recovery, $M = 3.71$, $t(8266) = −7.663, p < .001$. We find a similar pattern of results for revisit intent, $M = 4.54$ versus 4.30, $t(8266) = −3.792, p < .001$, and the likelihood to recommend, $M = 4.40$ versus 4.05, $t(8266) = −5.760, p < .001$. These results support our contention that the core service experience and the recovery efforts separately drive customers’ postpurchase response, and that the latter can mitigate, but not necessarily nullify, the negative impact of core failure. The results provide support for H1a, H2a, and H3a. Our findings do not support the service recovery paradox (McCollough & Bharadwaj, 1992), and are in line with other research that fails to find support for the effect (e.g., Andreassen, 2001; McCollough et al., 2000).

Participants in our study visited an average of almost five other brands (4.87) during the 30-day period leading up to the first wave. Given their frequent comparisons to other brands and the likelihood that some of these brands may not have not failed them, it is possible that the customers’ tolerance for
failure may be limited in this industry. Their easy access to and experience with other brands may also explain why the revisit intent and likelihood to recommend are lower for the satisfactory recovery group relative to the control group.

Impact of satisfactory recovery. To test H1b, H2b, and H3b, we again use planned comparisons across the three key measures between the satisfactory recovery group and the double deviation group. We find that the satisfactory recovery group exhibited significantly higher satisfaction, $M = 3.71$ versus 2.65, $t(492) = 8.151, p < .001$; revisit intent, $M = 4.30$ versus 3.23, $t(492) = 10.191, p < .001$; and likelihood to recommend, $M = 4.05$ versus 2.95, $t(492) = 10.060, p < .001$, than the double deviation group. These findings support our hypotheses and highlight the benefit of satisfactory recovery (see Table 1). Although the postpurchase response of the satisfactory recovery group did not exceed that of the control group, our findings nevertheless suggest that satisfactory recovery did have a positive impact on the customers.

The response of noncomplainers. In all, 8.1% of customers in our sample experienced a service failure, 57% of whom voiced a complaint. Therefore, only 43% were noncomplainers, which is considerably lower than what has been previously stated in the literature (Stephens & Gwinner, 1998). We used planned comparison t-tests to test H4a, H5a, and H6a using the three measures for the noncomplainers versus the satisfactory recovery group. The satisfactory recovery group exhibited higher satisfaction, $M = 3.71$ versus 3.09, $t(492) = -5.339, p < .001$; revisit intent, $M = 4.30$ versus 3.49, $t(492) = -9.017, p < .001$; and likelihood to recommend, $M = 4.05$ versus 3.18, $t(492) = -9.004, p < .001$, than the noncomplainer group. In other words, even though both groups experienced failure, a satisfactory recovery effort improved customers’ postpurchase response.

Finally, we used planned comparisons for the three measures for the noncomplainer group versus the double deviation group to test H4b, H5b, and H6b. The noncomplainers exhibited significantly higher satisfaction, $M = 3.09$ versus 2.65, $t(521) = -3.864, p < .001$; revisit intent, $M = 3.49$ versus 3.23, $t(521) = -2.382, p < .05$; and likelihood to recommend, $M = 3.19$ versus 2.95, $t(521) = 2.146, p < .05$, than the double deviation group. These results support H4b, H5b, and H6b and suggest that customers who did not complain had higher satisfaction and behavioral intentions than those who experienced a service failure but a failed recovery (see Table 2). The additional effort taken to complain magnified customers’ negative experience (Johnston & Fern, 1999) leading to the lower responses. These findings highlight the potential downside of encouraging customers to complain if they are dissatisfied. They show that if the recovery effort is subpar, then the brand may be worse off than if customers had not complained at all.

Impact on Future Behavior

As mentioned earlier, we asked participants in our longitudinal tracking study about their brand patronage behavior over 12 quarters following the initial period where we had identified service failures and recovery efforts for each participant’s focal brand. We refer to the initial wave as T0. In subsequent waves, each participant provided information about the same brand about which he or she had provided data in the initial wave. We measured brand patronage by asking the participants in every wave if they had visited their focal brand in the past 30 days. Not every respondent participated in every wave of the study. To be included in the analysis, a respondent had to participate in a minimum of seven waves. We used linear interpolation based on visitation prior to and after the missing value to replace the missing values (Kumar, Jones, Venkatesan, & Leone, 2011). Nine customers experienced a second failure from the same brand as in the first wave during the tracking period and were removed from the sample.

Model development and estimation. Given the context, our investigation faces two challenges. First, multiple factors presumably affect a customer’s visits to a restaurant. Because our data are limited to the variables in the multiwave survey instruments, it is possible that random and systematic errors due to unobserved variables may bias our parameter estimates. Second, unobserved heterogeneity among customers may result in some variance in their visitation patterns. To overcome these two data challenges, we use a finite mixture model of brand visits, and examine the differential effects of service failure and recovery outcomes.

We start with a logit model for a visit by customer $i$ in period $t$ and assume that the choice to visit or not depends on whether the consumer did or did not experience a service failure, the brand that failed, as well as his or her individual characteristics:

$$\frac{v_{it}}{1 - v_{it}} = \sum (\text{Service Failure Category})_i \sum (\text{Brand Information}) + \sum (\text{Customer Demographics}) + e,$$

(1)

Table 1. Attitudinal Measures by Service Recovery Group.

<table>
<thead>
<tr>
<th>Customer Group</th>
<th>Satisfaction</th>
<th>Revisit Intent</th>
<th>Likelihood to Recommend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$ (SD)</td>
<td>$M$ (SD)</td>
<td>$M$ (SD)</td>
</tr>
<tr>
<td>No failure</td>
<td>4.43 (0.93)</td>
<td>4.54 (0.72)</td>
<td>4.40 (0.81)</td>
</tr>
<tr>
<td>Failure recovered</td>
<td>3.71 (1.28)</td>
<td>4.30 (0.83)</td>
<td>4.05 (0.93)</td>
</tr>
<tr>
<td>Noncomplainer</td>
<td>3.09 (1.24)</td>
<td>3.49 (1.17)</td>
<td>3.19 (1.20)</td>
</tr>
<tr>
<td>Double deviation</td>
<td>2.65 (1.32)</td>
<td>3.23 (1.27)</td>
<td>2.95 (1.27)</td>
</tr>
</tbody>
</table>
### Table 2.
Logit Model Results.

<table>
<thead>
<tr>
<th>Wave T1</th>
<th>Wave T2</th>
<th>Wave T3</th>
<th>Wave T4</th>
<th>Wave T5</th>
<th>Wave T6</th>
<th>Wave T7</th>
<th>Wave T8</th>
<th>Wave T9</th>
<th>Wave T10</th>
<th>Wave T11</th>
<th>Wave T12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.021</td>
<td>0.496</td>
<td>-0.955</td>
<td>0.497</td>
<td>-2.080</td>
<td>0.542</td>
<td>-2.365</td>
<td>0.556</td>
<td>-2.063</td>
<td>0.555</td>
<td>-3.037</td>
</tr>
<tr>
<td>Control</td>
<td>0.064</td>
<td>0.249</td>
<td>-0.128</td>
<td>0.247</td>
<td>-0.181</td>
<td>0.254</td>
<td>-0.012</td>
<td>0.267</td>
<td>0.117</td>
<td>0.274</td>
<td>0.001</td>
</tr>
<tr>
<td>Noncomplainer</td>
<td>-1.147</td>
<td>0.305</td>
<td>-0.959</td>
<td>0.312</td>
<td>-1.026</td>
<td>0.333</td>
<td>-0.770</td>
<td>0.344</td>
<td>-0.497</td>
<td>0.348</td>
<td>-0.653</td>
</tr>
<tr>
<td>Double deviation</td>
<td>-1.186</td>
<td>0.293</td>
<td>-0.778</td>
<td>0.333</td>
<td>-0.942</td>
<td>0.347</td>
<td>-0.495</td>
<td>0.365</td>
<td>-0.204</td>
<td>0.368</td>
<td>-0.365</td>
</tr>
<tr>
<td>Gender</td>
<td>0.099</td>
<td>0.123</td>
<td>0.064</td>
<td>0.122</td>
<td>0.096</td>
<td>0.129</td>
<td>0.246</td>
<td>0.131</td>
<td>0.295</td>
<td>0.135</td>
<td>0.088</td>
</tr>
<tr>
<td>Age</td>
<td>0.004</td>
<td>0.005</td>
<td>0.006</td>
<td>0.005</td>
<td>0.014</td>
<td>0.006</td>
<td>0.013</td>
<td>0.006</td>
<td>0.009</td>
<td>0.006</td>
<td>0.009</td>
</tr>
<tr>
<td>Education</td>
<td>-0.021</td>
<td>0.047</td>
<td>0.052</td>
<td>0.047</td>
<td>0.138</td>
<td>0.049</td>
<td>0.108</td>
<td>0.050</td>
<td>0.136</td>
<td>0.051</td>
<td>0.083</td>
</tr>
<tr>
<td>Income</td>
<td>0.043</td>
<td>0.027</td>
<td>0.033</td>
<td>0.027</td>
<td>0.040</td>
<td>0.029</td>
<td>0.047</td>
<td>0.029</td>
<td>0.075</td>
<td>0.030</td>
<td>0.072</td>
</tr>
<tr>
<td>Brand 1</td>
<td>0.016</td>
<td>0.274</td>
<td>0.068</td>
<td>0.272</td>
<td>0.369</td>
<td>0.287</td>
<td>0.732</td>
<td>0.296</td>
<td>0.417</td>
<td>0.286</td>
<td>0.526</td>
</tr>
<tr>
<td>Brand 2</td>
<td>-0.044</td>
<td>0.394</td>
<td>0.701</td>
<td>0.389</td>
<td>-0.259</td>
<td>0.443</td>
<td>-0.661</td>
<td>0.466</td>
<td>-0.579</td>
<td>0.467</td>
<td>0.188</td>
</tr>
<tr>
<td>Brand 3</td>
<td>-0.027</td>
<td>0.301</td>
<td>-0.397</td>
<td>0.308</td>
<td>-0.176</td>
<td>0.325</td>
<td>-0.276</td>
<td>0.345</td>
<td>-0.527</td>
<td>0.334</td>
<td>-0.407</td>
</tr>
<tr>
<td>Brand 4</td>
<td>0.135</td>
<td>0.293</td>
<td>0.194</td>
<td>0.289</td>
<td>0.306</td>
<td>0.306</td>
<td>0.339</td>
<td>0.318</td>
<td>-0.045</td>
<td>0.311</td>
<td>0.247</td>
</tr>
<tr>
<td>Brand 5</td>
<td>-0.221</td>
<td>0.286</td>
<td>-0.161</td>
<td>0.288</td>
<td>-0.107</td>
<td>0.307</td>
<td>0.061</td>
<td>0.318</td>
<td>-0.441</td>
<td>0.315</td>
<td>-0.148</td>
</tr>
<tr>
<td>Brand 6</td>
<td>-0.581</td>
<td>0.322</td>
<td>-0.103</td>
<td>0.323</td>
<td>-0.150</td>
<td>0.347</td>
<td>0.042</td>
<td>0.357</td>
<td>-0.295</td>
<td>0.351</td>
<td>-0.177</td>
</tr>
<tr>
<td>Brand 7</td>
<td>0.240</td>
<td>0.288</td>
<td>0.088</td>
<td>0.284</td>
<td>0.200</td>
<td>0.301</td>
<td>0.519</td>
<td>0.309</td>
<td>-0.121</td>
<td>0.305</td>
<td>0.228</td>
</tr>
<tr>
<td>Brand 8</td>
<td>-0.236</td>
<td>0.297</td>
<td>-0.100</td>
<td>0.298</td>
<td>0.109</td>
<td>0.314</td>
<td>0.064</td>
<td>0.329</td>
<td>-0.281</td>
<td>0.322</td>
<td>-0.079</td>
</tr>
<tr>
<td>Brand 9</td>
<td>0.008</td>
<td>0.310</td>
<td>0.162</td>
<td>0.308</td>
<td>0.235</td>
<td>0.325</td>
<td>0.240</td>
<td>0.339</td>
<td>-0.219</td>
<td>0.337</td>
<td>0.275</td>
</tr>
</tbody>
</table>

Note. Bold is significantly different from the reference category at the 95% confidence level. Reference category = satisfactory recovery.

*Italic is significantly different from the reference category at the 90% confidence level.*
where $v_{it}$ is the probability of visit by customer $i$ in time $t$.

We rewrite this equation in a generic logistic model form where $y_{ij}$ is the binary dependent variable (Visit $= 1$, 0) for customer $i$, $X_{ij}$ are the $j = 1, 2, \ldots, J$ predictor variables, and $b$s are the corresponding model parameters:

$$ p(v_{it}) = \frac{1}{1 + e^{-(\sum_{j=1}^{J} x_{ij} b_j + v_{it})}}, \quad (2) $$

where $b_j$ indicates the estimated value of the $j$th coefficient in time $t$ to determine the probability of visits by customer $i$. Then, assuming that visit probability is distributed as a mixture of conditional normal densities, we can specify Equation 2 as

$$ v_{it} = f_{it}(v_{it} | X_{it}, \sigma^2, b_{jt}), \quad (3) $$

or

$$ z(v_{it} | x, \psi) = f_{it}(v_{it} | X_{it}, \sigma^2, b_{jt}), \quad (4) $$

where $v$ is the dependent variable with conditional density $z$, $\psi$ is a vector of all parameters, and $\sigma^2$ is the variance. We assume that the probability of visits $p$ depends on the vector of parameters, $\psi$ and predictor variables $x$. Given that the dependent variable of our analysis is binary, we consider odds (Visit $= 1$) $= p / (1 - p)$, and rewrite Equation 4 as a logit function:

$$ \logit(v_{it}) = \sum_{j=1}^{J} \log z(v_{it} | x_{it}, \psi) = \sum_{j=1}^{J} \log f_{it}(v_{it} | x_{it}, \psi), \quad (5) $$

or

$$ \logit(v_{it}) = \beta_0 + \beta_{Cat1} \times \text{Serv.Cat}_{1it} + \beta_{Cat2} \times \text{Serv.Cat}_{2it} + \beta_{Cat3} \times \text{Serv.Cat}_{3it} + \beta_{Cat4} \times \text{Serv.Cat}_{4it} + \sum_{k=1}^{K} \beta_{k} \times \text{BRAND}_{k} + \beta_{GENDER} \times \text{GENDER}_{it} + \beta_{AGE} \times \text{AGE}_{it} + \beta_{EDUCATION} \times \text{EDUCATION}_{it} + \beta_{INCOME} \times \text{INCOME}_{it} + e_{it}, \quad (6) $$

where Serv.Cat$_{1}$ = 1 when there was no service failure, and 0 otherwise. Similarly, Serv.Cat$_{3}$ = 1 where the customer is a noncomplainer and 0 otherwise, and Serv.Cat$_{4}$ = 1 where there is a failure and failed recovery and 0 otherwise. Service Category 2, where service failure occurred followed by successful recovery is the baseline in the model. BRAND = brand indicator variables ($K$-1 0-1 dummies for $K$ brands, where $K = 10$), GENDER = 1 = male, 0 = female, AGE = age of the customer, EDUCATION = level of education of the customers, and INCOME = income level of customers. We estimate the model separately for each of the 12 quarters, after T0, to examine how brand visitation varied over time and whether this pattern differed across the four service recovery categories. We repeated the entire analysis using Service Failure Category 3 as the baseline to enable additional comparisons among the categories and test the robustness of the findings. We use the estimated coefficients for the three dummy variables Serv.Cat$_{1}$, Serv.Cat$_{3}$, and Serv.Cat$_{4}$ to test the effect of belonging to these groups on brand patronage.

We estimate the posterior class probabilities for each observation and estimate Equation 6 using the posterior probability weights and expectation-maximization (EM) algorithm within the above maximum likelihood framework (Dempster, Laird, & Rubin, 1977). We repeated this process of estimation (E) and maximization (M) to obtain the final results. We used FlexMix library in R to estimate the above equations (see Leisch, 2004 for the estimation process).

### Results

Figure 1 shows the percentage of customers within each of the service recovery outcome groups who visited their focal restaurant during each wave. Both, the data in the figure as well the findings from our visitation model suggest that the pattern of the attitudinal response in T0 did not translate into fewer visits by the satisfactory recovery group than the control group for the first seven waves (see Figure 1 and Table 2). We find that the estimate of the dummy variable corresponding to the control group was not statistically significant in any of these waves. This may be because the core service and the recovery effort is evaluated separately by customers and weighed differently when they reported their attitudinal measures versus their actual visitation behavior. The pattern of results suggest that the successful recovery effort helped dissipate the negative emotions from the failure when customers were making visitation decisions (Pennebaker, 1990), and that the prompted recall of the failure during the survey temporarily raised the salience of the failure in their memory that adversely affected the reported attitudinal scores. In other words, it appears that the positive effect of a successful recovery compensated for the negative effect of the failure when customers were making their visitation decisions in the medium term corresponding to the first seven waves of data collection. None of the brand effects was statistically significant.

However, over the longer term, this compensatory positive effect of successful recovery weakened. Specifically, starting with Wave T8, we find that the coefficient of the control or no-failure group (relative to the satisfactory recovery group as a reference) in our model becomes
positive and statistically significant, and its magnitude generally becomes larger over time (see Table 2). In T8 itself, our model indicates the visitation propensity of the satisfactory recovery group is significantly lower than the control group ($Z = 1.964, p > .05$) but not different from the noncomplainer group ($Z = 0.250, p > .10$). That trend continued through T12 where the satisfactory recovery group is not significantly different than the noncomplainer ($Z = 1.019, p > .10$) and the double deviation groups ($Z = 0.022, p > .10$). The delayed decline may be due to the dissipation of the goodwill that is created during the recovery process.

We estimated our brand visitation model again but used the noncomplainer group as the reference or baseline. Although in T0, the noncomplainer group had higher revisit intention than the double deviation group ($M = 3.49$ vs. $3.23$), it did not translate into a difference in future visitation (see Table 3 where noncomplainer is the reference group in our model), with T8 as an exception. Jamieson and Bass (1989) identify a nonlinear relationship between stated purchase intentions and actual behavior with the probability of purchase increasing significantly at the upper end of the scale. Therefore, it is likely that, given that both groups left the experience dissatisfied, the difference in the low visit intention measures was not large enough to create a difference if future visitation.

Importantly, our longer term visitation results (Figure 1 and Tables 2 and 3) suggest that the pattern of behavioral differences across the four groups changes substantially. Recall that in the earlier periods, the visitation behavior of the successful recovery group was similar to that of the control group. The coefficients for the dummy variable corresponding to the control group were not statistically different from zero in the model where the baseline was the no-failure group. However, we find that over time, the behavior of the three groups that experienced failure converge and separate from that of the no-failure group. Note that the coefficients of the dummy variables corresponding to failed recovery as well as noncomplainers no longer remain statistically significant from T8 onward. This pattern of results suggests that in the longer run, the salience of the complaining behavior as well as the quality of the recovery effort reduces and what matters is whether or not the customer had experienced a failure or not.

**Investing in Service Recovery**

The results from our longer term patronage model suggest that it might be worthwhile for firms to just focus on failure prevention, rather than investing in either the quality of the recovery effort or encouraging dissatisfied customers to complain. However, the results from the short-term behavior show that, on average, the satisfactory recovery group made an additional 3.42 more visits than the double deviation group during the 21 months covering the seven quarters where the visitation behavior of the two groups differed. The average ticket size for the customers in our data is US$33.40 which leads to an additional revenue of US$113.63 from the additional visits. The results show that the medium-term revenue from satisfactory service recovery (vs. failed recovery) is US$116.90 more which may provide an incentive to invest in successful recovery efforts.
Table 3. Logit Model Results.

<table>
<thead>
<tr>
<th>Wave T1</th>
<th>Wave T2</th>
<th>Wave T3</th>
<th>Wave T4</th>
<th>Wave T5</th>
<th>Wave T6</th>
<th>Wave T7</th>
<th>Wave T8</th>
<th>Wave T9</th>
<th>Wave T10</th>
<th>Wave T11</th>
<th>Wave T12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.167</td>
<td>0.484</td>
<td>-1.864</td>
<td>0.492</td>
<td>-3.153</td>
<td>0.542</td>
<td>-3.142</td>
<td>0.550</td>
<td>-2.862</td>
<td>0.550</td>
<td>-3.339</td>
</tr>
<tr>
<td>Control</td>
<td>1.211</td>
<td>0.203</td>
<td>0.781</td>
<td>0.216</td>
<td>0.492</td>
<td>0.240</td>
<td>0.629</td>
<td>0.242</td>
<td>0.614</td>
<td>0.242</td>
<td>0.654</td>
</tr>
<tr>
<td>Satisfied Recovery</td>
<td>1.147</td>
<td>0.305</td>
<td>0.909</td>
<td>0.312</td>
<td>1.026</td>
<td>0.333</td>
<td>0.770</td>
<td>0.327</td>
<td>0.642</td>
<td>0.344</td>
<td>0.653</td>
</tr>
<tr>
<td>Double Deviation</td>
<td>-0.039</td>
<td>0.347</td>
<td>0.099</td>
<td>0.123</td>
<td>0.096</td>
<td>0.129</td>
<td>0.246</td>
<td>0.131</td>
<td>0.295</td>
<td>0.135</td>
<td>0.088</td>
</tr>
<tr>
<td>Gender</td>
<td>0.099</td>
<td>0.132</td>
<td>0.064</td>
<td>0.123</td>
<td>0.096</td>
<td>0.129</td>
<td>0.246</td>
<td>0.131</td>
<td>0.295</td>
<td>0.135</td>
<td>0.088</td>
</tr>
<tr>
<td>Age</td>
<td>0.004</td>
<td>0.005</td>
<td>0.006</td>
<td>0.005</td>
<td>0.014</td>
<td>0.006</td>
<td>0.013</td>
<td>0.006</td>
<td>0.009</td>
<td>0.006</td>
<td>0.011</td>
</tr>
<tr>
<td>Education</td>
<td>-0.021</td>
<td>0.047</td>
<td>0.052</td>
<td>0.047</td>
<td>0.138</td>
<td>0.049</td>
<td>0.108</td>
<td>0.050</td>
<td>0.136</td>
<td>0.051</td>
<td>0.083</td>
</tr>
<tr>
<td>Income</td>
<td>0.043</td>
<td>0.027</td>
<td>0.033</td>
<td>0.027</td>
<td>0.040</td>
<td>0.029</td>
<td>0.047</td>
<td>0.029</td>
<td>0.075</td>
<td>0.030</td>
<td>0.071</td>
</tr>
<tr>
<td>Brand 1</td>
<td>0.016</td>
<td>0.274</td>
<td>0.068</td>
<td>0.272</td>
<td>0.369</td>
<td>0.287</td>
<td>0.732</td>
<td>0.296</td>
<td>0.417</td>
<td>0.286</td>
<td>0.526</td>
</tr>
<tr>
<td>Brand 2</td>
<td>-0.044</td>
<td>0.394</td>
<td>0.701</td>
<td>0.389</td>
<td>-0.259</td>
<td>0.443</td>
<td>-0.061</td>
<td>0.448</td>
<td>-0.579</td>
<td>0.467</td>
<td>0.188</td>
</tr>
<tr>
<td>Brand 3</td>
<td>-0.026</td>
<td>0.301</td>
<td>-0.397</td>
<td>0.308</td>
<td>-0.176</td>
<td>0.325</td>
<td>-0.276</td>
<td>0.345</td>
<td>-0.527</td>
<td>0.334</td>
<td>-0.407</td>
</tr>
<tr>
<td>Brand 4</td>
<td>0.135</td>
<td>0.293</td>
<td>0.194</td>
<td>0.289</td>
<td>0.306</td>
<td>0.306</td>
<td>0.339</td>
<td>0.318</td>
<td>-0.045</td>
<td>0.311</td>
<td>0.247</td>
</tr>
<tr>
<td>Brand 5</td>
<td>-0.221</td>
<td>0.286</td>
<td>-0.161</td>
<td>0.288</td>
<td>-0.107</td>
<td>0.307</td>
<td>0.307</td>
<td>0.318</td>
<td>-0.544</td>
<td>0.315</td>
<td>-0.148</td>
</tr>
<tr>
<td>Brand 6</td>
<td>-0.581</td>
<td>0.322</td>
<td>-0.103</td>
<td>0.223</td>
<td>-0.150</td>
<td>0.347</td>
<td>0.042</td>
<td>0.357</td>
<td>-0.295</td>
<td>0.351</td>
<td>-0.177</td>
</tr>
<tr>
<td>Brand 7</td>
<td>0.240</td>
<td>0.288</td>
<td>0.088</td>
<td>0.284</td>
<td>0.200</td>
<td>0.301</td>
<td>0.519</td>
<td>0.309</td>
<td>-0.121</td>
<td>0.305</td>
<td>0.228</td>
</tr>
<tr>
<td>Brand 8</td>
<td>-0.236</td>
<td>0.297</td>
<td>-0.100</td>
<td>0.298</td>
<td>0.019</td>
<td>0.314</td>
<td>0.064</td>
<td>0.329</td>
<td>-0.281</td>
<td>0.322</td>
<td>-0.079</td>
</tr>
<tr>
<td>Brand 9</td>
<td>0.008</td>
<td>0.310</td>
<td>0.162</td>
<td>0.308</td>
<td>0.235</td>
<td>0.325</td>
<td>0.240</td>
<td>0.339</td>
<td>-0.219</td>
<td>0.337</td>
<td>0.275</td>
</tr>
</tbody>
</table>

Note: Bold is significantly different from the reference category at the 95% confidence level. Reference category = noncomplainer.
*Italic is significantly different from the reference category at the 90% confidence level.
Similarly, the difference between the satisfactory recovery group relative to the noncomplainers over the same course of 21 months was an average of 3.62 total visits that translates into a potential revenue gain of US$120.75 per customer. This finding suggests that it might be beneficial for service firms to invest in encouraging customers to voice their complaints to avail the opportunity to successfully recover these customers from failures in the core service. This approach is especially important because we find that in our data, the noncomplainers were the largest group among all those who experienced a failure. These results only partially support H7a, H7b, and H7c.

These data demonstrate the immediate and strong negative reaction by customers who have an experience that ends in dissatisfaction (noncomplier and double deviation). These customers patronized about five brands during the same time period. Therefore, switching costs in this industry tend to be low (Burnham, Frels, & Mahajan, 2003) do not exist. The customers simply avoid the risk of experiencing another service failure at the same restaurant may opt for safer alternatives in their brand portfolio.

**General Discussion**

In this research, we examined the impact of service failure and recovery on customers’ attitudinal and behavioral response toward a restaurant brand. In contrast to previous research on service failure, we adopted a more comprehensive approach and accounted for the response of not only complainers but also noncomplainers as well as those who never experienced failure. Furthermore, we examined changes in the behavior of customers across the four categories over time to provide more insights into the short- and long-term benefits of initial service quality and service recovery. Overall, we find that while recovery efforts might be beneficial in the short and medium term, the long-term behavior of the market may be largely dependent on whether or not a service firm failed at delivering its core service rather than whether the customers’ complaining behavior or the quality of the firm’s recovery efforts. Nevertheless, our data suggest that the financial returns from encouraging dissatisfied customers to complain as well as from improving the quality of the recovery effort are likely to be positive.

**Theoretical Implications**

Our findings have implications for several streams of literature, including service recovery, complaining behavior, and postfailure attitude-behavior linkages. To begin with, we provide new and comprehensive insights into the service recovery paradox (McCollough & Bharadwaj, 1992). Although previous research on the issue is somewhat focused on the immediate aftermath of a failure, and tends to use only attitudinal measures, our findings suggest that a more complete treatment of the phenomenon should involve an analysis of medium- and long-term brand patronage behavior as well. We find that while successful recovery may adequately compensate for failure in the short run, the effect of the quality of the recovery effort fades away over time. Specifically, in the medium run, the behavioral response of those who are successfully recovered may resemble that of those who never experienced failure. However, in the long run, the responses of those who experienced failure appear indistinguishable and independent of their complaining behavior or recovery efforts. Our data suggest that even the service recovery paradox might hold in the short run, it is unlikely to hold over time.

These findings are consistent with the assumption that, because of emotional dissipation, the goodwill created during the recovery might disappear over time (Hart, Heskett, & Sasser, 1990). However, the unfavorable effect of failure itself is likely to endure and adversely affect behavior over a longer time horizon. This pattern of results is consistent with the “duality of service quality” and our assumption that customers are likely to maintain a distinction between the unconditional core service quality and the conditional recovery effort. The effects of these quality perceptions are unlikely to be unconditionally addictive. We find that, in the medium term, customer’s own complaining behavior and the quality of the firm’s response are likely to influence brand patronage. However, the incremental effects of these factors are likely to dissipate over time, and the success versus failure at delivering core quality is likely to drive long-term patronage behavior.

Our findings also add to the complaint management literature which has generally downplayed the role of noncomplainers. We show that these silent customers do exhibit a negative shift in patronage behavior in the short run and this drift persists in the long term. Our initial postexperience attitudinal results are consistent with previous research (Voorhees et al., 2006) in that we find that the double deviation group is less satisfied, and less likely to revisit or recommend than the noncomplainer group. However, we did not see a difference in future visitation patterns between the two groups. This may be because the noncomplainer group switched instead of expending the effort to seek recovery. This avoidance mitigated the negative attitudinal perceptions associated with a second failure but the decision to leave the brand was already made. Given the very low switching costs and large number of available, competing brands, switching patronage is easy.

Taken together, our results from the long-term impact of the various service failure and recovery outcomes suggest that we may be need fewer customer classifications than envisaged in service recovery outcome matrix (Oliver, 1987). Although the matrix is an intuitive classification of the possible failure and recovery outcomes, our data suggest that long-term customer response can be grouped into
just two rather than four categories. However, in the medium term, which corresponds to seven quarters within our context, we again find that participants could be classified into two, albeit different groups based on the brand patronage behavior. Specifically, one group consisted of those who were satisfied and included those who did not experience failure or were recovered successfully. The other group consisted of dissatisfied customers and included noncomplainers and those in the double deviation group. However, we find that the effects of the complaining behavior and the recovery process dissipate over time and the participants can be divided into just the no-failure and failure group, with the latter consisting of the satisfactory recovery, noncomplainer, and double deviation groups.

**Managerial Implications**

Although managers recognize the potential long-term benefits of satisfactory service recovery, few know the return on their recovery investment. Our findings point to a somewhat paradoxical result that successful recovery has a positive effect in the medium term but not necessarily in the long term. Nevertheless, firms might be able to recover their investments in recovery. For example, the restaurant chains in our study attracted about 225,000 customers annually per restaurant and the overall average failure rate was 8.1% or about 18,225 failures per year, per restaurant. Even a small expenditure of US$5 to recover each failure amounts to over US$90,000 per restaurant or over US$45 million a year for a 500 restaurant chain. Our findings show that satisfactory recovery is worth 3.4 (double deviation) to 3.6 times the average customer expenditure during the initial failure over a 2-year period. This suggests that firms can invest substantial resources in service recovery and achieve a high return on such investments.

In addition, our results demonstrate the size and cost of service failures when customers choose not to complain. Given the loss of revenue for each noncomplainer, resources should be directed toward developing proactive procedures that identify and address undisclosed service failures. For example, the internal research at one firm in the restaurant industry shows that satisfaction declines considerably when the food is delivered more than 15 min after the order was placed. In such situations, instead of waiting for customers to complain, servers can proactively “compensate” customers for the delay and reduce the number noncomplainers.

The double deviation group also visits significantly less than the satisfactory recovery group. Given the considerable revenues gained by saving such customers, monetary and nonmonetary efforts should be made to insure that the recovery is satisfactory. It is unlikely that a firm will be able to recover 100% of its failures. In fact, our study indicated that only 46% of the customers who complained were satisfied with their recovery which is consistent with the Best and Andreasen (1977) study. But with the knowledge of the return on the investment, firms can commit more resources that could significantly enhance the success rate.

**Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, or publication of this article.

**Funding**

The author(s) received no financial support for the research, authorship, or publication of this article.

**References**


Author Biographies

Tim Norvell is an assistant professor of Marketing in the Love School of Business at Elon University. His research interests include: services marketing, service personnel practices, customer satisfaction and loyalty measurement.

Piyush Kumar is an associate professor of Marketing in the Terry College of Business at the University of Georgia. His research interests include: service strategy, marketing analytics, social media, brand strategy, and internet models.

Mayukh Dass is the J.B. Hoskins professor of Marketing in the Rawls College of Business at Texas Tech University. His research interests include: valuation issues in dynamic economies, analytical approaches to enhance brands, and analyzing and modeling art markets.