

MAARTEN J. GIJSENBERG, HARALD J. VAN HEERDE,
and PETER C. VERHOEF*

Service providers sometimes face mass service failures. These problems occur across service industries, ranging from severe Internet outages to major delays for airlines or trains. The literature has not yet addressed the following key question: How do service crises affect perceived service quality (PSQ) over time? To answer this question, the authors introduce a Double-Asymmetric Structural Vector Autoregressive model. It captures not only the short- and long-term effects of objective service performance on PSQ but also the differential effects of service crises versus service restoration. The authors analyze a unique data set from a major European railway company, spanning seven years of monthly observations. During this period, severe winter weather caused dramatic service crises. The authors find that performance losses loom larger than gains in the short run and also have permanent negative effects on PSQ in the long run. Consequently, a crisis followed by a restoration will result in a net negative long-term effect on PSQ. The impact of a crisis also depends on the prior trend in objective service performance.

Keywords: service crises, service quality, time-series models, prospect theory

Losses Loom *Longer* Than Gains: Modeling the Impact of Service Crises on Perceived Service Quality over Time

*Maarten J. Gijzenberg is Associate Professor of Marketing, Department of Marketing, University of Groningen (e-mail: M.J.Gijzenberg@rug.nl). Harald J. van Heerde is Research Professor of Marketing, Massey Business School, Massey University, and Extramural Fellow at CentER, Tilburg University (e-mail: heerde@massey.ac.nz). Peter C. Verhoef is Professor of Marketing, Department of Marketing, University of Groningen (e-mail: P.C.Verhoef@rug.nl). The authors thank Ruth Bolton, Marnik Dekimpe, and Jenny van Doorn for their valuable comments and suggestions on this article, as well as the *JMR* review team and seminar participants at the Frankfurt School of Management, the University of Münster, the BI Norwegian Business School, the IDC Arison Business School, the 2014 AMA ECMI EMAC Marketing & Innovation Symposium at Erasmus University, the Marketing Science Conference 2014 at Emory University, and the Marketing Dynamics Conference 2014 in Las Vegas. They also thank a major European train service provider for sharing the data, which they used to advise this company on how to improve customer satisfaction levels. Maarten J. Gijzenberg gratefully acknowledges financial support from the Marie Curie program of the European Commission under grant CIG 321372. This article was partly written while the first author was a visiting scholar at Monash University. A. Parasuraman served as associate editor for this article.

Service providers sometimes face extreme service failures that have a profound impact on their customers. For example, in October 2011, BlackBerry owners around the globe were confronted with severe and enduring Internet connection problems. Consequently, these customers could not access Internet or e-mail using their BlackBerries for several days (Connors, Dummett, and Lawton 2011; Gar-side 2011). Mike Lazaridis, founder of Research in Motion Limited (RIM, BlackBerry's parent company), appeared in a YouTube video stating, "I apologize for the service outages this week, we have let many of you down" (Potter 2011), and RIM cochief Jim Balsillie promised, "We worked 12 years since the launch of the BlackBerry to win the trust of our 70 million customers and we are going to fully commit to win that trust back to 100%" (Connors and Worthen 2011). Nevertheless, customers began looking for alternatives, resulting in a strong decline of RIM's market share, profits, and stock price (Benoit 2013; Connors 2012), a

devolution that was further fueled by a Europe-wide outage in September 2012 (Vitorovich and Hodgson 2012). These issues are not unique to the technology field, however; severe service problems also occur in other sectors, such as extensive delays for airlines or trains (e.g., Goglia 2014; Huckman, Pisano, and Fuller 2007; Van Doorn and Verhoef 2008; Wangenheim and Bayón 2007).

Objective service performance (OSP)¹ is the extent to which a company succeeds in delivering the service it is promising. It is an important determinant of perceived service quality (PSQ) and customer satisfaction (Gupta and Zeithaml 2006). What is much less well-known is how major failures in OSP affect PSQ in the short run and the long run, and we address this question in this study.

Service failures have received some attention in the service recovery literature. However, this literature has mainly focused on how firms attempt to address service failures and how this affects PSQ and customer satisfaction (e.g., Smith, Bolton, and Wagner 1999). In this prior research, the unit of analysis is the individual customer, who is individually confronted with a service failure. As the previous examples show, in a service crisis, extreme service failures occur for all customers at the same time (e.g., all BlackBerry users were confronted with a failing Internet connection). This not only directly affects customers but also becomes an important news item in the media for brands such as BlackBerry, especially when the service crisis is severe and lasts for a long time.

A service crisis thus involves a severe and sustained drop in OSP that affects many customers at the same time. After such a crisis, firms will aim to return to a normal OSP, hoping that they can solve the resulting PSQ problems. The question is whether this can actually be achieved and whether firms can reach, in the long run, the precrisis PSQ levels through a recovery in OSP. Perceived service quality levels, as implied by the performance-expectation paradigm, are influenced by the (dis)confirmation of expected service performance (Bolton and Drew 1991; Oliver 1977, 1980; Rust et al. 1999; Szymanski and Henard 2001). In line with prospect theory (Kahneman and Tversky 1979), negative experiences can have a particularly strong impact on perceived (service) quality and customer satisfaction (e.g., Anderson and Sullivan 1993; Rust et al. 1999). There is, however, a lack of studies that consider the short- and long-term consequences of service crises on PSQ, though several studies have analyzed longitudinal customer satisfaction and service quality data (e.g., Bolton and Drew 1991; Mittal, Kumar, and Tsiros 1999; Van Doorn and Verhoef 2008). Although Rego, Morgan, and Fornell (2013) analyze the development of aggregate satisfaction scores over time and report strong lagged effects, they do not examine the dynamic effects of OSP on PSQ.

A key related question is whether the trend in OSP plays a role in the long-term effects of a service crisis. An upward trend of improving OSP may put consumers in a very different mindset than a downward trend. Because consumer expectations are built on prior performance evolutions (Hansen and Danaher 1999), the effect of a (positive or

negative) shock in OSP on PSQ may depend on the prior trend in performance.

This study investigates the impact of service crises and subsequent restoration on PSQ over time. We focus on PSQ because it is one of the most important metrics in services and because it relates directly to customer satisfaction, customer loyalty, and firm value (Gupta and Zeithaml 2006; Rust, Moorman, and Dickson 2002; Rust, Zahorik, and Keiningham 1995; Zeithaml, Berry, and Parasuraman 1996). We specifically address the following research questions: (1) What are the short- and long-term effects of OSP changes on PSQ? (2) Do losses in OSP not only loom larger than gains (e.g., Kahneman and Tversky 1979) but also loom longer? and (3) Do these effects depend on the trend in OSP?

To address these questions, we develop a Double-Asymmetric Structural Vector Autoregressive (DASVAR) model, capitalizing on recent developments in the time-series literature. This model captures not only the short- and long-term effects of OSP on PSQ but also the differential effects of service crises versus service restoration (first asymmetry). It also has different lags across equations (second asymmetry). The model offers two benefits that existing (structural) vector autoregressive ([S]VAR) models do not offer. First, it allows for the possibility that a negative shock (i.e., a service crisis) followed by a same-size positive shock (i.e., restoration to the same precrisis performance level) leads to a net long-term loss (or gain) in PSQ. Second, unlike VAR models, the effect of a one-time shock (e.g., an OSP drop) is allowed to depend on the trend in OSP before the shock.

We analyze a unique data set from a major European railway company, spanning nearly seven years of monthly OSP and PSQ observations. During this period, severe winter weather caused dramatic drops in OSP measured by the number of successful connections. This setting offers ample opportunity to study the dynamic impact of service crises on PSQ over time.

We contribute to the literature in several ways. To our knowledge, this is the first study that considers the potentially asymmetric impact of service crises on PSQ over time. We provide insights on this issue by introducing the DASVAR model to the literature. This model enables us to confirm the well-known assertion that losses loom *larger* than gains. Moreover, we also show that losses loom *longer* than gains, because short-term losses lead to permanently lower long-term PSQ, whereas short-term gains do not have persistent effects. As such, we show that a crisis followed by a restoration will result in a net negative long-term effect on PSQ. In addition, we provide evidence that the nature and strength of these effects depends on the *trend* of OSP development—upward, downward, or stable.

The article is organized as follows. We begin with an overview of the literature on the effects of changes in OSP on PSQ and customer satisfaction. Next, we present our data and the initial analyses of asymmetries in the evolution of PSQ. We then present the DASVAR model to assess the impact of OSP on PSQ. A discussion of the results of the analyses follows, with special attention paid to the role of the trend in OSP development. We conclude with a discussion of the insights of our work.

¹We use the term “objective service performance” as a synonym for “operational service performance.”

RELEVANT LITERATURE

This study has its roots in the product-harm crisis and service-satisfaction literature. Although PSQ is our main outcome measure, it is strongly linked to, and a very close antecedent of, customer satisfaction (e.g., Bitner and Hubbert 1994). Because these constructs are so connected, we discuss literature on the effects of OSP on both PSQ and customer satisfaction. Table 1 shows how the current research fits in the literature.

Product-harm crises are defined as cases in which products are found to be defective, causing harm to their users and often leading to costly product recalls (Ahluwalia, Burnkrant, and Unnava 2000). Product-harm crises can potentially inflict serious damage to firms. Several studies have added to the understanding of how these crises affect firm performance. Rubel, Naik, and Srinivasan (2011) show how firms can prepare for the consequences of potential product-harm crises. Other studies use empirical data to address important issues such as consumer responses to a product recall (Cleeren, Dekimpe, and Helsen 2008; Zhao, Zhao, and Helsen 2011), changes in marketing effectiveness resulting from a product-harm crisis (Van Heerde, Helsen, and Dekimpe 2007), and how the characteristics of the product-harm crisis affect brand and category purchasing behavior (Cleeren, Van Heerde, and Dekimpe 2013). Each of these studies has examined *product-harm crises*—that is, cases in which a physical *good* has a problem (e.g., salmonella poisoning in peanut butter; Van Heerde, Helsen, and Dekimpe 2007). A key distinguishing aspect that differentiates a good from a service is that, for a good, production and consumption are separated in time; in contrast, for many (albeit not all) services, production and consumption occur simultaneously (Huang and Rust 2013; Zeithaml, Bitner, and Gremler 2012). This implies that when a defective good is identified, firms can quickly begin a recall of the good by locating batches in the distribution chain, often preventing any harm to consumers. However, when a service fails, con-

sumers are most often affected instantaneously. This means that a service failure will typically affect all consumers who are using that service at the moment of the failure, whereas a product-harm crisis will typically only affect a very small fraction of the consumer base, if any.

The problem of a service failure is compounded when the failing service is a mass-consumption service. In individual service encounters, a failing service such as an automatic teller machine that is out of order or a suit that is damaged in a dry-cleaning service will annoy the individual customer and lead to a decrease in satisfaction and probably some negative word of mouth. Some critical incidents can even make customers defect to competitors (Smith, Bolton, and Wagner 1999; Van Doorn and Verhoef 2008). The service literature has investigated remedial actions a service provider can offer to overcome a service failure for an individual customer (e.g., Gelbrich and Roschk 2011; Smith, Bolton, and Wagner 1999). However, these isolated cases, if properly contained, are unlikely to affect overall (aggregate) service quality perceptions and satisfaction. This is likely to be very different for a mass-service setting, such as Internet service providers, telecom services, and railway companies. In such a setting, a service failure often affects the entire customer base (e.g., the BlackBerry case), and even if only a portion of customers is affected, there may be negative consequences for the company's market value (Malhotra and Kubowicz Malhotra 2011).

Despite the importance of mass service failures, the marketing and service literature has not addressed the question of how they affect PSQ over time. Obviously, we would expect a massive and immediate drop in PSQ. However, for a firm it is also important to know the long-term developments in PSQ: Is the drop just temporary or does it persist into the future?

Another key aspect of mass service failures is service restoration. If the firm is able to overcome the mass service failure and bring the OSP level back to normal levels, does PSQ return to precrisis levels as well? In specific circumstances, excellent recovery can even lead to higher satisfaction than before the service failure (e.g., Smith and Bolton 1998). However, this is certainly not a general case (Smith and Bolton 1998). Prospect theory predicts that a loss in OSP will have a larger effect on PSQ and customer satisfaction than an OSP gain of the same magnitude (Kahneman and Tversky 1979). The service quality and customer satisfaction literature has reported similar asymmetries when examining the effects of the difference between expected, “normal” performance and delivered performance of service attributes (e.g., waiting times) on PSQ and customer satisfaction (e.g., Antonides, Verhoef, and Van Aalst 2002; Inman, Dyer, and Jia 1997; Rust et al. 1999). Recent work by Knox and Van Oest (2014) also shows that recovery after a customer complaint counters the effect but does not completely offset it. These studies imply that a drop in OSP followed by a full restoration could lead to a long-term decrease in PSQ or, in other words, that losses not only loom *larger* than gains, they also loom *longer* than gains.

The impact of a service crisis on PSQ, however, could also depend on the trend in OSP. For example, Smith and Bolton (1998, p. 77) argue that customers apply “what have you done for me lately” heuristics. Sivakumar, Li, and Dong

Table 1
CLASSIFICATION OF OUR STUDY WITHIN PRODUCT AND SERVICE FAILURE (CRISIS) LITERATURE (SELECTED RESEARCH)

Aggregation Level	Type of Crisis	
	Product-Harm Crises	Service Crises/Failures
Individual level	Consequences for customer attitudes and loyalty (Cleeren, Dekimpe, and Helsen 2008; Cleeren, Van Heerde, and Dekimpe 2013; Dawar and Pillutla 2000; Klein and Dawar 2004)	Service failure consequences for customer satisfaction and customer loyalty (Anderson and Sullivan 1993; Oliver 1977, 1980; Sivakumar, Li, and Dong 2014; Smith, Bolton and Wagner 1999; Van Doorn and Verhoef 2008)
Mass level	Product-harm consequences for sales and firm value (Chen, Ganesan, and Liu 2009; Van Heerde, Helsen, and Dekimpe 2007)	Impact of customer information breaches on market value of firms (Malhotra and Kubowicz Malhotra 2011) This study: Impact of service crises on PSQ over time

(2014) propose that the effects of a service failure on PSQ depend on the development in the reference level (increasing vs. decreasing) of OSP. This reference level usually increases/decreases when OSP increases/decreases, as customers adapt to the new situations (Helson 1964). Drawing on intertemporal choice theory, Hansen and Danaher (1999) suggest that the overall service judgments (i.e., PSQ) after an improvement in OSP should be more positive when there is a positive trend in OSP than when there is a negative trend. Therefore, the trend in OSP may affect the customers' mindsets.

A downward trend may make customers cynical, and so a further drop in performance does not further depress PSQ too much, because customers may expect a continuation of the trend. In addition, prior beliefs not only affect PSQ through the immediate performance expectation but also affect the perceptions of the new OSP experiences themselves (Boulding, Kalra, and Staelin 1999; Rust et al. 1999). Of key interest, then, is what happens when a performance trend is broken—when an upward evolution is unexpectedly interrupted by a service failure or a downward evolution is suddenly interrupted by a positive surprise in OSP. Both of these unexpected changes can increase the variance in OSP, which may have additional negative effects on PSQ (Rust et al. 1999).

Although this research is the first to analyze the asymmetric effects of service crises versus service improvements over time, (negative) shocks and both their long-term and asymmetric effects have received attention in the economics and finance literature, among others. Some of these articles have used (S)VAR models (e.g., Christiano, Eichenbaum, and Evans 1994; Kilian 2009) to investigate long-term effects of, for example, monetary and oil price shocks, but a wide range of other methods has been used as well (e.g., Gürkaynak, Sack, and Swanson 2005; Jermann and Quadrini 2009; Stephens 2001). Gambacorta and Iannotti (2007) and Kilian and Vigfusson (2011) have investigated asymmetric effects in SVAR models.

In summary, this is the first study to focus on the short- and long-term effects of service crises and subsequent recovery on PSQ using empirical data, with a special emphasis on the role of the trend in OSP. The latter is an important contribution, because prior research has only theorized that the effects of service failures on customer satisfaction and PSQ may depend on service history (Sivakumar, Lei, and Dong 2014) or has tested the notion only in experimental settings (e.g., Hansen and Danaher 1999).

DATA AND PRELIMINARY INSIGHTS

Empirical Context

To address our research questions, we use a unique data set composed of monthly data on OSP and PSQ. The data are provided by a major national European railway company that wishes to remain anonymous. They cover nearly seven years, from January 2006 to October 2012. The railway company transports millions of customers per day. For many customers, this is their main transportation mode. During the observation period, the company experienced several service crises caused by severe winter weather unprecedented in the recent past. Heavy snowfall led to

thousands of canceled journeys and severe delays for all customers using the service.

The railway company defines OSP as the percentage of successful connections. That is, for each transfer point (train station), a set of possible connections is determined as combinations of two rides for which the time between them is more than the minimum norm for that transfer point, and a combination of two rides is taken on a daily basis by at least 300 customers on average. A connection is considered successful if the actual time between the arrival of the first ride and the departure of the second is more than the minimum for that transfer point. This service performance measure is an objective measure, which is different from the subjective (often survey-based) service performance measures used by other studies in marketing (e.g., Mittal, Kumar, and Tsiros 1999).

The company also has an alternative measure for OSP: five-minute punctuality (the percentage of rides with a delay of less than five minutes). This measure has a high correlation with the percentage of successful connections: .86 ($p < .01$). Our discussions with experts from the company confirm that, among the two alternative measures, the percentage of successful connections is the critical measure. That is, most customers use the company's transport offering as part of a transport chain, with connections within the company's dense network or with other transport offerings (other train companies, trams, or buses). A missed connection can have severe consequences in terms of lost time and associated customer frustration and is therefore more likely to affect PSQ than, for example, the average number of minutes of delay in a single ride. As such, the percentage of successful connections is also a better reflection of critical service encounters that can result in critical incidents (e.g., Van Doorn and Verhoef 2008), and we therefore adopt it as the focal measure for OSP. Nevertheless, as a robustness check, we also estimate our model with the five-minute punctuality measure as an indicator of OSP and compare the predictive validity of the two measures.

We measured PSQ through the following survey question: "What is your general opinion/judgment about traveling per train?" The respondents answered this question on a ten-point scale (1 = "could not be worse," 2 = "very bad," 3 = "bad," 4 = "very inadequate," 5 = "inadequate," 6 = "sufficient/satisfactory," 7 = "more than sufficient/satisfactory," 8 = "good," 9 = "very good," and 10 = "excellent"). This is the official survey question for measuring service quality judgments, and the railway firm has used it for years. The company has an agreement with the government that the evaluation of its contractually required performance criteria (including customer satisfaction) is based on this question.² We consider the question for PSQ an appropriate measure, given that it reflects an overall judgment of the relative inferiority/superiority of the organization and its services (e.g., Bitner and Hubbert 1994). The question uses the same ten-point scale that has been used for many decades to grade assignments and exams throughout the education system (primary school, secondary school, and tertiary education)

²While the focal company labels this measure as customer satisfaction and interprets it as such, we treat the construct being measured as PSQ for the sake of conceptual rigor and consistency with the extant literature on service assessments.

in the country of interest, which makes it a very familiar scale for the survey respondents. Throughout the education sector, the scale is treated as an interval scale. The question used also mimics an item used in the multi-item scale for overall perceived quality from Cronin, Brady, and Hult (2000), which uses a nine-point Likert-type scoring format ranging from “poor” to “excellent.” Nevertheless, we acknowledge that the question is not the perfect measure for PSQ, which has frequently been measured with multi-item scales (e.g., Bitner and Hubbert 1994; Cronin, Brady, and Hult 2000). However, when using available company data, prior studies have also used single items to measure constructs such as customer satisfaction, which are commonly used in practice to reduce survey length (e.g., Bolton 1998; Van Doorn and Verhoef 2008).

Perceived service quality is measured on the basis of repeated cross-sections (e.g., Dekimpe et al. 1998; Fornell, Rust, and Dekimpe 2010; Srinivasan, Vanhuele, and Pauwels 2010). Specifically, on a monthly basis, a representative sample of more than 6,000 customers is surveyed while riding on the train, covering the totality of the national network. The company applies random quota sampling, with quota prescribing the (approximate) percentage of respondents traveling in first- versus second-class, in rush hour versus off-rush hours, and in specific parts of the country. Employees of the company subsequently go through the trains and randomly select respondents to match the prescribed quota. This approach mitigates concerns about self-selection issues of very (dis)satisfied respondents. The resulting national average PSQ rating across this sample is the PSQ score for that month.

Figure 1 shows the evolution of both OSP (gray line) and PSQ (black line) over time.³ Drops in OSP (indicated with

arrows) are associated with drops in PSQ, and the correlation between both series is quite strong (.70). However, whereas OSP quickly restores after a drop, PSQ takes more time to return to its previous levels. Drops in PSQ are deep and steep, whereas recovery is slow.

Asymmetries in the Evolution of Perceived Service Quality

As a first step in our analyses, we assess whether the evolution of PSQ shows asymmetries by exploring the third-order moment (i.e., the skewness statistic) of the series. If the series in levels show deepness asymmetry, we should find a negative skewness (i.e., fewer but stronger negative deviations). In a similar vein, in the case of steepness asymmetry, the first-differenced series would show negative skewness (e.g., Deleersnyder et al. 2004; Lamey et al. 2007).

To formally test for asymmetries, we use the nonparametric triples test proposed by Randles et al. (1980). The test is based on all possible triples (y_i, y_j, y_k) of observations of a series y_t . A series of T observations would thus comprise $\binom{T}{3}$ triples. Triples are considered left (right) triples when the middle observation is closer to the larger (smaller) observation than to the smaller (larger) observation. If the distribution is symmetric, the number of left triples would equal the number of right triples. When there are relatively more left triples, the distribution is negatively skewed. The formal test statistic of the triples test is given by the following:

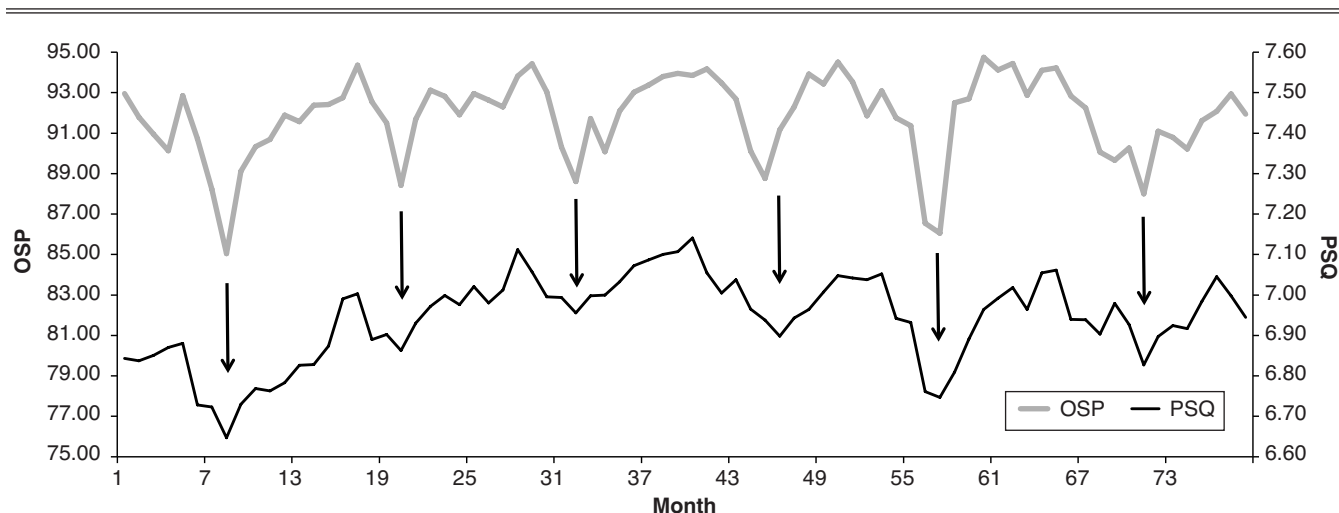
$$(1) \quad \frac{\hat{\eta}}{\sqrt{\frac{\hat{\sigma}_{\hat{\eta}}^2}{T}}},$$

where

$$(2) \quad \hat{\eta} = \frac{[(\text{number of right triples}) - (\text{number of left triples})]}{3 \binom{T}{3}}.$$

For the formal derivation of the test, we refer to Randles et al. (1980).

Figure 1
EVOLUTION OF OBJECTIVE SERVICE PERFORMANCE AND PERCEIVED SERVICE QUALITY OVER TIME



Notes: The arrows point to months with severe drops in OSP.

The test statistic follows an asymptotically normal distribution. If PSQ exhibits (negative) deepness asymmetry, the distribution is expected to be left-skewed and the test statistic would show a negative value for the PSQ level series. Similarly, (negative) steepness asymmetry would be reflected in a negative test statistic value for the first difference of the PSQ series.

As we expected, PSQ shows considerable asymmetries in its evolution over time. The triples test showed a significant negative sign for both the deepness ($-.036$; $p < .05$) and steepness ($-.049$; $p < .05$) statistics. These results indicate that drops in PSQ related to service performance crises are stronger and faster than the recovery of PSQ after such crises.

MODEL FOR THE IMPACT OF OBJECTIVE SERVICE PERFORMANCE ON PERCEIVED SERVICE QUALITY

Model Requirements

To address the research questions of this study, we model the dynamic effects of mass service failures and recoveries on PSQ. We use a two-equation extension of a VAR model for OSP and PSQ. The model needs to address three key challenges that arise in the dynamic interrelationship between OSP and PSQ. First, whereas PSQ is a customer-led metric that is likely to be directly affected by OSP, OSP is a business-led metric, often the outcome of complex processes, and it is unlikely to be immediately affected by PSQ. Thus, we use an SVAR model (e.g., Bernanke 1986; Stock and Watson 2001) to capture the direct effect of OSP on PSQ, but not the other way around. Second, because OSP losses potentially have different short-term and long-term impacts than OSP gains, the model needs to allow for asymmetric effects for losses and gains (Kilian and Vigfusson 2011). Third, another asymmetry is needed because whereas a mass service failure can often be solved in a limited period of time, customer judgments could be stickier (e.g., Baumeister et al. 2001; Skowronski and Carlston 1989). Thus, the model also needs to allow for different (or asymmetric; Ozcicek and McMillin 1999) lag lengths for the OSP and PSQ equations. We address these challenges by estimating a model that we dub a “Double-Asymmetric Structural VAR” (DASVAR) model.

Definition of the Basic SVAR Model

A key first step in the analysis is to determine whether the time series is stationary or shows a unit root (e.g., Dekimpe and Hanssens 1995). To assess the stationarity of the series, we analyzed them with Phillips and Perron’s (1988) test, using an intercept and a trend as exogenous variables. Results show that the OSP series is stationary at the 5% level. However, PSQ has a unit root, and thus the series is evolving. We specify our model with the OSP series in levels and the PSQ series in first differences.

The starting point is an SVAR model for PSQ and OSP:

$$(3a) \quad \Delta PSQ_t = \beta_{1,0} + \sum_{l=1}^L \beta_{1,1,l} OSP_{t-l} + \sum_{l=1}^L \beta_{1,2,l} \Delta PSQ_{t-l} + \varepsilon_{1,t}, \text{ and}$$

$$(3b) \quad OSP_t = \beta_{2,0} + \sum_{l=1}^L \beta_{2,1,l} OSP_{t-l} + \sum_{l=1}^L \beta_{2,2,l} \Delta PSQ_{t-l} + \varepsilon_{2,t}.$$

In this model, ΔPSQ_t represents the change in PSQ, and OSP_t represents the OSP. While a standard VAR model allows for instantaneous, bidirectional effects through the error covariance structure, in an SVAR model theoretical knowledge of the phenomenon is used to specify unidirectional effects. This means that in Equation 3a, OSP will immediately affect PSQ, captured by $\beta_{1,1,0}$. Because it is unlikely that PSQ will have an immediate effect on OSP in the same month, we restrict the reverse effect in Equation 3b to zero.⁴

Because our model is specified with the OSP series in levels and the PSQ series in first differences, any temporary changes in the OSP of the company can have enduring consequences on PSQ, a case of “hysteresis” (see, e.g., Dekimpe and Hanssens 1995; Slotegraaf and Pauwels 2008; Srinivasan et al. 2009). As such, short temporary service failures may lead to permanent negative effects on PSQ. To assess whether this is the case, we will obtain impulse response functions (IRFs) for the estimated model.

Allowing for Asymmetries

We extend this model (Equations 3a and 3b) by allowing for two types of asymmetries. First, we introduce asymmetries between losses and gains. As such, we are the first to introduce asymmetric effects in (S)VAR models in marketing. We also allow for asymmetries in the lag length of the two equations. We next recap why both asymmetries matter.

Prospect theory (Kahneman and Tversky 1979) suggests that human decision making is less influenced by actual levels of certain factors than by changes in these factors. Moreover, such changes will have a stronger impact when they are negative compared with positive. To allow for such asymmetric effects of deteriorations in OSP relative to improvements, we include in our model not only OSP level but also OSP loss, that is, the change in OSP from $t-1$ to t when this change is negative. We formally define Δ^-OSP_{t-l} as follows:

$$(4) \quad \Delta^-OSP_{t-l} = \begin{cases} 0 & \text{if } OSP_{t-l} - OSP_{t-l-1} \geq 0 \\ -1 \times (OSP_{t-l} - OSP_{t-l-1}) & \text{if } OSP_{t-l} - OSP_{t-l-1} < 0 \end{cases}$$

To facilitate parameter interpretation, the (negative) difference in Equation 4 is multiplied by -1 . The change variable thus assumes a zero or a positive value, and it is expected to

⁴A nonrestricted version of our model is the DAVAR (Double-Asymmetric VAR) model that does not impose restrictions on the direction of the immediate effects but models them through the covariance structure of the error terms. We estimated the DAVAR model, and it shows positive contemporaneous correlations of the error terms. However, its fit is worse than for the DASVAR model, both in-sample (mean absolute percentage error [MAPE] of PSQ: .583% vs. .454% for the DASVAR model) and out-of-sample (MAPE of PSQ: .661% versus .489% for the DASVAR model). This is why we retained the DASVAR model.

have a negative parameter estimate, reflecting the negative impact of a drop in OSP on PSQ.

The resulting DASVAR model is

$$(5a) \quad \Delta PSQ_t = \beta_{1,0} + \sum_{l=0}^{L_1} \beta_{1,1,l} \text{OSP}_{t-l} + \sum_{l=1}^{L_1} \beta_{1,2,l} \Delta PSQ_{t-l} \\ + \sum_{l=0}^{L_1} \beta_{1,3,l} \Delta^- \text{OSP}_{t-l} + \varepsilon_{1,t}, \text{ and}$$

$$(5b) \quad \text{OSP}_t = \beta_{2,0} + \sum_{l=1}^{L_2} \beta_{2,1,l} \text{OSP}_{t-l} \\ + \sum_{l=1}^{L_2} \beta_{2,2,l} \Delta PSQ_{t-l} + \varepsilon_{2,t}.$$

Because $\Delta^- \text{OSP}_{t-l}$ is directly inferred from OSP_{t-l} , we do not need to model it separately (Kilian and Vigfusson 2011).⁵ Importantly, unlike a standard (S)VAR model, this model allows for the possibility that a negative shock (OSP drop) followed by the same-size positive shock (e.g., recovery to the same precrisis performance level) leads to a net long-term loss (or gain) in PSQ.⁶

In addition to the asymmetries in the reactions to losses versus gains, we allow for asymmetries in the number of lags to be included in Equations 5a and 5b: L_1 and L_2 are allowed to be different (e.g., Ozcicek and McMillin 1999). Because the SVAR model specifies contemporaneous effects through the structural part of the model, the error terms of the equations are specified as uncorrelated (Kilian and Vigfusson 2011). We determine the optimal number of lags for each equation separately, on the basis of the Bayesian information criterion (BIC), and subsequently combine the optimal numbers of lags in the final DASVAR model.

Deriving the IRFs

Although $\beta_{1,1,0}$ and $\beta_{1,3,0}$ provide insights into the immediate effects of the OSP level and loss, respectively, they do not show the impact of OSP shocks over time. To obtain insights on this, we derive IRFs based on the estimated parameters. However, traditional methods to derive these functions cannot be applied in this case (Kilian and Vigfusson 2011). These traditional methods do not take into account the history preceding the shock, which becomes relevant in the case of asymmetric effects. The Appendix provides a detailed description of the calculation of the IRFs.

Because our model allows for differential effects of OSP deterioration versus improvements, we derive two types of IRFs. The first investigates the effect of an OSP performance shock δ (improvement), whereas the second examines the

effect of a negative OSP shock $-\delta$ (deterioration). We note that the outcome variable is defined in first differences. To obtain insights about the effect on the actual level of PSQ, we analyze the cumulative effects over time of a one-time shock in OSP.

EMPIRICAL RESULTS ON THE IMPACT OF OBJECTIVE SERVICE QUALITY ON PERCEIVED SERVICE QUALITY

Determining the Number of Lags

We first determine the optimal number of lags. Table 2 reports the BIC statistics for different numbers of lags for the two equations. We test specifications with up to ten lags. The results show that one lag is optimal for the OSP equation, whereas two lags are required for the PSQ equation. This is in line with the expectation that changes in OSP are shorter lived than changes in PSQ.

Model Diagnostics

Before presenting the results of our model, we first provide insights on its fit. We present both full-sample and holdout-sample diagnostics for our focal model and three rival specifications, which do not allow for asymmetries in losses versus gains and/or lag structures. In addition, we compare the performance of our model with a rival model using the five-minute punctuality measure to explain PSQ.

Full-sample diagnostics. To judge the quality of our DASVAR model, we compare it with three benchmark SVARs. Benchmark 1 assumes symmetric effects for losses and gains in OSP but with asymmetric lags. Benchmark 2 allows for asymmetric effects but assumes symmetric lags. Benchmark 3 assumes symmetry in effects and lags. We evaluate these models on the basis of their Akaike information criterion (AIC), BIC, and geometric mean of the relative absolute error (GMRAE; Armstrong and Collopy 1992). The fit statistics appear in Figure 2.

The addition of the asymmetric loss effect clearly adds explanatory power, as the AIC and BIC of the PSQ equation improve relative to Benchmark 1 (AIC: -3.474 vs. -3.248 ; BIC: -3.203 vs. -3.067 , respectively). The GMRAE improves (decreases) as well (.474 vs. .486) and is well below the critical value of 1 (Armstrong and Collopy 1992). Allowing for asymmetries in the number of lags relative to Benchmark 2 also noticeably improves model fit, with the DASVAR model outperforming on all statistics (AIC: -3.474 vs. -3.311 ; BIC: -3.203 vs. -3.131 ; GMRAE: .474 vs. .526).

Table 2
LAG SELECTION FOR DASVAR MODEL

Number of Lags	OSP Equation (BIC)	PSQ Equation (BIC)
1	<i>3.845</i>	-3.131
2	3.895	<i>-3.203</i>
3	3.988	-3.065
4	4.056	-3.038
5	4.107	-2.921
6	4.223	-2.888
7	4.312	-2.720
8	4.357	-2.646
9	4.422	-2.600
10	4.403	-2.518

Notes: Italicized cells represent the best (lowest) values and indicate the chosen number of lags for the two equations.

⁵Phillips and Perron's (1988) test, using an intercept and a trend as exogenous variables, shows that $\Delta^- \text{OSP}_{t-l}$ is stationary at the 5% level.

⁶Because theory suggests an asymmetric effect of OSP on PSQ but not the other way around, our model allows for the former asymmetry but not for the latter.

Figure 2
FULL-SAMPLE DIAGNOSTICS FOR PSQ

	Asymmetric Lags	Symmetric Lags
Asymmetric Losses and Gains	Focal Model: DASVAR	Benchmark 2: Asymmetric Effect SVAR
	AIC -3.474	AIC -3.311
	BIC -3.203	BIC -3.131
	GMRAE .474	GMRAE .526
Symmetric Losses and Gains	Benchmark 1: Asymmetric Lag SVAR	Benchmark 3: Symmetric SVAR
	AIC -3.248	AIC -3.242
	BIC -3.067	BIC -3.122
	GMRAE .486	GMRAE .495

The DASVAR model also dominates the fully symmetric Benchmark 3.

Holdout diagnostics. To judge the relative in-sample versus out-of-sample performance of the models, we split the sample in a 60-month estimation sample and an 18-month validation holdout sample. We subsequently reconstruct the level series on the basis of the predictions of the first-differenced models. The resulting in-sample and holdout sample MAPEs for the four models are shown in Figure 3.

In-sample MAPE values are .454% for the focal model, and .473%, .472%, and .474% for the rival models, respectively. Out-of-sample, these values increase to .489% for the focal model and .552%, .543%, and .553% for the benchmark models, respectively. Although the error statistics increase out-of-sample relative to in-sample, the increases are small and confirm the predictive power of the models, the DASVAR model in particular.

Figure 4 shows the out-of-sample forecasts of our focal model and Benchmark 2 with asymmetric lags but without the asymmetric loss effect. This figure confirms that, also

Figure 3
IN-SAMPLE AND OUT-OF-SAMPLE MAPE FOR PSQ

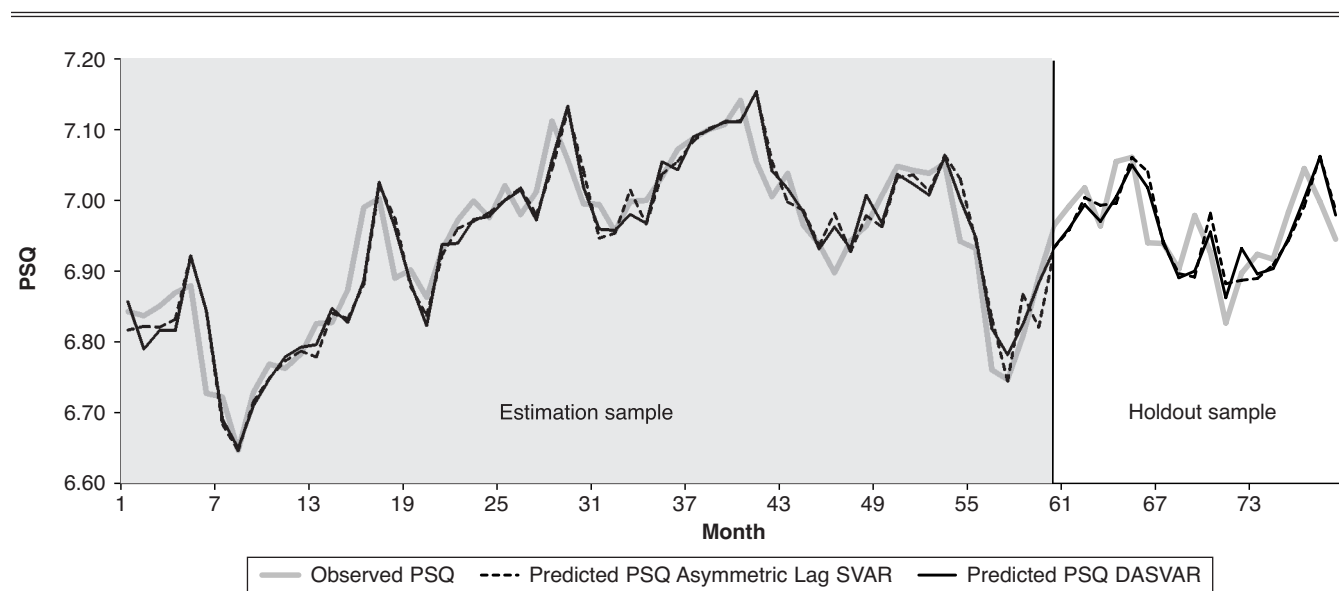
	Asymmetric Lags	Symmetric Lags
Asymmetric Losses and Gains	Focal Model: DASVAR	Benchmark 2: Asymmetric Effect SVAR
	In-sample .454%	In-sample .472%
	Out-of-sample .489%	Out-of-sample .543%
Symmetric Losses and Gains	Benchmark 1: Asymmetric Lag SVAR	Benchmark 3: Symmetric SVAR
	In-sample .473%	In-sample .474%
	Out-of-sample .552%	Out-of-sample .553%

out-of-sample, the inclusion of the asymmetric loss effect enables us to make better predictions of PSQ as an outcome of OSP.

Alternative performance measure. As we argue in the “Data and Preliminary Insights” section, an alternative measure of OSP would be five-minute punctuality. Estimation results of a rival model using this performance measure instead of the connections measure confirm the company experts’ insights that five-minute punctuality is less critical in shaping PSQ than (un)successful connections.⁷ The model shows a considerable drop in R^2 (.475 vs. .568 for the focal model) and serious deterioration in both AIC and BIC (AIC: -3.279 vs. -3.474 for the focal model; BIC: -3.007 vs. -3.203 for the focal model), confirming the superiority of the connections-based OSP measure in explaining PSQ. Summarizing these results, we can conclude that our model shows a good fit, both in-sample and out-of-sample, and outperforms alternative specifications.

⁷A rival model with the five-minute punctuality measure included in the model in addition to the connections measure showed a slightly increased R^2 (.584) but made the penalized fit worse (AIC = -3.359; BIC = -2.906).

Figure 4
IN-SAMPLE AND OUT-OF-SAMPLE PREDICTIVE PERFORMANCE



Substantive Insights

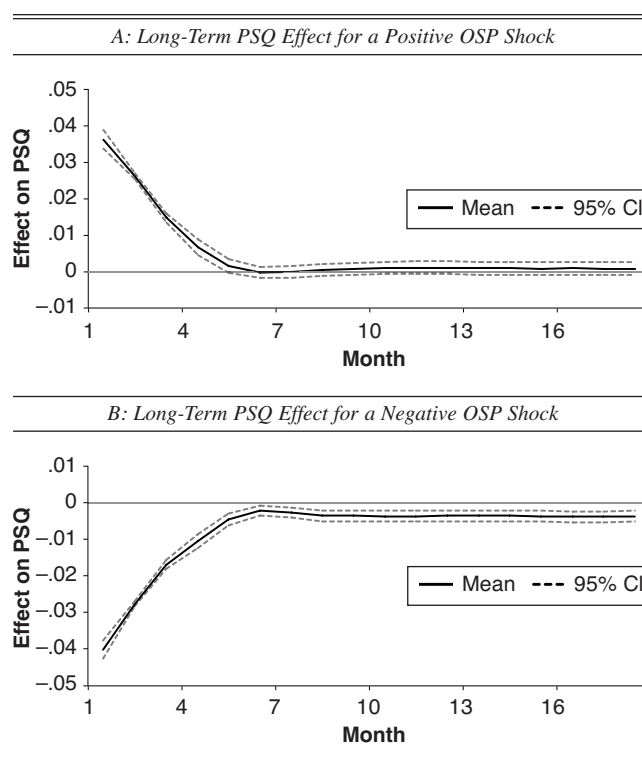
Table 3 reports the estimation results of the focal DAS-VAR model. The subsequent discussion focuses on the PSQ equation because this is the main point of interest of this article. We first discuss the short-term effect of OSP changes on PSQ and then share insights on the long-term effects.

Losses loom larger than gains. Table 3 (Equation 5a) shows that, as we expected, an OSP improvement has an immediate positive effect on PSQ (.011; two-sided $p < .10$ [we use two-sided tests throughout]). For a performance decrease, the immediate effect on PSQ is the combination of the negative of the gain effect (−.011; two-sided $p < .10$) and the extra effect due to the loss term (−.027, two-sided $p < .05$). Thus, an OSP deterioration leads to a stronger PSQ decrease than the increase caused by the equivalent OSP improvement. As such, losses loom larger than gains, at least in the short run. But what happens in the long run?

Losses loom longer than gains. To judge the long-term effects of OSP on PSQ, we derive the impulse response functions according to the steps presented in the Appendix. Because of the asymmetries, we derive two IRFs: one for a positive shock in the OSP of the company and one for a negative shock. The results appear in Figure 5, Panels A and B, respectively. Given that PSQ has a unit root, the dependent variable is defined in first differences. Thus, the cumulative IRFs provide insights on any possibly lasting effects on PSQ of temporary changes in OSP. The solid black lines show the average IRF over all histories in the data set, whereas the dotted lines represent the 95% confidence intervals based on the standard deviation of the mean as obtained in step 5 of the algorithm, described in the Appendix. We apply shocks of one standard deviation to the error term.

Increases in OSP have no significant lasting effects on PSQ. As such, service improvements seem to be considered normal, a simple delivery on promises. No special perma-

Figure 5
CUMULATIVE IMPULSE RESPONSE FUNCTIONS



Notes: CI = confidence interval.

nent rewards are attached to them. Decreases in OSP, on the other hand, show clear permanent negative effects on PSQ. A temporary drop in OSP can thus have long-lasting consequences for the company. Losses thus not only loom larger but also loom longer than gains. Note that although these findings show strong face validity, standard, symmetric, (S)VAR models would not have been able to capture these effects. Only by introducing asymmetric loss effects to the model specification can we uncover these differential dynamics.

Combining the insights on the short- and long-term effects, we conclude the following. Losses loom *larger* than gains in the short run, with OSP drops having a stronger immediate negative impact relative to the immediate positive impact of equally sized OSP gains. In addition, we show that losses loom *longer* than gains because temporary decreases in OSP have lasting negative consequences for PSQ, whereas effects of service improvements are short lived.

THE ROLE OF THE TREND IN OBJECTIVE SERVICE PERFORMANCE

The actual impact of OSP on PSQ may depend on the trend in the OSP of the company (e.g., Inman, Dyer, and Jia 1997; Sivakumar, Li, and Dong 2014). Our model captures this impact by allowing for the lagged effects of both OSP itself and the OSP loss variable on PSQ. Due to the inclusion of both terms, evolutions of sustained OSP gains will have a different impact on current PSQ compared with evolutions of sustained OSP losses or mixed evolutions. We

Table 3
RESULTS FOR DASVAR MODEL

	Parameter Estimate	SE	p-Value
<i>PSQ Equation</i>			
Constant	−.060	.412	.884
OSP _t	.011*	.006	.091
Δ-OSP _t	−.027**	.011	.016
ΔPSQ _{t−1}	.002	.110	.983
OSP _{t−1}	−.000	.007	.951
Δ-OSP _{t−1}	.023**	.010	.033
ΔPSQ _{t−2}	−.308**	.110	.007
OSP _{t−2}	−.010*	.005	.058
Δ-OSP _{t−2}	−.019**	.008	.020
R ²	.568		
AIC	−3.474		
BIC	−3.203		
<i>OSP Equation</i>			
Constant	44.487**	8.475	.000
ΔPSQ _{t−1}	9.063**	3.219	.006
OSP _{t−1}	.515**	.092	.000
R ²	.423		
AIC	3.755		
BIC	3.845		

* $p < .10$.

** $p < .05$.

therefore investigate the extent to which effects are different for different OSP evolutions.

We define three types of OSP trends a company can face when experiencing a sudden and large shock in OSP. The trends are based on the evolution over the four periods before the shock period. Considering a history of four periods (months) enables us to capture true trends and reduces the probability of basing our insights on short-lived positive or negative sequences that would dominate shorter histories. We define the first scenario, “Business as Usual,” as a relatively stable OSP situation, in which there is no clear direction in OSP (up and down or down and up), reflected in two increases and two decreases in OSP over the four-period history. The second scenario, “Sustained Gains,” contains a positive trend in OSP as reflected in at least three of the four preceding periods showing an increase in OSP. The third scenario is labeled “Sustained Losses” and covers a situation of deteriorating OSP with at least three of the four preceding periods showing a decrease in OSP. These definitions enable us to realistically capture the overall trend the company may be experiencing while encountering the shock. Figures 6–8 show the results for these three scenarios.

Business as Usual

In the Business as Usual scenario (Figure 6), we observe the same picture as for the overall case (Figure 5): a sudden positive shock in OSP has no significant long-term effects (Figure 6, Panel A), but a sudden OSP drop will have a last-

ing negative effect on customers’ judgment of the service quality. Short-term decreases in OSP will thus cause permanent losses in PSQ (Figure 6, Panel B).

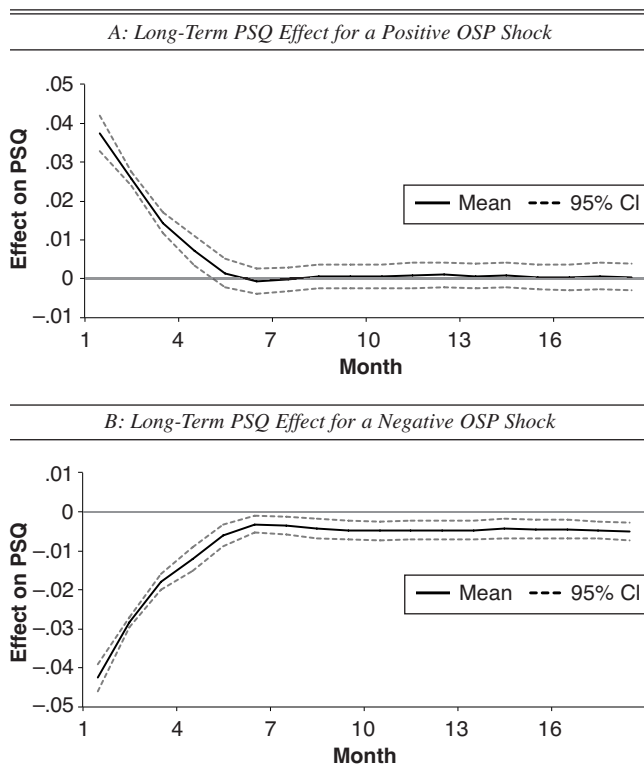
Sustained Gains

When a company experiences an upward trend of sustained gains (Figure 7), the company may have succeeded in creating a considerable amount of goodwill among its customers. Customers know that they can expect good service. If the company continues to improve its service performance, after a period of sustained gains, it may be delivering even more than what its customers expect, a case of positive expectancy disconfirmation potentially leading to customer delight (e.g., Bolton and Drew 1991; Oliver, Rust, and Varki 1997). The positive feelings customers already had toward the company will be reinforced, creating a lasting significant positive effect on PSQ (Figure 7, Panel A), in line with Hansen and Danaher (1999).

However, given a trend of sustained gains in OSP, customers will certainly not expect OSP to deteriorate. A sudden drop in performance (Figure 7, Panel B) will thus constitute a case of extreme negative expectancy disconfirmation, leading to very negative feelings (e.g., Inman, Dyer, and Jia 1997). The feeling that the company is not keeping its implicit promise of continually improving OSP is considerably stronger than just not keeping its promises in a Business as Usual scenario (Skowronski and Carlston 1989). Moreover, a series of increases in OSP followed by a drop

Figure 6

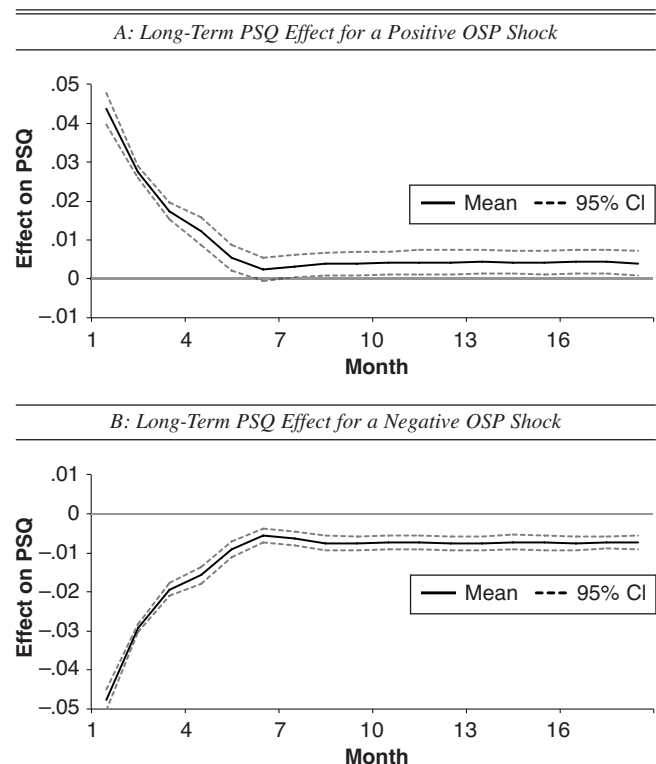
CUMULATIVE IMPULSE RESPONSE FUNCTIONS FOR
BUSINESS AS USUAL: FLUCTUATING OSP



Notes: CI = confidence interval.

Figure 7

CUMULATIVE IMPULSE RESPONSE FUNCTIONS FOR
SUSTAINED GAINS: POSITIVE TREND IN OSP



Notes: CI = confidence interval.

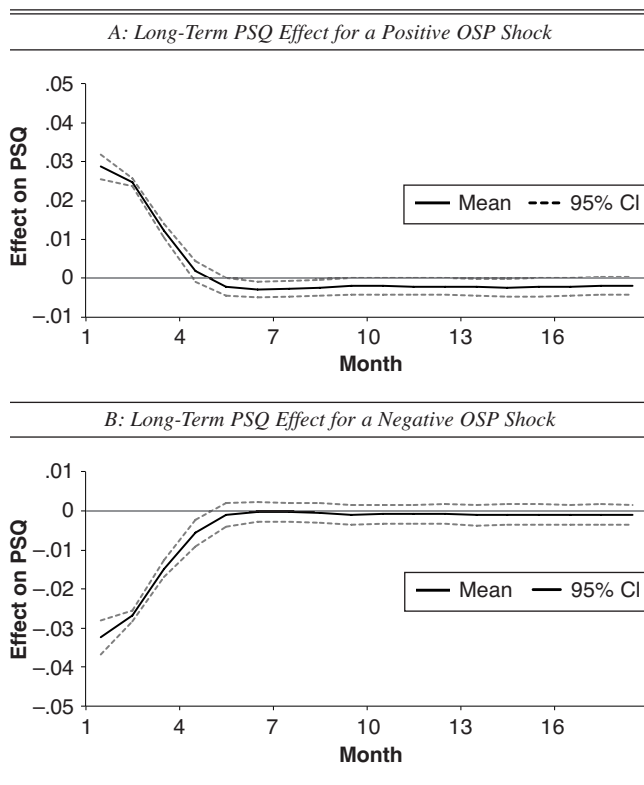
increases the variance, which can adversely affect PSQ (Rust et al. 1999). In line with these arguments, it is not surprising that the lasting negative effect on PSQ is stronger in this case (Figure 7, Panel B) than in the Business as Usual case (Figure 6, Panel B).

Sustained Losses

When the prior trend in OSP is down (Figure 8), we observe some remarkable results. Sudden improvements in OSP have a short positive effect and result in an insignificant ($p > .05$) but negative long-term effect on PSQ (Figure 8, Panel A). Although we also find an insignificant long-term effect in the overall case (Figure 5) and the Business as Usual case (Figure 6, Panel A), the negative sign of the effect is counterintuitive. It is possible that the sudden increase in OSP increases the variance in expected OSP, which in turn increases the perceived risk. Such increased risk is valued negatively and reduces the perceived utility (Rust et al. 1999). Another, more psychological explanation is that negative events have much stronger influence than positive events. Although positive events may improve attitudes, they cannot fully compensate for negative events in the past, and the contrast between events may make these negative events even more salient (Baumeister et al. 2001). We emphasize that the result for this scenario should be interpreted carefully and that more in-depth research is needed.

Figure 8

CUMULATIVE IMPULSE RESPONSE FUNCTIONS FOR
SUSTAINED LOSSES: NEGATIVE TREND IN OSP



Notes: CI = confidence interval.

After a prior history of sustained losses, another drop in OSP will not have lasting negative consequences (Figure 6, Panel B). This result may seem counterintuitive at first, but consider that the prior drops (which, in this case, happen in three or four of the preceding four periods) will have their own permanent negative effects on PSQ because they typically follow a Business as Usual or Sustained Gains scenario. A new drop in OSP in period 5 will then not add much information for the customer. In a sense, the company is living up to its customers' experience-based expectations of bad and even worsening services (e.g., Baumeister et al. 2001), and PSQ could not drop much further.

Comparing the three scenarios, we find that negative shocks in OSP will lead to permanent negative effects on PSQ in cases of stable or sustained increases in OSP. In the latter case, a deterioration will be a difficult wake-up call for customers, showing them that an ever-improving service does not exist. However, in cases of sustained decreasing OSP, yet another drop will not add much information for customers, because they may have become cynical of the actual OSP. Increases in OSP, in contrast, only have a lasting positive effect on PSQ in cases of sustained increases in OSP. In all other situations, OSP improvements will not produce any enduring positive effects on PSQ. As such, it seems that customers negatively evaluate shocks that go against the trend and reverse history. This might be explained by the fact that these changes increase the variance in OSP, creating more risk for customers and resulting in enduring lower levels in PSQ (Rust et al. 1999).

DISCUSSION

Mass service failures are among the worst nightmares for service providers. Whereas the impact on customers in case of product-harm crises can often be mitigated by a timely product recall, there is no escape in case of service failures when the production and consumption of a service coincides. Unlike individual (one-on-one) service failures, mass service failures are massive events with potentially devastating consequences for PSQ. Despite the importance of a deep understanding of the dynamic impact of service failures and recoveries on PSQ, no study to date has offered empirical insights on this impact or developed a suitable methodology to measure it. This study fills this gap in the literature and thus offers several substantive and methodological advances.

From a substantive point of view, we deepen the insights on the dynamics of loss aversion. In line with prospect theory, we find that the short-term (immediate) PSQ effect of an OSP loss looms larger than an OSP gain. Moreover, we test the expectation (and find evidence) that losses loom *longer* than gains, with service failures followed by a recovery leading to a permanent negative effect on PSQ. The finding that negative shocks in OSP have long-lasting effects is in line with extant psychological research showing that negative (life) events have a longer-lasting impact on a person's well-being than a positive event (Baumeister et al. 2001; Sheldon, Ryan, and Reis 1996).

Theoretical notions in the service literature postulate that the effects of service failures may depend on the prior trend in service levels (Sivakumar, Li, and Dong 2014). Ours is the first empirical study to show the important role of OSP evolutions on how an OSP change affects PSQ over time.

We find that a drop in OSP is especially detrimental for PSQ when the prior trend in OSP was upward, a case of extreme negative disconfirmation. A drop in OSP does not further decrease PSQ when it follows a long string of losses in OSP—the damage has already been done by these prior losses.

In the case of an upward trend in OSP, a further increase in OSP has a long-term positive effect on PSQ, in line with customer delight. This extends prior experimental findings of Hansen and Danaher (1999). A somewhat counter-intuitive result is that an OSP improvement after a sustained decline has a negative (but insignificant) long-term effect on PSQ. Additional research on this initial finding is definitely required. This research could use lab studies to further validate this result and uncover the roles of the consumer psychology processes: positive versus negative disconfirmation and quality variance reduction versus quality variance increase.

From a methodological perspective, this study builds on the previous literature by specifying an SVAR model with direct effects of OSP on PSQ, but not vice versa, and allowing for asymmetric (different) lag lengths across equations (e.g., Villanueva, Yoo, and Hanssens 2008). However, the proposed DASVAR model goes beyond these models in that it is the first to allow for asymmetric effects of OSP losses versus gains. We find statistical evidence for the importance of allowing for both types of asymmetries, in terms of not only the significance of individual parameters and in-sample (penalized) fit but also the holdout sample fit. The asymmetries in losses versus gains imply that the IRFs of this model depend on recent prior changes in OSP. This property, which is not shared by a standard (S)VAR model without this type of asymmetry, enables us to condition the IRFs on the basis of recent OSP evolutions.

The DASVAR model is also suited to other contexts in which there are reasons to expect differential effects of losses versus gains and/or differences in lag length across equations. In marketing, researchers are often interested in the drivers of marketing instruments or intermediate customer metrics on firm performance. Our model can help answer questions that involve asymmetries in dynamic effects. For example, is the long-term impact of an increase in advertising expenditures on sales the symmetric opposite of the effect of the same decrease? Is the dynamic effect of an increase in the number of Facebook likes on website traffic the same as the effect of a decrease by the same number?

From a managerial point of view, the insights can be of significant value for service providers. An obvious recommendation is that firms should aim to avoid service crises because they have enduring negative effects on PSQ. If these crises occur, the question is how to overcome a sharp decrease in PSQ. Because OSP losses loom larger and longer than gains, a major implication of our study is that the service recovery needs to do more than overcome the service failure if the objective is to keep the long-term PSQ at a constant level. Thus, a service failure means that the bar for future performance is raised.

A service provider also needs to be mindful about the trend in its recent OSP trajectory. Exceeding prior OSP levels is especially relevant in a case in which recent OSP increases are followed by a sudden mass service failure. A

silver lining is that when the trend is downward to begin with, an additional service deterioration does not further depress PSQ in the long run; however, the downside of this is that the prior drops will have their own negative long-term effects. Scenarios with relatively stable patterns are better for long-term PSQ than strong up–down or down–up scenarios. These insights are key to consider if a firm reports regularly about its recent OSP in press releases or direct-to-customer communication. Emphasizing—without compromising the truth—the steadiness of the OSP (or deemphasizing the instability) may engender favorable perceptions of service quality.

Our model allows service providers not only to quantify the current period drop in PSQ but also to document the long-term (and even permanent) consequences. By combining the estimated short-term and long-term PSQ effects with the well-studied link between PSQ and financial performance, service providers can obtain an estimate of the total financial loss due to the mass service failure. These insights are also very important in lawsuit cases in which an external supplier is at fault for the mass service failure, because they give a qualified monetary estimate for which the external supplier can be sued.

We hope that this study sparks further research on the dynamic effect of service failures on PSQ and other firm performance metrics. It would be worthwhile to further investigate and validate the study's findings with a conventional, multiple-item scale for PSQ and/or customer satisfaction as the dependent measure. Insights could be further enhanced by focusing on the links between OSP, PSQ, satisfaction, sales, and profits (for which we do not have data), similar to the literature on the service–profit chain (e.g., Heskett et al. 1994; Kamakura et al. 2002). In addition, it would be useful to study the moderating influences of crisis characteristics on dynamic effects. For example, does the duration of the service crisis aggravate the negative long-term consequences? Do (more or less) self-inflicted service crises caused by, for example, strikes show different, possibly stronger negative consequences than crises caused by external factors such as the weather?

APPENDIX: DERIVATION OF THE IRFS

In the case of asymmetric effects, the history will determine the value of the asymmetry variables in the model and, thus, the IRFs. To derive IRFs that take these asymmetries into account, we adopt a five-step approach (see Kilian and Vigfusson 2011). The idea is to calculate the IRF for shocks that are applied at different moments in time. For a given starting moment (time = t), the history before t will determine the value of the asymmetry variables, which will then affect the response to a shock.

Step 1

To determine the impact of the history of the variables in our model on the IRFs, we first define a set of histories. Let N = the number of time series observations and L = $\max(\text{lags } \text{OSP}_t, \text{lags } \Delta \text{PSQ}_t)$. For $t = L + 1, \dots, N$, we define a history $\Omega^t = (\text{OSP}_{t-L}, \dots, \text{OSP}_{t-1}; \Delta \text{PSQ}_{t-L}, \dots, \Delta \text{PSQ}_{t-1})$ of consecutive values of OSP_t and ΔPSQ_t . We do so by using a moving window of length L in our data set, leading to $T (= N - L)$ different histories.

Step 2

After defining the histories, the following step to judge the effect of a shock in the OSP on PSQ, relative to a “normal” situation, is the simulation of two situations: one without the shock and one with the shock. For each history Ω^t , we simulate two time paths for OSP_{t+s} and ΔPSQ_{t+s} , with $s = 0, \dots, S$, with S equal to the simulation span of the IRF.

The first time path $\psi^{t,1} = (OSP_{t+0}^1, \dots, OSP_{t+S}^1; \Delta PSQ_{t+0}^1, \dots, \Delta PSQ_{t+S}^1)$ constitutes the benchmark time path in which we draw the realizations of $\varepsilon_{1,t+s}$ and $\varepsilon_{2,t+s}$ independently from their marginal empirical distributions. The second time path $\psi^{t,2} = (OSP_{t+0}^2, \dots, OSP_{t+S}^2; \Delta PSQ_{t+0}^2, \dots, \Delta PSQ_{t+S}^2)$ constitutes the “shock” path. Similar to the benchmark path, the realizations of $\varepsilon_{1,t+s}$ are drawn from its marginal empirical distribution. However, we introduce the shock by setting $\varepsilon_{2,t+0}$ to a prespecified value δ , after which the subsequent $\varepsilon_{2,t+s}$ are again drawn from the marginal empirical distribution of $\varepsilon_{2,t}$.

Step 3

Because we now know the evolutions of PSQ for the situations without and with the shock in OSP, we can determine the extent to which PSQ is different in the two situations. We therefore calculate the difference between the time paths of ΔPSQ_{t+s} : $\Delta^t = (\Delta PSQ_{t+0}^2 - \Delta PSQ_{t+0}^1, \dots, \Delta PSQ_{t+S}^2 - \Delta PSQ_{t+S}^1)$.

Step 4

We repeat the simulation in Steps 2 and 3 10,000 times and average the difference across these repetitions:

$$\bar{\Delta}^t = \left(\overline{\Delta PSQ_{t+0}^2 - \Delta PSQ_{t+0}^1}, \dots, \overline{\Delta PSQ_{t+S}^2 - \Delta PSQ_{t+S}^1} \right).$$

This provides us with the effect of an OSP shock at time t on PSQ, conditional on the specific history Ω^t preceding the shock.

Step 5

Steps 2–4 have provided us with the effect for a specific history preceding the shock. However, because we want to obtain insights across all possible histories generated in Step 1, we repeat Steps 2–4 for the T different histories in our data set and average across these T histories. This provides us with an average effect for a representative set of histories of an OSP shock on PSQ:

$$\bar{\Delta} = \left(\overline{\Delta PSQ_0^2 - \Delta PSQ_0^1}, \dots, \overline{\Delta PSQ_S^2 - \Delta PSQ_S^1} \right).$$

Because our model allows for differential effects of OSP deterioration versus improvements, we apply this five-step approach twice. The first sequence investigates the effect of a positive OSP shock δ , whereas the second sequence covers the effect of a negative OSP shock $-\delta$.

REFERENCES

Ahluwalia, Rohini, Robert E. Burnkrant, and H. Rao Unnava (2000), “Customer Response to Negative Publicity: The Moderating Role of Commitment,” *Journal of Marketing Research*, 37 (May), 203–214.

- Anderson, Eugene W. and Mary W. Sullivan (1993), “The Antecedents and Consequences of Customer Satisfaction for Firms,” *Marketing Science*, 12 (2), 125–43.
- Antonides, Gerrit, Peter C. Verhoef, and Marcel van Aalst (2002), “Customer Perception and Evaluation of Waiting Time: A Field Experiment,” *Journal of Customer Psychology*, 12 (3), 193–202.
- Armstrong, J. Scott and Fred Collopy (1992), “Error Measures for Generalizing About Forecasting Methods: Empirical Comparisons,” *International Journal of Forecasting*, 8 (1), 69–80.
- Baumeister, Roy F., Ellen Bratslavsky, Catrin Finkenauer, and Kathleen D. Vohs (2001), “Bad Is Stronger Than Good,” *Review of General Psychology*, 5 (4), 323–70.
- Benoit, David (2013), “BlackBerry’s Black Friday: Results Plunge,” *The Wall Street Journal*, (September 20), (accessed June 10, 2015), [available at <http://blogs.wsj.com/moneybeat/2013/09/20/blackberrys-black-friday-results-plunge/>].
- Bernanke, Ben S. (1986), “Alternative Explanations of the Money–Income Correlation,” *Carnegie-Rochester Conference Series on Public Policy*, 25 (1), 49–99.
- Bitner, Mary Jo and Amy R. Hubbert (1994), “Encounter Satisfaction Versus Overall Satisfaction and Service Quality,” in *Handbook of Service Quality*, Roland T. Rust and Rich L. Oliver, eds. Thousand Oaks, CA: Sage Publications, 72–94.
- Bolton, Ruth 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.
- and James H. Drew (1991), “A Longitudinal Analysis of the Impact of Service Changes on Customer Attitudes,” *Journal of Marketing*, 55 (January), 1–9.
- Boulding, William, Ajay Kalra, and Richard Staelin (1999), “The Quality Double Whammy,” *Marketing Science*, 18 (4), 463–84.
- Chen, Yubo, Shankar Ganesan, and Yong Liu (2009), “Does a Firm’s Product Recall Strategy Affect Its Financial Value? An Examination of Strategic Alternatives During Product-Harm Crises,” *Journal of Marketing*, 73 (November), 214–26.
- Christiano, Lawrence J., Martin Eichenbaum, and Charles Evans (1994), “The Effects of Monetary Policy Shocks: Some Evidence from the Flow of Funds,” NBER Working Paper No. 4699, National Bureau of Economic Research.
- Cleeren, Kathleen, Marnik G. Dekimpe, and Kristiaan Helsen (2008), “Weathering Product-Harm Crises,” *Journal of the Academy of Marketing Science*, 36 (2), 262–70.
- , Harald J. van Heerde, and Marnik G. Dekimpe (2013), “Rising from the Ashes: How Brands and Categories Can Overcome Product-Harm Crises,” *Journal of Marketing*, 77 (March), 58–77.
- Connors, Will (2012), “BlackBerry Maker in Turmoil,” *The Wall Street Journal*, (March 30), (accessed June 10, 2015), [available at <http://online.wsj.com/news/articles/SB10001424052702303404704577312010085459278>].
- , Ben Dummett, and Christopher Lawton (2011), “For BlackBerry Maker, Crisis Mounts,” *The Wall Street Journal*, (October 13), (accessed March 11, 2014), [available at <http://online.wsj.com/news/articles/SB10001424052970203914304576626451110144140>].
- and Ben Worthen (2011), “RIM Tries to Regain Trust from Its Customers,” *The Wall Street Journal*, (October 14), (accessed June 10, 2015), [available at <http://online.wsj.com/news/articles/SB10001424052970204002304576628003438184440>].
- Cronin, J. Joseph, Jr., Michael K. Brady, and G. Tomas M. Hult (2000), “Assessing the Effects of Quality, Value and Customer Satisfaction on Consumer Behavioral Intentions in Service Environments,” *Journal of Retailing*, 76 (2), 193–218.
- Dawar, Niraj and Madan M. Pillutla (2000), “Impact of Product-Harm Crises on Brand Equity: The Moderating Role of Customer Expectations,” *Journal of Marketing Research*, 37 (May), 215–26.

- Dekimpe, Marnik G. and Dominique M. Hanssens (1995), "The Persistence of Marketing Effects on Sales," *Marketing Science*, 14 (1), 1–21.
- , Linda M. van de Gucht, Dominique M. Hanssens, and Keiko I. Powers (1998), "Long-Run Abstinence After Narcotics Abuse: What Are the Odds?" *Management Science*, 44 (11, Part 1), 1478–92.
- Deleersnyder, Barbara, Marnik G. Dekimpe, Miklos Sarvary, and Philip M. Parker (2004), "Weathering Tight Economic Times: The Sales Evolution of Customer Durables over the Business Cycle," *Quantitative Marketing and Economics*, 2 (4), 347–83.
- Fornell, Claes, Roland T. Rust, and Marnik G. Dekimpe (2010), "The Effect of Customer Satisfaction on Consumer Spending Growth," *Journal of Marketing Research*, 47 (February), 28–35.
- Gambacorta, Leonardo and Simonetta Iannotti (2007), "Are There Asymmetries in the Response of Bank Interest Rates to Monetary Shocks?" *Applied Economics*, 39 (19), 2503–2517.
- Garside, Juliette (2011), "Service Restored, but BlackBerry May Never Repair Its Reputation," *The Guardian*, (October 16), (accessed June 10, 2015), [available at <http://www.theguardian.com/technology/2011/oct/16/service-restored-but-blackberry-may-never-repair-its-reputation>].
- Gelbrich, Katja and Holger Roschk (2011), "A Meta-Analysis of Organizational Complaint Handling and Customer Responses," *Journal of Service Research*, 14 (1), 24–43.
- Goglia, John (2014), "Why Extreme Cold Is Causing Extreme Problems for Airlines," *Forbes*, (January 7), (accessed June 10, 2015), [available at <http://www.forbes.com/sites/johngoglia/2014/01/07/why-extreme-cold-is-causing-extreme-problems-for-the-airlines/>].
- Gupta, Sunil and Valarie Zeithaml (2006), "Customer Metrics and Their Impact on Financial Performance," *Marketing Science*, 25 (6), 718–39.
- Gürkaynak, Refet S., Brian Sack, and Eric Swanson (2005), "The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models," *American Economic Review*, 95 (1), 425–36.
- Hansen, David E. and Peter J. Danaher (1999), "Inconsistent Performance During the Service Encounter: What's a Good Start Worth?" *Journal of Service Research*, 1 (3), 227–35.
- Helson, Harry (1964), *Adaptation Level Theory*. Oxford, UK: Harper & Row.
- Heskett, James L., Thomas O. Jones, Gary W. Loveman, W. Earl Sasser Jr., and Leonard Schlesinger (1994), "Putting the Service-Profit Chain to Work," *Harvard Business Review*, 72 (2), 164–74.
- Huang, Ming-Hui and Roland T. Rust (2013), "IT-Related Service: A Multidisciplinary Perspective," *Journal of Service Research*, 16 (3), 251–58.
- Huckman, Robert S., Gary P. Pisano, and Virginia A. Fuller (2007), "JetBlue Airways: Valentine's Day 2007," Harvard Business School Case Studies 608001-PDF-ENG. Boston: Harvard Business Publishing.
- Inman, J. Jeffrey, James S. Dyer, and Jianmin Jia (1997), "A Generalized Utility Model of Disappointment and Regret Effects on Post-Choice Valuation," *Marketing Science*, 16 (2), 97–111.
- Jermann, Urban and Vincenzo Quadrini (2009), "Macroeconomic Effects of Financial Shocks," NBER Working Paper No. 15338, National Bureau of Economic Research.
- Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47 (2), 263–91.
- Kamakura, Wagner A., Vikas Mittal, Fernando de Rosa, and Jose A. Mazzon (2002), "Assessing the Service Profit Chain," *Marketing Science*, 91 (3), 294–317.
- Kilian, Lutz (2009), "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market," *American Economic Review*, 99 (3), 1053–69.
- and Robert J. Vigfusson (2011), "Are the Responses of the US Economy Asymmetric in Energy Price Increases and Decreases?" *Quantitative Economics*, 2 (3), 419–53.
- Klein, Jill and Niraj Dawar (2004), "Corporate Social Responsibility and Customers' Attributions and Brand Evaluations in a Product-Harm Crisis," *International Journal of Research in Marketing*, 21 (3), 203–217.
- Knox, George and Rutger van Oest (2014), "Customer Complaints and Recovery Effectiveness: A Customer Base Approach," *Journal of Marketing*, 78 (September), 42–57.
- Lamey, Lien, Barbara Deleersnyder, Marnik G. Dekimpe, and Jan-Benedict E.M. Steenkamp (2007), "How Business Cycles Contribute to Private-Label Success: Evidence from the United States and Europe," *Journal of Marketing*, 71 (January), 1–15.
- Malhotra, Arvind and Claudia Kubowicz Malhotra (2011), "Evaluating Customer Information Breaches as Service Failures: An Event Study Approach," *Journal of Service Research*, 14 (1), 44–59.
- Mittal, Vikas, Pankaj Kumar, and Michael Tsiros (1999), "Attribute-Level Performance, Satisfaction, and Behavioral Intentions over Time: A Consumption-System Approach," *Journal of Marketing*, 63 (April), 88–101.
- Oliver, Richard L. (1977), "Effect of Expectation and Disconfirmation on Postexposure Product Evaluations: An Alternative Interpretation," *Journal of Applied Psychology*, 62 (4), 480–86.
- (1980), "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research*, 17 (November), 460–69.
- , Roland T. Rust, and Sajeew Varki (1997), "Customer Delight: Foundations, Findings, and Managerial Insight," *Journal of Retailing*, 73 (3), 311–36.
- Ozcek, Omer and W. Douglas McMillin (1999), "Lag Length Selection in Vector Auto-Regressive Models: Symmetric and Asymmetric Lags," *Applied Economics*, 31 (4), 517–24.
- Phillips, Peter C.B. and Pierre Perron (1988), "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75 (2), 335–46.
- Potter, Ned (2011), "BlackBerry Outage: RIM Apologizes, Says Service Returning," ABCNews.com, (October 13), (accessed June 10, 2015), [available at <http://abcnews.go.com/Technology/blackberry-outage-rim-ceo-apologizes-service-returning-normal/story?id=14727816>].
- Randles, Ronald H., Michael A. Fligner, George E. Policello, and Douglas A. Wolfe (1980), "An Asymptotically Distribution-Free Test for Symmetry Versus Asymmetry," *Journal of the American Statistical Association*, 75 (369), 168–72.
- Rego, Lopo L., Neil A. Morgan, and Claes Fornell (2013), "Reexamining the Market Share–Customer Satisfaction Relationship," *Journal of Marketing*, 77 (September), 1–20.
- Rubel, Olivier, Prasad A. Naik, and Shuba Srinivasan (2011), "Optimal Advertising When Envisioning a Product-Harm Crisis," *Marketing Science*, 30 (6), 1048–65.
- Rust, Roland T., J. Jeffrey Inman, Jianmin Jia, and Anthony Zahorik (1999), "What You Don't Know About Customer-Perceived Quality: The Role of Customer Expectation Distributions," *Marketing Science*, 18 (1), 77–92.
- , Christine Moorman, and Peter R. Dickson (2002), "Getting Return on Quality: Revenue Expansion, Cost Reduction, or Both?" *Journal of Marketing*, 66 (October), 7–24.
- , Anthony J. Zahorik, and Timothy L. Keiningham (1995), "Return on Quality: Making Service Quality Financially Accountable," *Journal of Marketing*, 59 (April), 58–70.
- Sheldon, Kennon M., Richard Ryan, and Harry T. Reis (1996), "What Makes for a Good Day? Competence and Autonomy in the Day and in the Person," *Personality and Social Psychology Bulletin*, 22 (12), 1270–79.

- Sivakumar, K., Mei Li, and Beibei Dong (2014), "Service Quality: The Impact of Frequency, Timing, Proximity, and Sequence of Failures and Delights," *Journal of Marketing*, 78 (January), 41–58.
- Skowronski, John J. and Donal E. Carlston (1989), "Negativity and Extremity Biases in Impression Formation: A Review of Explanations," *Psychological Bulletin*, 105 (1), 131–42.
- Slotegraaf, Rebecca J. and Koen Pauwels (2008), "The Impact of Brand Equity and Innovation on the Long-Term Effectiveness of Promotions," *Journal of Marketing Research*, 45 (June), 293–306.
- Smith, Amy K. and Ruth 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.
- , ———, and Janet Wagner (1999), "A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery," *Journal of Marketing Research*, 36 (August), 356–72.
- Srinivasan, Shuba, Koen Pauwels, Jorge Silva-Risso, and Dominique M. Hanssens (2009), "Product Innovations, Advertising, and Stock Returns," *Journal of Marketing*, 73 (January), 24–43.
- , Marc Vanhuele, and Koen Pauwels (2010), "Mind-Set Metrics in Market Response Models: An Integrative Approach," *Journal of Marketing Research*, 47 (August), 672–84.
- Stephens, Melvin, Jr. (2001), "The Long-Run Consumption Effects of Earnings Shocks," *Review of Economics and Statistics*, 83 (1), 28–36.
- Stock, James H. and Mark W. Watson (2001), "Vector Autoregressions," *Journal of Economic Perspectives*, 15 (4), 101–115.
- Szymanski, David M. and David H. Henard (2001), "Customer Satisfaction: A Meta-Analysis of the Empirical Evidence," *Journal of the Academy of Marketing Science*, 29 (1), 16–35.
- Van Doorn, Jenny and Peter C. Verhoef (2008), "Critical Incidents and the Impact of Satisfaction on Customer Share," *Journal of Marketing*, 72 (July), 123–42.
- Van Heerde, Harald J., Kristiaan Helsen, and Marnik G. Dekimpe (2007), "The Impact of a Product-Harm Crisis on Marketing Effectiveness," *Marketing Science*, 26 (2), 230–45.
- Villanueva, Julian, Shijin Yoo, and Dominique M. Hanssens (2008), "The Impact of Marketing-Induced Versus Word-of-Mouth Customer Acquisition on Customer Equity Growth," *Journal of Marketing Research*, 45 (February), 48–59.
- Vitorovich, Lilly and Jessica Hodgson (2012), "RIM Restores BlackBerry Service in Europe," *The Wall Street Journal*, (September 21), (accessed June 10, 2015), [available at <http://online.wsj.com/news/articles/SB10000872396390444165804578009801352709628>].
- Wangenheim, Florian v. and Tomás Bayón (2007), "Behavioral Consequences of Overbooking Service Capacity," *Journal of Marketing*, 71 (October), 36–47.
- Zeithaml, Valarie A., Leonard L. Berry, and A. Parasuraman (1996), "The Behavioral Consequences of Service Quality," *Journal of Marketing*, 60 (April), 31–46.
- , Mary J. Bitner, and Dwayne D. Gremler (2012), *Services Marketing*, 6th ed. New York: McGraw-Hill.
- Zhao, Yi, Ying Zhao, and Kristiaan Helsen (2011), "Consumer Learning in a Turbulent Market Environment: Modeling Consumer Choice Dynamics After a Product-Harm Crisis," *Journal of Marketing Research*, 48 (April), 255–67.