11-2021

Does active service intervention drive more complaints on social media? The roles of service quality and awareness

Shujing SUN

Yang GAO
Singapore Management University, ygao@smu.edu.sg

Huaxia RUI

DOES ACTIVE SERVICE INTERVENTION DRIVE MORE COMPLAINTS ON SOCIAL MEDIA? THE ROLES OF SERVICE QUALITY AND AWARENESS

Shujing Sun, University of Texas at Dallas, shujing.sun@utdallas.edu
Yang Gao, Singapore Management University, ygao@smu.edu.eg
Huaxia Rui, University of Rochester, huaxia.rui@simon.rochester.edu, corresponding author

Last revised: May 11, 2021

ABSTRACT

Despite many advantages of social media as a customer service channel, there is a concern that active service intervention encourages excessive service complaints. Our paper casts doubt on this misconception by examining the dynamics between social media customer complaints and brand service interventions. We find service interventions indeed cause more complaints, yet this increase is driven by service awareness rather than chronic complaining. Due to the publicity and connectivity of social media, customers learn about the new service channel by observing customer service delivery to others – a mechanism that is unique to social media customer service and does not exist for traditional call centers. Importantly, high-quality service reduces future complaints, thereby proactive customer service is a sound strategy on social media, as long as firms dedicate to service quality. Hence, firms should be less concerned about *whether* to respond and more focused on *how* to respond to customer complaints.

KEY WORDS AND PHRASES: social media, customer service, complaint management, awareness enhancement, chronic complainer, Twitter

---

Authors' names and postal addresses:

Shujing Sun, JSOM 3.208, University of Texas at Dallas, 800 W. Campbell Rd, Richardson, TX 75080
Yang Gao, 80 Stamford Road, Singapore Management University, Singapore 178902
Huaxia Rui, CS 3-310 Carol Simon Hall, University of Rochester, Rochester, NY, 14627

INTRODUCTION

Customer complaints on social media are prevalent today. An industry report in 2017 showed that "*brands receive 146% more social messages needing responses than they did three years ago*" [39]. The number of customer service interactions on Twitter also increased by 250% from 2015 to 2017 [16]. Because of the vast publicity and the large audience, firms naturally worry about customers' rising complaints and possible negative publicity. As quoted in Form 10-K filed by Delta airlines to SEC in 2019, "*[a]dverse publicity, whether justified or not, can rapidly spread, including through social or digital media. In particular, passengers can use social media to provide feedback about their interaction with us in a manner that can be quickly and broadly disseminated*" [50, p. 16]. Indeed, delivering customer service on social media is a delicate business, the public and connected nature amplifies the impact of each customer complaint and renders firm responses vulnerable to open scrutiny in a way never seen before. Hence, a natural but fundamental question faced by many firms is: What is driving customers' heightened interests to complain on social media?

One possibility is that more customers are simply switching from traditional call centers to new digital channels. According to a *Wall Street Journal* (WSJ) article in June 2020, the multiplication of digital channels explains some of the upticks in the rising customer complaints. "*Customers with complaints are increasingly using email, live chat, and social media – 43%, compared with 12% in 2017*", says Scott Broetzmann, president and chief executive of Alexandria, Virginia-based Customer Care Measurement [4]. The channel switching following this scenario may benefit firms because, according to McKinsey & Company, the handling cost per customer on Twitter is about one-sixth of the handling cost through traditional call centers.[1] Moreover, the proper handling of a customer's complaint can not only turn a service failure into a positive brand experience for that particular customer but it also creates a positive brand perception among other social media users who have observed the customer service interaction, thanks to the nature of social media.

Another possibility is that customers who previously did not bother to complain are more willing to voice their dissatisfaction because of the convenience of voicing on social media. Instead of waiting for hours to get a call picked up by call center agents, customers can effortlessly send a tweet to a brand's social media account and communicate with agents asynchronously at their convenience. This is also a benign scenario to firms because, according to Hirschman's theory of *exit, voice, and loyalty* [23], a customer's voicing decision is an alternative to exiting a firm's business. Hence, providing a convenient voicing channel can help retain those customers who would otherwise withdraw from the business relationship without any attempt to repair, thereby improving customer-brand relationships and reducing future customer acquisition costs [28, 43].

There is yet another possibility that can be alarming. The service intervention itself might have been fueling the rise of complaints on social media. In fact, Scott Broetzmann expects a further rise of customer complaints on social media as companies become more proactive, according to the same WSJ article [4]. After studying how managerial responses to negative online review affect future negative reviews, Chevalier et al. [13, p. 707] concluded that managers face a "tough dilemma" because, "*on one hand, a managerial response to a negative review may neutralize the possible negative effect of the original review on future bookings,*" but "*on the other hand, firms also run the risk of encouraging further critical reviews by responding to customers, which may hurt future bookings.*" Another paper by Ma et al. [31] similarly found that customers who received service intervention on Twitter would complain more in the future because of their increased expectations for service. This finding further exacerbates the concern about chronic complainers. Following the paper's publication, several business reporters directly warned that "*[c]ompanies engaging with customers on Twitter beware: responding to complaints on social media has the side effect of triggering new complaints"* [38, 44] and "*Twitter replies to customer grievances could trigger more complaints*" [36]. The fact that on average, only 10% of customer complaints on social media are responded by firms [39] while 55% of complainers who reached out on social media without receiving a response said they expected the company to respond [4], also seems to validate firms' concern about chronic complainers.

Does active service intervention drive more service complaints on social media? If so, should managers be concerned about this while forming their social media strategies? The key to answering these questions is to clarify the underlying mechanisms, as the implications behind each mechanism can be drastically different. While the findings in Ma et al. [31] suggest that service interventions on social media increase customers' tendencies to complain more in the future, their study is based on a fixed sample of 714 customers from one telecommunications firm. Before delving into a more in-depth analysis, we found that 98% of customers sent less than two complaints per year over our sample period. This model-free statistic further sheds doubt on the mechanism of "chronic complainers", which is likely a stumbling block that holds up firms' active management of service requests on social media. Therefore, we are motivated to offer managers a complete picture by exploring alternative mechanisms behind customers' "seemingly heightened" interests to complain on social media. Our findings shed light on these crucial questions facing many managers who wish to harness the power of social media for their brands while avoiding the pitfalls.

Based on the analyses of a large data set consisting of all customer service interactions between 40 major international airlines and their customers on Twitter from January 2014 to September 2019, we found that service intervention, when measured as the extent (i.e., volume) of a firm's responses to customer complaints, indeed led to more future complaints on social media. However, when service intervention was measured by the quality of response (i.e., the promptness and effectiveness of brand responses), it actually led to fewer future complaints.

To understand the seemingly paradoxical findings, we further analyze the mechanism behind the positive relationship between service intervention volume and future complaints. First, by examining customers who complained multiple times over the years, we found that they only constituted a small proportion of all complaining customers. Also, customers did not complain excessively conditional on receiving brand service interventions. Therefore, the issue revealed by Ma et al. [31] and warned about by business media seems to be exaggerated at the aggregate level, at least in the airline industry for which customer service is particularly important. Second, unlike customer service through traditional channels that are always private, there is a dual role of each service intervention on social media: addressing a focal customer's complaint and increasing awareness of the new service channel among customers connected to the complaining customer. Thus, this awareness mechanism naturally connects the volume of service interventions and future complaints on social media.

To test this mechanism, we leverage an obscure technical feature of Twitter through which certain complaints are made more visible to other customers. We constructed a variable that measures this "awareness stock" and performed a joint test with both service volume and service awareness included. Just as we have conjectured, the estimation results confirm awareness stock's significant effect on future complaints, conditional on the actual service volume to focal customers. The awareness enhancement mechanism is unique to social media customer service due to its public and connected nature, which are unlikely to exist in other service channels lacking a social network aspect, such as traditional private call centers and online review platforms.

The managerial implication from our findings is thus clear: firms have more reasons to embrace social media customer service and should be less concerned about the channel being abused by disgruntled customers. As of 2017, firms responded to about 10% of customer complaints on social media. Even for the airline industry that pioneered the adoption of social media customer service, nearly half of all customer complaints were not responded to [20]. Given that more people are adopting social media and the cost of delivering customer service through social media is much lower, we believe the current level of service intervention is sub-optimal, resulting in plenty of wasted opportunities. Moreover, to avoid the potential negative externality, firms should provide quality customer service on social media. According to our findings, more prompt responses and a higher resolution rate actually reduced future complaints.

RESEARCH BACKGROUND AND HYPOTHESES

In this section, we introduce the research background. Then we develop two sets of hypotheses corresponding to the *service volume* (i.e., the number of brand service interventions to customer-initiated redress-seeking posts) and the *service quality* (i.e., the promptness and effectiveness of brand responses).

Research Background

Customer complaint management has long been recognized as critical to firms by researchers and practitioners. There is extensive Marketing literature studying motivations underlying customers' voicing decisions. According to Hirschman [23], when customers perceive deteriorated service or product quality, they choose to either exit from a firm's business or voice to elicit a brand response and solicit individual compensation. Based on Hirschman's theory, Fornell and Wernerfelt [17] proposed an economic model of a defensive marketing strategy for complaint management. They suggested that the objective to reduce the number of customer complaints about a product is questionable. Instead, firms maximize complaints from dissatisfied customers subject to certain cost restrictions, as defensive marketing (e.g., complaint management) can lower the total marketing expenditure by substantially reducing the cost of offensive marketing (e.g., advertising). Follow-up literature has shown that customers' word-of-mouth is driven by *intrinsic factors*, such as an individual's desire for social interaction, economic incentives, self-enhancement, and concerns of others' perceptions [22, 46]; and *external factors*, such as product characteristics [8] and brand characteristics [30].

Advancements in information technology and mobile devices have enabled customers to voice their opinions about firms' products and services at the greatest convenience. Among various online channels, online review platforms have drawn increasing attention from practitioners and researchers. From the perspective of customer satisfaction, Gu and Ye [18] found that online management responses are highly effective among low satisfaction customers but have limited influence on other customers. A few recent papers in the Marketing literature studied the positive consequences of managerial responses to online hotel reviews. For instance, Wang and Chaudhry [51] found that managerial responses to negative hotel reviews can positively influence subsequent opinions if those responses are observable at the time of reviewing. Proserpio and Zervas [37] showed that managerial response indeed increases a hotel's star rating. In contrast, a recent paper by Chevalier et al. [13] argued that a customer is motivated to write reviews not only because reviews may impact other customers, but also because reviews may impact the management and the quality of the service. Therefore, the managerial response will stimulate negative reviews that are seen as more impactful in the eyes of customers. Using data from online customer reviews of travel agents, Yang et al. [55] developed a stochastic differential equation model that describes how average review ratings react to the arrival of new reviews and brand response. They suggested that responding to every customer review will likely be either too costly or ineffective if the responses are not adequate, thereby suggesting firms respond in a selective manner.

Beyond managerial responses, firms can now provide actual customer service through various service-oriented technology innovations and applications, such as web-based self-service portals and social media platforms [6, 29, 47]. Besides the benefits of cutting costs and improving efficiency, those customer-led service channels have been shown to improve customer satisfaction and profitability [54]. Given the great promise and the drastic differences from conventional call centers, the social media customer service channel has drawn increasing attention from Information Systems and Marketing researchers. Utilizing Twitter data from a major U.S. airline, Gunarathne et al. [19] identified several determinants of customers' post-complaint satisfaction via social media customer service. The paper suggested that customers' online influence, previous experience, and complaint type jointly shape customer feelings toward brand-customer interactions.

From the organizational perspective, Gunarathne et al. [20] and Hu et al. [25] found evidence of firms' differential service treatment by customers' online influence and politeness. Ma et al. [31] examined the effect of service intervention on customer voices to a telecommunications firm on Twitter. The paper found that redress seeking is a major driver of complaints, and the firm's service interventions will encourage repeated complaints from individual customers. Mousavi et al. [34] also used the Twitter data of telecommunications firms and identified factors and external events that can influence the effectiveness of customer care. The paper suggested a clear separation of the four firms in their customer service provision and that customers expect better customer care for higher-priced firms. Moreover, seemingly unrelated events, such as signing an exclusive contract with a celebrity, can impact digital customer care.

We extend the literature on the emerging technology-enabled service channels by focusing on social media-based customer service, where firms use social media platforms to assist customers with service requests and concerns. The context relates to brand managerial responses to online reviews but differs in several fundamental aspects. First, despite both online review platforms and social media platforms serving as public channels for customers to voice, social media customer service has equivalent functions to a conventional call center, while brand managerial responses to online reviews do not typically involve actual service interventions. As a result, the underlying motivations are different between customers posting online reviews and customers requesting service support on social media. Second, unlike social media platforms, online review platforms do not have the "social aspect", which likely results in different dynamics between brand service intervention and customer complaints. For instance, because of the social network, customer service provision on social media may affect not only a focal complainer but also individuals connected to the complainer. To the best of our knowledge, few papers have investigated such dynamics. The only exception is the study by Ma et al. [31], which analyzed how a firm's service intervention changes individual customers' subsequent decisions to voice. Through a dynamic choice model on a sample of 714 customers from a telecommunications firm, the authors found that although service intervention improved customer-brand relationships, it also encouraged more future complaints from individual customers.

Compared to the pioneering works about the social media customer service channel, our work has several notable differences. First, while previous literature primarily focused on individuals' motivations to voice, we are interested in the relationships of aggregated customer complaints and brand service interventions within a dynamic setting. Our unique data set allows us to identify the evolving process of brand-level complaint behavior, accounting for the dynamic changes in customer population, service quality, and online and offline shocks to a brand. Moreover, due to the public and connected nature of social media, individuals' decisions are no longer independent. Therefore, we believe an aggregated level analysis is critical for a firm's resource allocation strategy.

Second, instead of considering the service intervention as a binary treatment, we extract multi-dimensional service strategy measures from the textual information of customer-brand conversations. This allows us to conduct a more granular analysis of service strategy from various aspects and identify their differential effects on customers' complaint behavior. Such a detailed analysis is also crucial for identification purposes, as customer complaint behaviors are influenced by multiple endogenous factors, of which service quality is an important aspect. Lastly, we collected an inclusive data set covering 40 major international airlines throughout almost six years. As the airline industry is a leading customer service provider on social media, we believe findings built on this comprehensive and inclusive dataset will be less subject to external validity concerns. Accordingly, we believe the insights on social media customer service from this paper will be informative to researchers and practitioners.

Development of Hypotheses

To understand how *service volume* might affect the aggregated customer complaint volume, we distinguish two groups of customers: *focal customers* who complain to firms on social media and *bystanders* who observe the service encounters between focal customers and the involved firms. As a brand response credibly signals its service availability, we argue that the volume of brand service can affect customer complaints in the following ways.

First, evidence on the effect of service volume on a focal customer's voicing propensity is mixed. On the one hand, active responses signal a firm's care about the customer, which contributes to her commitment and loyalty to the brand. Previous studies showed that effective management of complaints fosters customer-brand relationships and prevents customer churn through service recovery [11, 28]. Recent studies also found that online management responses positively affect customer satisfaction [18], hotel's star rating [37], and customers' subsequent reviews [51]. With the improved customer-brand relationship and enhanced brand perception, the focal customer is less likely to complain to the brand in the future [43]. On the other hand, brand service interventions may incur more subsequent complaints from the focal customer, as prior experience boosts her confidence about a firm's accountability and the likelihood of resolving problems. In fact, a recent paper by Ma et al. [31] suggested that more service interventions encourage repeated complaints from individual customers to a telecommunications firm on Twitter. As such, even with a fixed customer base, a firm may still experience an increasing trend in customer complaints as the service volume increases.

Second, due to the public nature of social media, a focal customer's complaint and the associated brand service interventions are observable to bystanders. This helps increase the bystanders' awareness of the new service channel and redirect them to social media for future service requests. Blodgett et al. [10] suggested that a primary determinant on whether a customer seeks to redress is the perceived likelihood of success. Even customers who would otherwise exit are more likely to seek redress if it is clear that the seller is willing to remedy the problem. As the service volume increases, a brand essentially advertises its service availability and care about customers, which further enhances bystanders' expectation of getting customer support through social media. Collectively, the increased service awareness may drive more customers to social media for future redress seeking. Moreover, survey data show that focal customers are more likely to share their experiences, both online and offline, after receiving a firm's response [26]. Such word-of-mouth generated by focal customers also facilitates the overall awareness of a brand's social media customer service.

Although service intervention may affect focal customers' complaints in competing directions, we believe that more brand interventions will likely increase overall complaints to a firm. Because a focal customer may have hundreds or even thousands of followers, the size of potential bystanders is much larger than the size of focal customers. Consequently, the awareness enhancement effect should play a dominating role in driving customer complaints to a brand. To examine this empirically, we propose the following hypothesis for a statistical test.

- **Hypothesis 1 (Service Volume Effect):** *A higher service volume in terms of more brand interventions will lead to more customer complaints.*

Service Quality and Customer Complaints

Much like the conventional calls at an 800 number, customers expect quick and effective responses on social media. While customers certainly benefit from high-grade service intervention, it remains unclear how service quality influences their propensities to complain in the future. Holding everything else the same, we hypothesize that *service quality* may affect customer complaints through two competing mechanisms, though it is not obvious which one will be true with a specific setting or empirical data.

Anecdotal evidence implies that 31% of customers in the U.S. expect a response within 24 hours or less, 24% expect a response within an hour, and 20% expect to get a response immediately [40]. As social media enables convenient and almost instant sharing among friends, a prompt and effective response can not only turn a service failure into a positive brand experience for the complaining customer but also deliver a positive brand image to bystanders. Besides, a prompt intervention helps avoid new problems or aggravate existing ones [33]. Compared to the effect of active service responses, service quality is more likely to improve the customer-brand relationship. Accordingly, we expect that high-quality customer support can better facilitate broadcasting the positive brand image and, as a result, more likely to reduce future complaint occurrences to a large extent.

On the other hand, given the same service volume, higher customer service quality may attract customers to shift from conventional call centers to social media for redress seeking, especially when they perceive the overall quality through the social media channel as higher than that of traditional call centers. In such cases, prompt and effective care might *encourage* future complaints. Nonetheless, there is an important caveat in this mechanism. On Twitter, although a conversation between a focal user and a third-party is shown in her followers' Twitter feed in real-time, the conversation by default is displayed as truncated. Despite that Twitter has adopted steady improvements that help follow conversations,[2] a bystander still has to take extra steps to expand the full conversation. As such, for bystanders who do not bother to track and read through a focal customer's service conversation, they cannot observe the service quality. Hence, the strength of this mechanism depends on the number of bystanders who actively monitor focal customers' service encounters.

Based on these arguments, we propose the following competing hypotheses for empirical testing:

- **Hypothesis 2A (Snowball Effect of Service Quality):** *Higher customer service quality leads to more customer complaints.*
- **Hypothesis 2B (Neutralizing Effect of Service Quality):** *Higher customer service quality leads to fewer customer complaints.*

DATA

We used the real-time Twitter API to collect all tweets posted and received by 40 international airlines from January 2014 to September 2019. We collected data from Twitter as it is extensively leveraged for social media customer service.[3] Please see Appendix A (Table A1) for the full list of firms included in the sample. For firms with multiple Twitter accounts, we included its main account in our sample, which is the most influential (i.e., with the largest number of followers). For non-English tweets, we first used the Google Translation API to translate them to English before extracting relevant features, such as sentiment. We included only customer-initiated tweets to a brand, and constructed service conversations consisting of a thread initiator (i.e., customer-initiated tweet) and follow-up brand-customer communications associated with the thread. Next, we adopted a lexicon proposed for social media customer service [20] and implemented a bag-of-words approach to categorize customer-initiated tweets into complaints, compliments, and informational posts. The details are described in Appendix B. We then aggregated the conversational data to the firm-week level.

Table 1 reports the definitions and summary statistics of the key variables.

Table 1. Variable Definition and Summary Statistics

Variable	Observation	Mean	SD	Definition
Customer Voice				
logComplaints	11,584	4.23	1.32	Volume of customer complaints (log-transformed)
neuRatio	11,584	59.84	11.39	Percentage (%) of neutral voices among all the customer voices
Service Volume				
logReplies	11,584	3.29	1.47	Number of brand replies (i.e., service interventions) to customer-initiated complaints (log-transformed)
Service Quality				
Delay	11,584	4.31	1.63	Time (in minutes) from a brand's first reply to customer-initiated complaints (log-transformed)
Resolution	11,584	0.39	0.29	Ratio of redress seeking conversations ended with a resolution
CustomerGratitude	11,584	0.09	0.11	Ratio of redress seeking conversations with customers expressing their gratitude to agents
Other Controls				
ReplyLength	11,584	19.92	3.35	Average number of words per brand reply
ConversationLength	11,584	1.71	0.62	Average number of brand replies per service intervention
DirectMessage	11,584	0.16	0.14	Ratio of redress seeking conversations with direct messages
Please	11,584	0.17	0.13	Ratio of agents' usage of "please" in replies
Apology	11,584	0.14	0.12	Ratio of agents' apology to customers in replies
Reasoning	11,584	0.02	0.04	Ratio of agents' explicit reference to reasons in replies
Reassurance	11,584	0.003	0.010	Ratio of agents' efforts to minimize customers' concerns in replies
Gratitude	11,584	0.18	0.16	Ratio of agents' expressing appreciation to customers in replies
Brand Control				
OfflineIncidents	11,584	0.23	0.57	Number of brands' offline incidents
GoogleTrend	11,584	52.80	19.08	Google search volume index for a brand
logFollowers	11,584	12.55	1.33	Number of brand's Twitter followers (log-transformed)
Service Awareness				
AwarenessStockM	11,584	4.95	1.796	Cumulative awareness through mentions (log-transformed)
AwarenessStockR	11,584	4.57	1.795	Cumulative awareness through replies (log-transformed)
ATCR Controls				
PassengerVolume	1,979	4.62	4.48	Number of enplaned passengers (in million)
%FlightDelay	1,979	19.70	7.08	Percentage of flights that are delayed
%FlightCancellation	1,979	1.16	1.27	Percentage of scheduled flights that are canceled
%BaggageClaim	1,979	2.88	1.48	Percentage of passengers claimed for mishandled baggage
Instrumental Variable				
$\Delta NeuVoice$	11,584	0.0004	0.0702	Changes in the fraction of neutral customer voices

Note. This table reports summary statistics and definitions of key variables at the firm-week level, which corresponds to the sample of firms listed in Appendix A (Table A1) from January 2014 to September 2019. SD stands for standard deviation.

The dependent variable is the log-transformed volume of customer-initiated complaints directed to firm i at week t ($logComplaints_{i,t}$). All thread initiators were counted in the dependent variable construction, regardless of whether a service request received a brand response or not. Corresponding to the hypotheses, we consider two aspects of brand service intervention as independent variables: *service volume* and *service quality*.

We measure *service volume* by the number of service interventions to customer-initiated complaints by firm *i* at week *t*, with the logarithmic transformation ($logReplies_{i,t}$). Similar to the construction of $logComplaints_{i,t}$, only the first brand response to the initiated customer tweet is counted when calculating $logReplies_{i,t}$, which is essentially the total number of customer-initiated tweets responded by firm *i* at week *t*.

We measure *service quality* from several aspects. The first dimension is $Delay_{i,t}$, which captures the timeliness of the service and is calculated as the average delay in minutes from a brand's first reply to a customer-initiated complaint. Although a prompt response is an important signal for service quality, a quick but very generic response may not fully capture the overall customer service quality. Accordingly, we propose the second measure, service effectiveness, constructed as the ratio of redress seeking conversations that end with a resolution ($Resolution_{i,t}$). To determine the resolution of each customer service encounter, we employ a supervised classifier trained with a labeled data set, the predictive performance of which is reported in Table C2 of Appendix C. The larger the resolution rate, the higher the service quality is. We also constructed an alternative measure for service effectiveness as the ratio of service encounters that end with customers' explicit expressions of gratitude to agents ($CustomerGratitude_{i,t}$). To identify whether a customer explicitly appreciates an agent's efforts, we implemented a lexicon-based approach following Yeomans et al. [56].

In addition to the brand service volume and service quality, we included various controls for average agents' response characteristics that are constructed from the conversational data. We use $ReplyLength_{i,t}$ to measure agents' efforts in responding to complaints. Longer replies imply more efforts. We use $ConversationLength_{i,t}$ to capture how efficient a service agent can handle a complaint. We considered firms that address a complaint with fewer replies as more efficient. We used a lexicon-based method to check if a customer service interaction includes a *direct message* and constructed $DirectMessage_{i,t}$ as the ratio of redress seeking conversations with communications in direct messages. Following Yeomans et al. [56], we further created a list of linguistic features to quantify the politeness of agents' replies, which may have direct consequences on a service intervention.

Note that the total complaints to a brand are influenced by multiple endogenous factors. For example, if there is a shock to the customer population or brand service quality, there could be a spike in both customer voices and the brand service volume. This will lead to a positive relationship between these two factors, even if there is no direct relationship. To identify the effect of brand service volume and quality, we included three variables to account for firm heterogeneity and alleviate endogeneity concerns. Specifically, we use $OfflineIncidents_{i,t}$ and $GoogleTrend_{i,t}$ to capture any offline and online shocks that could simultaneously affect customer voices and brand service strategy.[4] We used brand Twitter followers ($logFollowers_{i,t}$) to proxy for a firm's Twitter customer base that closely correlates with the volume of customer voices and number of offline customers. For the robustness checks that we will perform, we further collected the Air Travel Consumer Report (ATCR) from the U.S. Department of Transportation, which measures the offline service performance of the major U.S. airlines at monthly intervals.

EMPIRICAL ANALYSIS

In this section, we first discuss the empirical strategy and then present the empirical results on the dynamic response of customers' complaints to firms' service strategies.

Model Specification

To examine the effect of service intervention on the volume of customer complaints, we use the following empirical specification:

$$Y_{i,t} = \beta_1 ServiceVolume_{i,t-1} + \beta_2 ServiceQuality_{i,t-1} + \beta_3 X_{i,t} + \beta_4 TimeTrend_t + \alpha_i + \delta Seasonality_t + \epsilon_{i,t}$$

where $Y_{i,t}$ is the log-transformed volume of customer-initiated service complaints to firm *i* at week *t*.[5] The key independent variables include lagged measures for service volume (i.e., $logReplies_{i,t-1}$) and service quality (i.e., $Delay_{i,t-1}$, $Resolution_{i,t-1}$, and $CustomerGratitude_{i,t-1}$), the structure of which precludes a potential reverse-causality explanation. Given that the data has a long panel structure with many periods for relatively few firms ($N = 40, T = 298$ weeks),[6] we took advantage of the natural ordering of time (in weeks) and include the linear time trend to account for the platform growth. On top of various controls, we include firm fixed effects, α_i, to capture unobserved time-invariant differences across brands. We also control for seasonality effects through year and month fixed effects, which capture the unobserved common shocks in customer voicing decisions and brand reply strategies.

A major concern for long-panel data is that error terms may not be independent and identically distributed. The failure to correct for serial correlation of errors, if present, can cause the standard errors of the estimates to be smaller compared to their actual values, thus leading to incorrect tests of hypotheses [9]. As a first step, we refer to the Wooldridge test, and the result rejected the null hypothesis that there is no serial correlation in the residuals of the fixed effects model ($F(1,39) = 171.329, p < 0.01$) [15, 52]. To account for the autocorrelation in the disturbance term, we specified the error term, $\epsilon_{i,t}$, as autoregressive with order one AR(1):

$$\epsilon_{i,t} = \rho\epsilon_{i,t-1} + \eta_{i,t}$$

where $|\rho| < 1$ and $\eta_{i,t}$ is independent and identically distributed [5]. The AR(1) error structure captures persistence in customers' and firms' behaviors, such as word-of-mouth among customers, growth of brands' online social networks, evolution in the social media service labor, etc. We implemented the analysis using the STATA procedure XTREGAR [12, 24]. All the results remain robust when we imposed panel-specific AR(1) disturbance and are available upon request. However, since Beck and Katz [7] recommended against estimating panel-specific AR parameters, as opposed to one AR parameter for all panels, we report results assuming common AR(1) disturbance across panels in the current paper.

Baseline Results

Table 2 reports the estimation results on the effects of brand service intervention on customer complaints. Take column (1) as an example, the coefficient of *service volume* ($logReplies_{i,t-1}$) is positive and statistically significant, thereby supporting the Service Volume Effect Hypothesis (H1), which states that a higher service volume leads to more customer complaints in the following period. In terms of the magnitude, a 1% increase in the service volume leads to about a 0.13% increase in customer complaints.

Table 2. Effect of Brand Service Intervention on Customer Complaints

Variables	(1)	(2)	(3)	(4)
Service Volume				
$logReplies_{i,t-1}$	0.132*** (0.008)	0.137*** (0. 008)	0.125*** (0. 008)	0.130*** (0. 008)
Service Quality				
$Delay_{i,t-1}$	0.013*** (0.004)	0.012** (0.004)	0.013*** (0.004)	0.012** (0.004)
$Resolution_{i,t-1}$	-0.081*** (0.017)	-0.082*** (0.017)		
$CustomerGratitude_{i,t-1}$			-0.092* (0.040)	-0.094* (0.040)
Brand Control				
$OfflineIncidents_{i,t}$	4.506*** (0.763)	4.493*** (0.764)	4.498*** (0.763)	4.485*** (0.764)
$GoogleTrend_{i,t}$	0.534*** (0.040)	0.528*** (0.040)	0.540*** (0.040)	0.533*** (0.040)
$logFollowers_{i,t}$	0.272*** (0.024)	0.217*** (0.028)	0.273*** (0.024)	0.218*** (0.028)
Linear Time Trend	Y	Y	Y	Y
Quadratic Time Trend	N	Y	N	Y
R^2	0.156	0.158	0.152	0.155
ρ_{ar}	0.302	0.298	0.307	0.302

Note. *** $p < 0.001$ ** $p < 0.01$* $p < 0.05, + p < 0.1$. This table reports the results on the dynamics between brand service interventions and customer complaints. The number of observations is 11,544. The number of observations was reduced by 40 because of the panel-by-panel Cochrane–Orcutt transformation. The dependent variable is $logComplaints_{i,t}$. The estimation is a fixed-effects linear model with the first-order autocorrelation AR(1) disturbance across panels. *OfflineIncidents* and *GoogleTrend* are rescaled by a factor of 1/100 to allow for more non-zero digits in the estimation. Throughout the analyses, we accounted for other controls, and firm, year, month fixed effects.

Regarding *service quality* measures, we found a significant and positive coefficient estimate for $Delay_{i,t-1}$[7] and a significant and negative coefficient estimate for $Resolution_{i,t-1}$. The results suggest that delayed responses can lead to more future complaints, and better resolution can reduce future complaints. Both findings support the Neutralizing Effect of Service Quality Hypothesis (H2B). In terms of the magnitude, a 1% delay in brands' first responses to service complaints associates with a 0.01% increase in future complaints, and a 1% increase in resolution rate associates with a 0.08% reduction in future complaints. The Snowball Effect of Service Quality Hypothesis (H2A) was not supported. Overall, the findings suggest that high service quality reduces rather than increases future complaints. Consistent with the literature that effective service intervention improves customer-brand relationships [28], the improved brand image may attenuate the immediate negativity of a focal complaint and reduce future complaint occurrences.

Customers' voicing decisions were influenced by multiple factors that are likely endogenous, which made it necessary to control for offline service shocks and online customer growth. From the regression analysis, we observed the expected coefficient estimates of key brand controls. For example, shocks to both offline and online performance correlated with more future complaints. The increasing number of brand followers also explains a large proportion of the increasing volume of customer complaints. Considering that firms may have nonlinear growth patterns in the social media customer service provision, we further imposed a quadratic time trend in column (2). Our main findings remain robust. In columns (3) - (4), we replaced $Resolution_{i,t-1}$ with the alternative measure, $CustomerGratitude_{i,t-1}$, for service effectiveness. Under different specifications, the estimates remain qualitatively the same, demonstrating the robustness of the results.

Instrumental Variable Analysis

Despite a wide range of controls, a remaining threat to the identification is unobserved, time-varying factors that affect the sensitivity of customer complaints to brand service efforts. To alleviate the concern that the findings may be driven by spurious correlation rather than causality, we applied an instrumental variable (IV) analysis. Specifically, we constructed the IV $\Delta NeuVoice_{i,t-1}$ for $logReplies_{i,t-1}$, which represent lagged changes in the composition of customer voices directed to firm i:

$$NeuRatio_{i,t} = \frac{Neutral\ Customer\ Voices_{i,t}}{All\ Customer\ Voices_{i,t}}$$

$$\Delta NeuVoice_{i,t-1} = NeuRatio_{i,t-1} - NeuRatio_{i,t-2}$$

The underlying assumption is that $\Delta NeuVoice_{i,t-1}$ affects customer complaints ($logComplaints_{i,t}$) only through brand service volume ($logReplies_{i,t-1}$). The logic is as follows: given the limited servers (i.e., number of agents), any shocks from the demand side requests (i.e., customer voices composition) will lead to the re-allocation of agent resources (i.e., brand replies to customer voices of different types). Since $\Delta NeuVoice_{i,t-1}$ reflects the changes in neutral voices, it directly affects the level of brand replies to complaints ($logReplies_{i,t}$). In the meantime, as neutral voices correspond to informational inquiries, they are unlikely to affect future customer complaints.

The F-test and Stock-Yogo weak IV test (see Appendix A, Table A2) indicate the strong relevance of IV to the endogenous variable. Table 3 reports the estimation results using the STATA command XTIVREG2 with autocorrelation-robust standard errors. As expected, the estimate for the IV was negative and significant in the first stage, suggesting that an increase in customers' neutral voices (i.e., informational inquiries) negatively affected a firm's resources allocated to complaints ($logReplies_{i,t}$). In the second stage, the coefficient estimates of the service volume and service quality were consistent with the baseline results, reconfirming the robustness of the findings.

Table 3. Instrumental Variable Analysis

Variables	(1)	(2)	(3)	(4)
Service Volume				
$logReplies_{i,t-1}$	0.1574* (0.071)	0.1573* (0.070)	0.1533* (0.070)	0.1532* (0.070)
Service Quality				
$Delay_{i,t-1}$	0.0316*** (0.005)	0.0299*** (0.005)	0.0314*** (0.005)	0.0296*** (0.005)
$Resolution_{i,t-1}$	-0.103*** (0.027)	-0.102*** (0.027)		
$CustomerGratitude_{i,t-1}$			-0.119* (0.052)	-0.122* (0.051)
Brand Control				
$OfflineIncidents_{i,t}$	6.442*** (0.873)	6.396*** (0.877)	6.468*** (0.873)	6.421*** (0.877)
$GoogleTrend_{i,t}$	0.417*** (0.054)	0.413*** (0.055)	0.420*** (0.054)	0.415*** (0.055)
$logFollowers_{i,t}$	0.3009*** (0.040)	0.2538*** (0.053)	0.3005*** (0.040)	0.2526*** (0.053)
1st-stage Results				
$\Delta NeuVoice_{i,t-1}$	-0.8413*** (0.076)	-0.8420*** (0.077)	-0.8512*** (0.077)	-0.8516*** (0.078)
Linear Time Trend	Y	Y	Y	Y
Quadratic Time Trend	N	Y	N	Y
R^2	0.275	0.276	0.272	0.273

Note. *** $p < 0.001$ ** $p < 0.01$ * $p < 0.05, + p < 0.1$. This table reports the results from the IV analysis corresponding to different specifications. The dependent variable is $logComplaints_{i,t}$. The estimation model is fixed-effects linear model with autocorrelation-robust standard errors, which was implemented using the STATA procedure XTIVREG2. *OfflineIncidents* and *GoogleTrend* were rescaled by a factor of 1/100 to allow for more non-zero digits in the estimation. Similarly, IV is rescaled from percentage to fraction to allow for more non-zero digits in the estimation. Throughout the analyses, we have accounted for other controls, and firm, year, month fixed effects.

ROBUSTNESS CHECKS

To further alleviate the endogeneity concern and evaluate the validity of our main findings, we dedicate this section to robustness checks. We start with estimating the main model with an alternative method, negative binomial regression, and then perform a subsample analysis with only airlines operating in English-speaking countries. To alleviate concerns about unobserved offline confounding factors, we further control for airlines' actual offline service performance. Moreover, we use different intervals to construct the lagged independent variables to examine how sensitive the results are to the variable construction.

Alternative Model for Limited Dependent Variable

In the main analysis, we log-transformed the volume of service complaints as the dependent variable. Considering that the untransformed outcome variable represents over-dispersed count data, we performed additional analysis to check if our findings are sensitive to models that account for the limited dependent variable. More specifically, we adopted the negative binomial regression and account for AR(1) disturbance in estimating the error term. We implemented the analysis using the STATA procedure XTGEE, specifying *family(nbinomial)* and *corr(AR1)*. As shown in Table A3 of Appendix A, all results remain robust.

Subsample Analysis

In the main analysis, we include all major international airlines, the majority of which are designed to address service requests from English-speaking customers. However, several airlines from non-English speaking countries may have received non-English tweets. Before we constructed the conversation-level data set, we first used the Google Translate API to translate non-English tweets into English. Although Google's API is the best tool that is available, one might worry that the translation inaccuracy could cause systematic bias to our findings, especially considering that several key service quality variables (e.g., resolution and agents' reply quality) are based on classifiers trained by English tweets.

To alleviate such a concern, we conducted a robustness check by excluding airline accounts with non-English tweets from the sample. More specifically, we excluded Avianca, Lufthansa, Air France, Garuda Indonesia, Brussels Airlines, Swiss International Air Lines, KLM, Philippine Airlines, SAS Scandinavian Airlines, and Thai Airways from the main sample. Table A4 of Appendix A reports the estimation results on this subsample, which are qualitatively the same as the main results, thereby validating our main findings.

Offline Service Quality

When estimating the causal effect of brand service interventions on customer complaints, we accounted for multiple factors that are likely endogenous, including a firm's time-varying customer base on Twitter (*logFollowers*) and shocks to brand service quality (*OfflineIncidents* and *GoogleTrend*). We also included brand and time fixed effects throughout the analyses to control for unobserved *time-invariant* firm heterogeneities and *common* time trends.

Despite all these efforts, the main findings may still suffer from potential omitted variable bias if there are any time-varying shocks to an airline's actual customer size or offline service quality. For example, if there is a quality deterioration in offline service provisions, such as more delayed departures, mishandled baggage, or unexpected flight cancellations, we will observe a spike in both customer complaints volume and brand service volume. Similarly, the massive influx of customer complaints may overwhelm agents and negatively affect their service effectiveness beyond what has been captured in our current service quality controls. In such situations, the identification of the relationship between service volume and quality on customer complaints will be biased. To alleviate concerns due to time-varying omitted variables, we managed to collect airlines' offline service performance data from the ATCR. As a further check on the validity of our findings, we performed analyses using the U.S. sample, with additional controls for each airline's time-varying volume of enplaned passengers, delayed flights, flight cancellations, and mishandled baggage. As shown in Table A5 of Appendix A, all results remained consistent with the main results, which supports the robustness of our findings to omitted variable bias.

Choice of Intervals in Independent Variable Construction

In the main analysis, we use one-week lag as explanatory variables. Considering that brand service interventions may take longer to take effects, we performed a robustness check using different intervals to construct the lag term. More specifically, for time *t*, we calculated the independent variables as the moving average of the two-weeks lagged data (Appendix A Table A6, columns 1 & 2) and the four-weeks lagged data (Appendix A Table A6, columns 3 & 4). All results are consistent with our main analysis, suggesting that our findings are not sensitive to the choice of time intervals in constructing the service volume measure.

MECHANISM TESTS FOR THE AWARENESS ENHANCEMENT EFFECT

To understand why brand service intervention affects customers' complaints, we dedicate this section to mechanism tests. We first examined the customers' composition and checked how customers' voicing propensities shift depending on brand service interventions. We then empirically tested the awareness enhancement mechanism.

Chronic and Non-Chronic Complainers

To shed light on whether active service intervention encourages chronic complaining, we traced each individual-initiated service complaints to all the airlines' Twitter accounts during our study period. Considering that a customer's complaint propensity could shift over time, we constructed the measure in a "rolling" fashion. More specifically, we define customer *i* as a "chronic complainer" if her total complaints in the past 12 months (i.e., $t-11$, $t-10$, $t-9$, ..., $t-1$, t) is greater than a specified cutoff. Following this definition, we decomposed complainers by type and summarize the composition in Table 4. If we specified the cutoff as one, 11.02% of customers were chronic complainers, contributing 19.27% of the complaints. As we increased the cutoff, we saw a significant reduction in both the ratio of chronic complainers and the number of complaints from the chronic complainers. Specifically, when the cutoff is 3, 0.92% of customers were considered chronic complainers, contributing 4.01% of the complaints. When the cutoff increased to 5, 0.19% of customers were considered chronic complainers, contributing 1.81% of the complaints. Overall, the results suggest that chronic complainers are not the primary driving forces in the overall complaints.

Table 4. Composition of Complainers and Complaints

Cutoff	Composition of Complainers			Composition of Complaints		
	#Complainer	#Chronic	%Chronic	#Complaint	#Chronic	%Chronic
1	1,387,438	152,833	11.02%	2,138,447	412,084	19.27%
3	1,387,438	12,770	0.92%	2,138,447	85,718	4.01%
5	1,387,438	2,586	0.19%	2,138,447	38,708	1.81%

Note. This table reports composition of complainers and complaints with different cutoff values in defining chronic complainers.

To examine how individual customers' complaint propensities reacted to brand service interventions, we conducted a nonparametric analysis. For each customer-airline combination, we created two monthly time series: (1) whether a customer initiated any complaints at time *t* ($isVoice_t \in \{0,1\}$); and (2) whether a customer received any brand replies to her requests at time *t* ($isReplied_t \in \{0,1\}$). For the k^{th} month forward ($t+k$), we then calculated the average customers' complaint propensity as the percentage of customers that complained, conditional on the brand reply status at time *t*: $\Pr[isVoice_{t+k} = 1 | isVoice_t = 1, isReplied_t = 0]$, $\Pr[isVoice_{t+k} = 1 | isVoice_t = 1, isReplied_t = 1]$.

We replicated this analysis by specifying different cutoffs when defining chronic complainers. Table 5 reports the results corresponding to the cutoff being 1. Tables A7 and A8 in Appendix A report the results corresponding to alternative cutoff values. Regardless of the parameter set-up, we can see that across different samples, customers' complaint propensities are not much affected by brand service interventions. For example, conditional on brand service interventions at t, the probability of complaining in the following month ($t + 1$) is 11.40% for chronic complainers who received brand replies, which is 0.81% (= 12.21% – 11.40%) smaller than those without a reply. For non-chronic complainers, the probability of complaining in the following month ($t + 1$) slightly increases by 0.21% (= 3.89% – 3.68%) when a brand replied to them at time t.

Table 5. Effect of Brand Reply on Customer Complaint Propensity (cutoff = 1)

Time	Chronic		Non-chronic		All	
	Without Reply	With Reply	Without Reply	With Reply	Without Reply	With Reply
$t + 1$	12.21%	11.40%	3.68%	3.89%	7.63%	7.29%
$t + 2$	10.80%	9.89%	2.92%	3.03%	6.60%	6.19%
$t + 3$	9.84%	9.22%	2.71%	2.86%	6.10%	5.84%
$t + 4$	9.31%	8.82%	2.53%	2.62%	5.75%	5.60%
$t + 5$	8.98%	8.52%	2.52%	2.64%	5.63%	5.44%
$t + 6$	8.76%	8.22%	2.33%	2.45%	5.48%	5.26%
$t + 7$	8.44%	8.03%	2.26%	2.50%	5.26%	5.19%
$t + 8$	8.18%	7.73%	2.22%	2.40%	5.20%	5.04%
$t + 9$	7.78%	7.51%	2.26%	2.40%	5.02%	4.90%
$t + 10$	7.66%	7.40%	2.18%	2.35%	4.94%	4.89%
$t + 11$	7.55%	7.24%	2.20%	2.41%	4.85%	4.80%
$t + 12$	7.59%	7.38%	2.44%	2.54%	5.01%	4.94%

Note. This table reports customers' complaint propensities in the following periods $t + k$, conditioned on voicing at time t.

We further conducted t-tests and formally examine whether complaint propensities significantly differ for customers who received brand replies versus those who did not receive brand responses. As shown in Table A9 of Appendix A, there is no significant difference in customers' complaint propensities *with* or *without* brand replies. Overall, the model-free results suggest customers' complaint propensities do not change much with (or without) brand service interventions. Therefore, the positive effect of service intervention on customer complaints is unlikely driven by the mechanism proposed by Ma et al. [31]. In contrast to the telecommunications firm in Ma et al. [31], airlines face a customer base that does not receive services very frequently. Hence, the mechanism of more service interventions driving more complaints from chronic complainers may not be practically significant for other industries, at least not for the airline industry. Consequently, a natural follow-up question is to examine the mechanism that drove our findings.

The Awareness Enhancement Effect

Online social ties and interactions can affect individuals' behaviors from various aspects, such as the diffusion of YouTube videos [41], peer music consumption [14], participation in charitable social movements [42], and individual funding behavior [45]. Similarly, online social ties can affect customer complaint behavior in the context of social media customer service. Each time a firm responds to a service inquiry on social media, customers who are connected to the focal redress-seeking customer can observe the service intervention. Previous literature has shown that customers are more likely to seek redress if it is clear that the seller is willing to remedy the problem [10]. Further, customers who enter a complaint situation, knowing how fellow customers have been treated in similar circumstances, are likely to expect similar treatment [43]. Accordingly, brands' active service responses combined with customers' social ties will enhance the overall awareness of firms' service availability, thereby driving more customers to seek redress via social media.

To empirically test this *awareness enhancement effect*, we took advantage of an obscure technological nuance of Twitter [49]. As shown in Table 6, on Twitter, one can start a conversation with an airline via two types of posts, and the probability that the customer's tweet will appear in her followers' home timeline depends on the position of the "@" symbol. Specifically, a *mention* (*M*) is a tweet that contains another user's Twitter handle, preceded by the "@" symbol. If a customer posts a complaint in such format, "I need help with ... @Delta ...", anyone on Twitter who is following the customer will see the tweet in their home timeline. A *reply* (*R*) is similar to a mention, but the tweet begins with "@username". If a customer posts a complaint in this format, "@Delta. I need help with ...", only those who follow both the customer and Delta will see such tweets in their home timeline.

Table 6. Types of Tweets and Visibility

Type	Your Tweet	Who Sees It
Mention (*M*)	I need help with … @Delta …	You, Delta, and all of your followers.
Reply (*R*)	@Delta, I need help with …	You, Delta, and your followers who also follow @Delta.

Since the awareness effect depends on the number of potential viewers of a redress seeking tweet, we construct two service awareness measures at the firm-week level, taking into account different levels of publicity:

$$\begin{cases} ServiceAwarenessM_{i,t} = \sum_{j=1}^{N_{i,t}} logFollowers_{j,t} \times TweetsM_{j,t} \\ ServiceAwarenessR_{i,t} = \sum_{j=1}^{N_{i,t}} logFollowers_{j,t} \times TweetsR_{j,t} \end{cases}$$

where $logFollowers_{j,t}$ measures the log-transformed number of followers of customer *j* at week *t*, $TweetsM_{j,t}$ and $TweetsR_{j,t}$ correspond to the number of customer *j*'s initiated complaints (that were replied to by firm *i*) in the form of mention and reply. $N_{i,t}$ is the number of unique users who complained to firm *i* at week *t*. If the service awareness mechanism does exist, we expect to see a larger effect through $ServiceAwarenessM_{i,t}$ because such service interactions are directly visible to all potential audiences and can inform them of the existence of the new service channel.

The Marketing literature has documented that for memory decay and related reasons, past advertising may not be as effective as recent advertising [35]. Similarly, potential customers' awareness through friends' usage of social media customer service may decay over time. Accordingly, we constructed the service awareness stock to measure the cumulative customer's knowledge of firm *i*'s service availability on Twitter up to time *t* as follows:

$$\begin{cases} AwarenessStockM_{i,t} = \delta AwarenessStockM_{i,t-1} + ServiceAwarenessM_{i,t} \\ AwarenessStockR_{i,t} = \delta AwarenessStockR_{i,t-1} + ServiceAwarenessR_{i,t} \end{cases}$$

where δ is the parameter for memory attrition overtime and $\delta \in (0,1)$. The awareness stock depends on both cumulative brand awareness stock till the last period and the service awareness in the current period. Table 7 reports the regression results when $\delta = 0.6$, which correspond to the regression model with the lowest AIC and BIC, and the highest log-likelihood (see Table A10 for model comparison statistics).

As shown in the table, conditional on the service efforts made for focal customers, only the coefficient estimates for $AwarenessStockM_{i,t}$ were consistently significant throughout the different specifications.

Table 7. Mechanism Test on Service Awareness

Variables	**(1)**	**(2)**	**(3)**	**(4)**
Awareness Stock				
$AwarenessStockM_{i,t-1}$	0.039*** (0.011)	0.036** (0.011)	0.038** (0.011)	0.034** (0.011)
$AwarenessStockR_{i,t-1}$	0.012 (0.012)	0.013 (0.012)	0.014 (0.012)	0.015 (0.012)
$logReplies_{i,t-1}$	0.082*** (0.010)	0.089*** (0.010)	0.074*** (0.010)	0.081*** (0.010)
Service Quality				
$Delay_{i,t-1}$	0.015*** (0.004)	0.014*** (0.004)	0.015*** (0.004)	0.014*** (0.004)
$Resolution_{i,t-1}$	-0.0915***(0.016)	-0.0924*** (0.016)		
$CustomerGratitude_{i,t-1}$			-0.130*** (0.039)	-0.131*** (0.039)
Brand Control				
$OfflineIncidents_{i,t}$	4.468*** (0.760)	4.463*** (0.761)	4.451*** (0.759)	4.449*** (0.760)
$GoogleTrend_{i,t}$	0.529*** (0.040)	0.522*** (0.040)	0.537*** (0.041)	0.530*** (0.040)
$logFollowers_{i,t}$	0.256*** (0.024)	0.204*** (0.028)	0.257*** (0.024)	0.206*** (0.028)
Linear Time Trend	Y	Y	Y	Y
Quadratic Time Trend	N	Y	N	Y
R^2	0.159	0.162	0.156	0.159
ρ_{ar}	0.311	0.306	0.317	0.312

Note. *** $p < 0.001$ ** $p < 0.01$ * $p < 0.05, + p < 0.1$. This table reports the test on the service awareness enhancement mechanism. The number of observations is 11,544. The number of observations reduces by 40 because of the panel-by-panel Cochrane–Orcutt transformation. The dependent variable is $logComplaints_{i,t}$. The estimation model is fixed-effects linear model with the first-order autocorrelation AR(1) disturbance across panels. *OfflineIncidents* and *GoogleTrend* were rescaled by a factor of 1/100 to allow for more non-zero digits in the estimation. Throughout the analyses, we have accounted for other controls, and firm, year, month fixed effects.

The findings support the awareness enhancement mechanism, implying that increasing awareness of firms' service availability through higher service efforts is an underlying driver for the increasing volume of customer complaints. Furthermore, with the additional control of *AwarenessStock*, the estimated coefficient of *logReplies* decreased significantly compared with the baseline results. Taking column 1 of the table as an example, the coefficient changes from 0.132 (cf. Table 2, column 1) to 0.082, corresponding to about a 38% reduction in the magnitude of the effect. The results further corroborate that the awareness indeed mediated a non-trivial proportion of the effect of service volume on service complaints.

As time goes by, more and more customers become aware of the social media customer service channel. If their memory is persistent, the effect of awareness shall decay over time. At the time when everyone becomes fully aware of the service channel, the awareness effect shall be gone. To test if there is a time-decaying awareness effect, we categorize a firm into different stages based on the duration since its provision of social media customer service and incorporate the interaction term between awareness stock and the stage dummies. More specifically, we first identify the time when an airline started providing social media customer service, to customer-initiated service requests as its SMCS adoption time, $adoptionTime_i$, which is before the starting of our sample period (i.e., before 2014). Next, for firm *i* at time *t*, we calculate the time since its adoption as ($t - AdoptionTime_i$). We then create different binary variables corresponding to the first, second, and third stages (i.e., *SMCS_Stage1*, *SMCS_Stage2*, and *SMCS_Stage3*) based on the terciles of ($t - AdoptionTime_i$).[8]

As reported in Table 8, the baseline effects of $AwarenessStockM_{i,t-1}$ were robust to different specifications also.

Table 8. Test of Service Awareness Effect over Time

Variables	**(1)**	**(2)**	**(3)**	**(4)**
Service Volume				
$AwarenessStockM_{i,t-1}$	0.042* (0.017)	0.054** (0.017)	0.041* (0.017)	0.052** (0.017)
$SMCS\_Stage2$	0.205*** (0.050)	0.191*** (0.049)	0.211*** (0.050)	0.197*** (0.050)
$SMCS\_Stage3$	-0.124* (0.063)	-0.121+ (0.062)	-0.117+ (0.063)	-0.113+ (0.062)
$SMCS\_Stage2 \times AwarenessStockM_{i,t-1}$	0.013 (0.024)	-0.003 (0.023)	0.016 (0.024)	-0.0003 (0.023)
$SMCS\_Stage3 \times AwarenessStockM_{i,t-1}$	-0.040+ (0.024)	-0.060* (0.024)	-0.032 (0.024)	-0.052* (0.024)
$AwarenessStockR_{i,t-1}$	0.024 (0.016)	0.009 (0.016)	0.025 (0.016)	0.010 (0.016)
$SMCS\_Stage2 \times AwarenessStockR_{i,t-1}$	-0.019 (0.024)	-0.007 (0.024)	-0.022 (0.024)	-0.010 (0.024)
$SMCS\_Stage3 \times AwarenessStockR_{i,t-1}$	0.066** (0.024)	0.082*** (0.024)	0.057* (0.024)	0.073** (0.024)
$logReplies_{i,t-1}$	0.098*** (0.011)	0.113*** (0.011)	0.090*** (0.011)	0.106*** (0.011)
Linear Time Trend	Y	Y	Y	Y
Quadratic Time Trend	N	Y	N	Y
R^2	0.150	0.160	0.146	0.156
ρ_{ar}	0.322	0.310	0.328	0.316

Note. *** $p < 0.001$ ** $p < 0.01$ * $p < 0.05, + p < 0.1$. This table reports the test results on the time-varying service awareness enhancement effect. We categorized each firm into different stages according to the distribution of the weeks since the initiation of its customer service practice on social media. $SMCS\_Stage1$ (i.e., the initial adoption stage) is the baseline. The number of observations is 11,544. The number of observations was reduced by 40 because of the panel-by-panel Cochrane–Orcutt transformation. The dependent variable is $logComplaints_{i,t}$. The estimation was done with a fixed-effects linear model with the first-order autocorrelation AR(1) disturbance across panels. Throughout the analyses, we accounted for other controls, and firm, year, month fixed effects.

The coefficient estimates were positive and significant for *SMCS_Stage2* and slightly negative for *SMCS_Stage3*, consistent with the general concave growth pattern of social networks. The coefficient of $SMCS\_Stage2 \times AwarenessStockM_{i,t-1}$ was not significant and the coefficient of $SMCS\_Stage3 \times AwarenessStockM_{i,t-1}$ was slightly negative. When we estimated the interaction effect of $SMCS\_Stage3 \times AwarenessStockR_{i,t-1}$, the coefficient was positive and significant. Overall, the results suggest that the evidence for the decaying awareness enhancement effect is unclear. There are two possible reasons for the finding. First, people tend to have shorter attention spans and memories due to the rapid-fire nature of news cycles on the Internet [53]. Therefore, customers' complaints can be spontaneously triggered by their friends. Second, the insignificant decaying pattern suggests that by 2019, airlines' customer service practice on social media was still not common knowledge to everyone. This is important to managers who have not considered the practice of customer service, given the potential massive online market and the low cost of customer service provision on social media.

CONCLUSION

Utilizing a unique data set with rich conversational details for each service interaction, we investigated the dynamics between brand service interventions and customer complaints on social media. We found that a higher service volume indeed caused more customer complaints on social media, but a higher customer service quality reduced, rather than encouraged future customer complaints. After accounting for various confounding factors and conducting multiple tests, we found that the awareness enhancement mechanism is an important driver of the positive impact of service volume on service complaints.

This paper extends the literature that studies the effect of managerial response on customers' reviewing behavior. Compared to traditional product or service review platforms, social media-based customer service goes a step further, by not only allowing customers to communicate with firms but also enabling firms to directly offer service support that is similar to what one receives from conventional service channels. In particular, this paper advances the nascent literature on the relationship between brand service intervention and customer voices on social media. The unique social network aspect amplifies each service intervention, making the service awareness mechanism possible and unique for customer service via social media. This mechanism clarifies the misconception that service intervention increases an individual's tendency to vent excessively, and therefore offers a crucial piece of knowledge for researchers and practitioners who are concerned about the abusive use of the channel by disgruntled customers.

This paper provides valuable insights to practitioners, especially for managers who strive to harness the power of social media for customer service. First, our insight on the awareness enhancement effect provides managers with a more complete picture when they consider how to devise effective social media strategies for customer service. Following Hirschman's exit, voice, and loyalty theory, a consumer's voice associates with a lower probability of exiting a firm's business compared to no voice. Viewing the underlying mechanism in terms of service awareness, firms should thus be less concerned about chronic complainers and adopt a more proactive customer service strategy instead. If many dissatisfied customers who would otherwise not bother contacting firms via traditional call centers end up seeking support through the social media channel, there should be less customer defection for the firms in the long run. One caveat of this recommendation is that during the launch of a new product, an overwhelming volume of social media complaints may prevent the firm from accumulating enough customers in the first place.

Second, it is crucial to bear in mind that the amplification effect can go either way. Service failures tend to grab more attention than service successes because negative information is often considered by consumers as more informative than positive information when they form their overall impressions of a brand [2]. Thus, just as in traditional complaint management, investment in service quality is essential to avoid an otherwise sound social media strategy from backfiring.

Finally, for the service awareness mechanism to work, service interventions must be observable to the public, or at least to those connected with the customer through online social network. Hence, our work should have external validity for any social media site that offers a public platform with a networking structure and a substantial audience, such as Facebook and Instagram.

Our work has several limitations that bear noting and offers opportunities for future research. First, the current study does not examine the direct impact of social media service intervention on firms' revenue or customer base. Future research may look at these important aspects and better guide firms to efficiently allocate their limited resources. Second, although previous literature suggested that customers prefer voicing complaints publicly [21], some customers may choose to contact firms via Twitter's direct messages because of privacy concerns. Our data does not include such private complaints which are akin to those complaints delivered through traditional channels. It would be interesting to explore how the public and private channels might affect each other when such data becomes available to researchers. Third, future studies may look at more nuanced aspects within the context of social media-based customer services, such as the different types of complaints.

ACKNOWLEDGMENTS

An early version of this study was previously circulated as "Chronic Complainers or Increased Awareness? The Dynamics of Social Media Customer Service," and appeared in the Strategy, Information, Technology, Economics, and Society (SITES) Mini-Track at the 2021 Hawaii International Conference on Systems Science (HICSS). The authors sincerely thank the Guest Editors, Robert Kauffman and Thomas Weber, as well as the anonymous reviewers, for their many insightful and constructive suggestions. All remaining errors are the authors' own.

NOTES

[1] McKinsey & Company reported that trained customer service agents can handle four to eight times the number of contacts received through social media as they can by phone; for more details, see [32].
[2] For example, Twitter has changed its interface design (see [27]) and provided different ways of connecting conversations (see [48]).
[3] 80% of social media customer service requests come from Twitter according to [3].
[4] We collected airlines' offline events from AeroInside (see [1]), which provides detailed reports about airline incidents, accidents, and crashes.
[5] Since the raw volume measures were highly skewed within and across firms, the log-transformed measures enable better interpretation and easier comparison of the estimates across firms.
[6] Some firms have a shorter time window due to missing data. For example, Virgin America was integrated into Alaska Airlines on April 24, 2018. Jet Airways ceased all flight operations in April 2019. Although their Twitter accounts still had rare activity after the event, we excluded these observations from the analysis.
[7] We also constructed an alternative measure, *Delay*, as the average delay in brand responses to customers within a service intervention. All results remain qualitatively the same and are available upon request.
[8] We tried different cutoffs (i.e., 30%, 70% and 20%, 80%) to distinguish among *SMCS_Stage1, SMCS_Stage2, and SMCS_Stage3*. The results remained consistent and are available upon request.

REFERENCES

1. AeroInside. Aviation incidents. https://www.aeroinside.com/incidents (last accessed on April 17, 2021).
2. Ahluwalia, R., Burnkrant, R.E., and Unnava, H.R. Consumer response to negative publicity: The moderating role of commitment. *Journal of Marketing Research*, *37*, 2 (May 2000), 203–214.
3. Akik. Making customer service even better on Twitter. 2016. https://blog.twitter.com/en_us/a/2016/making-customer-service-even-better-on-twitter.html (last accessed on April 17, 2021).
4. Alcantara, A.M. Customer complaints, and their ways of complaining, are on the rise. *The Wall Street Journal*, 2020. https://www.wsj.com/articles/customer-complaints-and-their-ways-of-complaining-are-on-the-rise-11591998939 (last accessed on April 17, 2021).
5. Baltagi, B.H. and Wu, P.X. Unequally spaced panel data regressions with AR (1) disturbances. *Econometric Theory*, *15*, 6 (December 1999), 814–823.
6. Bardhan, I.R., Demirkan, H., Kannan, P.K., Kauffman, R.J., and Sougstad, R. An Interdisciplinary Perspective on IT Services Management and Service Science. *Journal of Management Information Systems*, *26*, 4 (April 2010), 13–64.
7. Beck, N. and Katz, J.N. What to do (and not to do) with time-series cross-section data. *American Political Science Review*, *89*, 3 (September 1995), 634–647.
8. Berger, J. and Schwartz, E.M. What drives immediate and ongoing word of mouth? *Journal of Marketing Research*, *48*, 5 (October 2011), 869–880.
9. Bhargava, A., Franzini, L., and Narendranathan, W. Serial correlation and the fixed effects model. *The Review of Economic Studies*, *49*, 4 (October 1982), 533–549.
10. Blodgett, J.G., Wakefield, K.L., and Barnes, J.H. The effects of customer service on consumer complaining behavior. *Journal of Services Marketing*, *9*, 4 (October 1995), 31–42.
11. Bolton, R.N., Lemon, K.N., and Bramlett, M.D. The effect of service experiences over time on a supplier's retention of business customers. *Management Science*, *52*, 12 (December 2006), 1811–1823.
12. Cameron, A.C. and Trivedi, P.K. Microeconometrics using Stata. *Indicator*, *2*, (January 2010), 47.
13. Chevalier, J.A., Dover, Y., and Mayzlin, D. Channels of Impact: User reviews when quality is dynamic and managers respond. *Marketing Science*, *37*, 5 (September 2018), 688–709.
14. Dewan, S., Ho, Y.-J., and Ramaprasad, J. Popularity or proximity: Characterizing the nature of social influence in an online music community. *Information Systems Research*, *28*, 1 (March 2017), 117–136.
15. Drukker, D.M. Testing for serial correlation in linear panel-data models. *The Stata Journal*, *3*, 2 (June 2003), 168–177.

16. Elrhoul, M. Research: Four ways brands can build customer service relationships on Twitter. 2015. https://blog.twitter.com/en_us/a/2015/research-four-ways-brands-can-build-customer-service-relationships-on-twitter.html (last accessed on April 17, 2021).
17. Fornell, C. and Wernerfelt, B. A model for customer complaint management. *Marketing Science*, *7*, 3 (August 1988), 287–298.
18. Gu, B. and Ye, Q. First step in social media: Measuring the influence of online management responses on customer satisfaction. *Production and Operations Management*, *23*, 4 (April 2014), 570–582.
19. Gunarathne, P., Rui, H., and Seidmann, A. Whose and what social media complaints have happier resolutions? Evidence from Twitter. *Journal of Management Information Systems*, *34*, 2 (April 2017), 314–340.
20. Gunarathne, P., Rui, H., and Seidmann, A. When social media delivers customer service: Differential customer treatment in the airline industry. *MIS Quarterly*, *42*, 2 (June 2018), 489–520.
21. He, S., Lee, S.-Y., and Rui, H. Open voice or private message? The hidden tug-of-war on social media customer service. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*. 2019.
22. Hennig-Thurau, T., Gwinner, K.P., Walsh, G., and Gremler, D.D. Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? *Journal of Interactive Marketing*, *18*, 1 (January 2004), 38–52.
23. Hirschman, A.O. *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states*. Harvard University Press, 1970.
24. Hoechle, D. Robust standard errors for panel regressions with cross-sectional dependence. *The Stata Journal*, *7*, 3 (September 2007), 281–312.
25. Hu, Y., Tafti, A., and Gal, D. Read this, please? The role of politeness in customer service engagement on social media. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*. 2019.
26. Huang, W. Study: Twitter customer care increases willingness to pay. 2016. https://blog.twitter.com/en_us/topics/insights/2016/study-twitter-customer-care-increases-willingness-to-pay-across-industries.html (last accessed on April 17, 2021).
27. Ian Cairns. Making conversations easier to follow on the Tweet page. 2015. https://blog.twitter.com/en_us/a/2015/making-conversations-easier-to-follow-on-the-tweet-page.html (last accessed on April 17, 2021).
28. Knox, G. and Van Oest, R. Customer complaints and recovery effectiveness: A customer base approach. *Journal of Marketing*, *78*, 5 (September 2014), 42–57.
29. Lehrer, C., Wieneke, A., vom Brocke, J., Jung, R., and Seidel, S. How big data analytics enables service innovation: Materiality, affordance, and the individualization of service. *Journal of Management Information Systems*, *35*, 2 (April 2018), 424–460.
30. Lovett, M.J., Peres, R., and Shachar, R. On brands and word of mouth. *Journal of Marketing Research*, *50*, 4 (August 2013), 427–444.
31. Ma, L., Sun, B., and Kekre, S. The squeaky wheel gets the grease—An empirical analysis of customer voice and firm intervention on Twitter. *Marketing Science*, *34*, 5 (September 2015), 627–645.
32. Masri, M., Esber, D., Sarrazin, H., and Singer, M. Social care in the world of "now." 2015. https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/social-care-in-the-world-of-now# (last accessed on April 17, 2021).
33. Melancon, J.P. and Dalakas, V. Consumer social voice in the age of social media: Segmentation profiles and relationship marketing strategies. *Business Horizons*, *61*, 1 (January 2018), 157–167.
34. Mousavi, R., Johar, M., and Mookerjee, V.S. The voice of the customer: Managing customer care in Twitter. *Information Systems Research*, *31*, 2 (June 2020), 340–360.
35. Naik, P.A., Mantrala, M.K., and Sawyer, A.G. Planning media schedules in the presence of dynamic advertising quality. *Marketing Science*, *17*, 3 (August 1998), 214–235.
36. News 18. Twitter replies to customer grievances could trigger more complaints. 2015. https://www.news18.com/news/tech/twitter-replies-to-customer-grievances-could-trigger-more-complaints-study-1034870.html (last accessed on April 17, 2021).

37. Proserpio, D. and Zervas, G. Online reputation management: Estimating the impact of management responses on consumer reviews. *Marketing Science*, *36*, 5 (September 2017), 645–665.
38. ScienceDaily. Tweeting responses to complaints on social media triggers new complaints. 2015. https://www.sciencedaily.com/releases/2015/08/150806144559.htm (last accessed on April 17, 2021).
39. Sprout Social. The Sprout Social Index, Edition XII: Call-out culture. Blog post, *Sprout Social*. https://sproutsocial.com/insights/data/q3-2017/ (last accessed on April 17, 2021).
40. Statista Research Department. Expected response time for social media questions or complaints in U.S. & global 2018. 2019. https://www.statista.com/statistics/808477/expected-response-time-for-social-media-questions-or-complaints/ (last accessed on April 17, 2021).
41. Susarla, A., Oh, J.-H., and Tan, Y. Social networks and the diffusion of user-generated content: Evidence from YouTube. *Information Systems Research*, *23*, 1 (March 2012), 23–41.
42. Tan, X., Lu, Y., and Tan, Y. The impact of subscription reciprocity on charitable content creation and sharing: Evidence from Twitter on Giving Tuesday. *MIS Quarterly, Forthcoming*, (March 2021).
43. Tax, S.S., Brown, S.W., and Chandrashekaran, M. Customer evaluations of service complaint experiences: implications for relationship marketing. *Journal of Marketing*, *62*, 2 (April 1998), 60–76.
44. The Economic Times. Social media customer service may trigger chain of complaints. 2015. https://economictimes.indiatimes.com/news/international/business/social-media-customer-service-may-trigger-chain-of-complaints/articleshow/48411922.cms (last accessed on April 17, 2021).
45. Thies, F., Wessel, M., and Benlian, A. Effects of social interaction dynamics on platforms. *Journal of Management Information Systems*, *33*, 3 (July 2016), 843–873.
46. Toubia, O. and Stephen, A.T. Intrinsic vs. image-related utility in social media: Why do people contribute content to twitter? *Marketing Science*, *32*, 3 (May 2013), 368–392.
47. Turel, O., Yuan, Y., and Connelly, C.E. In justice we trust: Predicting user acceptance of E-Customer services. *Journal of Management Information Systems*, *24*, 4 (April 2008), 123–151.
48. Twitter Help Center. How to create a thread on Twitter and how to view. https://help.twitter.com/en/using-twitter/create-a-thread (last accessed on April 17, 2021).
49. Twitter Help Center. About different types of Tweets. https://help.twitter.com/en/using-twitter/types-of-tweets (last accessed on April 17, 2021).
50. United States Securities and Exchange Commission. Delta Air Lines, Inc. - Financials. 2019. https://d18rn0p25nwr6d.cloudfront.net/CIK-0000027904/5fcae838-aa00-4be0-95dd-1824e4f97799.pdf (last accessed on April 17, 2021).
51. Wang, Y. and Chaudhry, A. When and how managers' responses to online reviews affect subsequent reviews. *Journal of Marketing Research*, *55*, 2 (April 2018), 163–177.
52. Wooldridge, J.M. *Econometric Analysis of Cross Section and Panel Data*. MIT press, 2002.
53. Wright, M. and Zolfagharifard, E. Internet is giving us shorter attention spans and worse memories, major study suggests. *The Telegraph*, 2019. https://www.telegraph.co.uk/technology/2019/06/06/internet-giving-us-shorter-attention-spans-worse-memories-major/ (last accessed on April 17, 2021).
54. Xue, M., Hitt, L.M., and Harker, P.T. Customer efficiency, channel usage, and firm performance in retail banking. *Manufacturing & Service Operations Management*, *9*, 4 (October 2007), 535–558.
55. Yang, M., Zheng, Z., and Mookerjee, V. Prescribing response strategies to manage customer opinions: A stochastic differential equation approach. *Information Systems Research*, *30*, 2 (June 2019), 351–374.
56. Yeomans, M., Kantor, A., and Tingley, D. The politeness package: Detecting politeness in natural language. *R Journal*, *10*, 2 (December 2018), 489–502.

APPENDIX A. SUPPLEMENTARY ANALYSIS

Table A1. Sample of Firms in the Dataset

Airline (by continent)	Twitter Account	Country
Africa		
Kenya Airways	@KenyaAirways	Kenya
Kulula	@kulula	South Africa
South African Airways	@flysaa	South Africa
Americas		
Air Canada	@AirCanada	Canada
Alaska Airlines	@AlaskaAir	United States
American Airlines	@AmericanAir	United States
Avianca	@Avianca	Colombia
Hawaiian Airlines	@HawaiianAir	United States
JetBlue Airways	@JetBlue	United States
Southwest Airlines	@SouthwestAir	United States
United Airlines	@united	United States
Virgin America	@VirginAmerica	United States
Westjet	@WestJet	Canada
Asia		
AirAsia Berhad	@airasia	Malaysia
Cathay Pacific Airways	@cathaypacific	Hong Kong, China
Etihad Airways	@etihad	United Arab Emirates
Garuda Indonesia	@IndonesiaGaruda	Indonesia
IndiGo Airlines	@IndiGo6E	India
Jet Airways (India)	@jetairways	India
Malaysia Airlines	@MAS	Malaysia
Philippine Airlines	@flyPAL	Philippines
Singapore Airlines	@SingaporeAir	Singapore
Spicejet	@flyspicejet	India
Thai Airways	@ThaiAirways	Thailand
Europe		
Aer Lingus Irish Airlines	@AerLingus	Ireland
Air France	@airfrance	France
Austrian Airlines	@_austrian	Austria
Brussels Airlines	@FlyingBrussels	Belgium
Deutsche Lufthansa	@lufthansa	Germany
easyJet	@easyJet	United Kingdom
Icelandair	@Icelandair	Iceland
KLM	@KLM	Netherlands
SAS Scandinavian Airlines	@SAS	Sweden
Swiss International Air Lines	@FlySWISS	Switzerland
Virgin Atlantic Airways	@VirginAtlantic	United Kingdom
Air New Zealand	@FlyAirNZ	New Zealand
Jetstar Airways	@JetstarAirways	Australia
Qantas Airways	@Qantas	Australia
Tiger Australia	@TigerairAU	Australia
Virgin Australia	@VirginAustralia	Australia

Note. This table lists each airlines' most recent Twitter handle. As some firms changed Twitter handles over time, we identify each airline with its time-invariant Twitter metadata (i.e., account ID) during data collection and variable construction.

Table A2. IV Relevance Test

Test	(1)	(2)	(3)	(4)
F test of excluded instruments	123.90***	120.96***	121.79***	118.92***

Note. *** $p < 0.001$ ** $p < 0.01$ * $p < 0.05, + p < 0.1$. This table reports the IV relevance test corresponding to columns 1 - 4 of Table 3. The dependent variable is $logComplaints_{i,t}$. The Stock-Yogo Weak IV test critical values for 10%, 15%, 20%, and 25% maximal IV size are 16.28, 8.96, 6.66, and 5.53, respectively.

Table A3. Negative Binomial Model with AR(1) Disturbance

Variables	(1)	(2)	(3)	(4)
Service Volume				
$logReplies_{i,t-1}$	0.202*** (0.008)	0.207*** (0. 008)	0.197*** (0. 008)	0.202*** (0. 008)
Service Quality				
$Delay_{i,t-1}$	0.026*** (0.004)	0.024*** (0.004)	0.026*** (0.004)	0.024*** (0.004)
$Resolution_{i,t-1}$	-0.117*** (0.020)	-0.118*** (0.020)		
$CustomerGratitude_{i,t-1}$			-0.120* (0.050)	-0.124* (0.050)
Brand Control				
$OfflineIncidents_{i,t}$	9.361*** (0.890)	9.255*** (0.890)	9.375*** (0.892)	9.269*** (0.891)
$GoogleTrend_{i,t}$	0.527*** (0.040)	0.520*** (0.040)	0.530*** (0.040)	0.523*** (0.040)
$logFollowers_{i,t}$	0.287*** (0.025)	0.228*** (0.029)	0.286*** (0.025)	0.227*** (0.029)
Linear Time Trend	Y	Y	Y	Y
Quadratic Time Trend	N	Y	N	Y
ρ_{ar}	0.135	0.133	0.135	0.134

Note. *** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$ $+p < 0.1$. This table reports the results on the dynamics between brand service interventions and customer complaints. The number of observations is 11,584. The dependent variable is $logComplaints_{i,t}$. The estimation model is negative binomial model with the first-order autocorrelation AR(1) across panels. *OfflineIncidents* and *GoogleTrend* are rescaled by a factor of 1/100 to allow for more non-zero digits in the estimation. Throughout the analyses, we have accounted for other controls, and firm, year, month fixed effects.

Table A4. Subsample Analysis of English Twitter Accounts

Variables	(1)	(2)	(3)	(4)
Service Volume				
$logReplies_{i,t-1}$	0.093*** (0.009)	0.094*** (0.009)	0.089*** (0.009)	0.090*** (0.009)
Service Quality				
$Delay_{i,t-1}$	0.013** (0.005)	0.011* (0.005)	0.013** (0.005)	0.012** (0.005)
$Resolution_{i,t-1}$	-0.055** (0.019)	-0.055** (0.019)		
$CustomerGratitude_{i,t-1}$			-0.098* (0.042)	-0.097* (0.042)
Brand Control				
$OfflineIncidents_{i,t}$	5.852*** (0.861)	5.826*** (0.861)	5.860*** (0.860)	5.833*** (0.860)
$GoogleTrend_{i,t}$	0.512*** (0.047)	0.509*** (0.047)	0.516*** (0.048)	0.514*** (0.048)
$logFollowers_{i,t}$	0.345*** (0.027)	0.295*** (0.032)	0.347*** (0.027)	0.297*** (0.032)
Linear Time Trend	Y	Y	Y	Y
Quadratic Time Trend	N	Y	N	Y
R^2	0.150	0.152	0.149	0.151
ρ_{ar}	0.333	0.332	0.335	0.334

Note. *** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$ $+p < 0.1$. This table reports the results on the dynamics between brand service interventions and customer complaints. The number of observations is 8,585. The dependent variable is $logComplaints_{i,t}$. The estimation model is fixed-effects linear model with the first-order autocorrelation AR(1) disturbance across panels. *OfflineIncidents* and *GoogleTrend* are rescaled by a factor of 1/100 to allow for more non-zero digits in the estimation. Throughout the analyses, we have accounted for other controls, and firm, year, month fixed effects.

Table A5. U.S. Sample Analysis with Controls for Brand Offline Service Performance

Variables	(1)	(2)	(3)	(4)
Service Volume				
$logReplies_{i,t-1}$	0.094*** (0.019)	0.076*** (0.019)	0.093*** (0.019)	0.076*** (0.019)
Service Quality				
$Delay_{i,t-1}$	0.046*** (0.011)	0.053*** (0.011)	0.048*** (0.011)	0.054*** (0.011)
$Resolution_{i,t-1}$	-0.198* (0.100)	-0.192^{+} (0.099)		
$CustomerGratitude_{i,t-1}$			-0.245* (0.124)	-0.235^{+} (0.124)
Brand Control				
$OfflineIncidents_{i,t}$	4.681*** (1.036)	4.723*** (1.031)	4.697*** (1.036)	4.739*** (1.031)
$GoogleTrend_{i,t}$	0.363*** (0.079)	0.393*** (0.079)	0.362*** (0.079)	0.391*** (0.079)
$logFollowers_{i,t}$	0.462*** (0.074)	0.669*** (0.088)	0.466*** (0.074)	0.673*** (0.088)
ATCR Offline Service Performance				
$PassengerVolume_{i,t}$	0.004 (0.009)	0.005 (0.009)	0.005 (0.009)	0.005 (0.009)
$\%FlightDelay_{i,t}$	0.019*** (0.003)	0.018*** (0.003)	0.019*** (0.003)	0.018*** (0.003)
$\%FlightCancellation_{i,t}$	0.020^{+} (0.012)	0.017 (0.012)	0.020 (0.012)	0.016 (0.012)
$\%BaggageClaim_{i,t}$	-0.001 (0.017)	0.004 (0.017)	-0.002 (0.017)	0.004 (0.017)
Linear Time Trend	Y	Y	Y	Y
Quadratic Time Trend	N	Y	N	Y
R^2	0.261	0.268	0.261	0.268
ρ_{ar}	0.233	0.233	0.232	0.232

Note. *** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$ $+p < 0.1$. This table reports the results on the dynamics between brand service interventions and customer complaints. The number of observations reduces by 7 (from 1,979 to 1,972) because of the panel-by-panel Cochrane–Orcutt transformation for the seven U.S. airlines. The dependent variable is $logComplaints_{i,t}$. The estimation model is fixed-effects linear model with the first-order autocorrelation AR(1) disturbance across panels. *OfflineIncidents* and *GoogleTrend* are rescaled by a factor of 1/100 to allow for more non-zero digits in the estimation. Throughout the analyses, we have accounted for other controls, and firm, year, month fixed effects.

Table A6. Analysis Accounting for Alternative Time Lags in IDV

Variables	2-weeks Moving Average of IDV		4-weeks Moving Average of IDV	
	(1)	(2)	(3)	(4)
Service Volume				
$logReplies_{i,t-1}$	0.192*** (0.010)	0.186*** (0.010)	0.257*** (0.012)	0.251*** (0.012)
Service Quality				
$Delay_{i,t-1}$	0.022*** (0.006)	0.021*** (0.006)	0.028*** (0.007)	0.026*** (0.007)
$Resolution_{i,t-1}$	-0.143*** (0.026)		-0.203*** (0.037)	
$CustomerGratitude_{i,t-1}$		-0.181** (0.065)		-0.176* (0.084)
Brand Control				
$OfflineIncidents_{i,t}$	4.112*** (0.740)	4.116*** (0.740)	3.800*** (0.722)	3.790*** (0.722)
$GoogleTrend_{i,t}$	0.568*** (0.042)	0.575*** (0.042)	0.603*** (0.043)	0.612*** (0.043)
$logFollowers_{i,t}$	0.199*** (0.029)	0.200*** (0.029)	0.177*** (0.030)	0.180*** (0.030)
Linear Time Trend	Y	Y	Y	Y
Quadratic Time Trend	N	Y	N	Y
R^2	0.154	0.151	0.153	0.150
ρ_{ar}	0.356	0.359	0.356	0.359

Note. *** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$ $+p < 0.1$. This table reports the results on the dynamics between brand service interventions and customer complaints. IDV is short for independent variable. The number of observations is 11,544. The number of observations reduces by 40 because of the panel-by-panel Cochrane–Orcutt transformation. The dependent variable is $logComplaints_{i,t}$. The estimation model is fixed-effects linear model with the first-order autocorrelation AR(1) disturbance across panels. *OfflineIncidents* and *GoogleTrend* are rescaled by a factor of 1/100 to allow for more non-zero digits in the estimation. Throughout the analyses, we have accounted for other controls, and firm, year, month fixed effects.

Table A7. Effect of Brand Reply on Customer Complaint Propensity (cutoff = 3)

Time	Chronic		Non-chronic		All	
	Without Reply	With Reply	Without Reply	With Reply	Without Reply	With Reply
$t+1$	29.75%	25.57%	5.58%	5.71%	7.63%	7.29%
$t+2$	26.64%	23.30%	4.67%	4.66%	6.60%	6.19%
$t+3$	24.58%	22.21%	4.29%	4.38%	6.10%	5.84%
$t+4$	23.31%	20.57%	4.00%	4.20%	5.75%	5.60%
$t+5$	22.46%	20.38%	3.92%	4.07%	5.63%	5.44%
$t+6$	21.12%	19.50%	3.79%	3.88%	5.48%	5.26%
$t+7$	20.73%	18.60%	3.61%	3.87%	5.26%	5.19%
$t+8$	19.85%	17.74%	3.60%	3.77%	5.20%	5.04%
$t+9$	18.61%	17.42%	3.50%	3.63%	5.02%	4.90%
$t+10$	17.80%	17.05%	3.47%	3.62%	4.94%	4.89%
$t+11$	17.58%	16.55%	3.39%	3.53%	4.85%	4.80%
$t+12$	16.84%	16.61%	3.62%	3.70%	5.01%	4.94%

Note. Table reports customers' complaint propensities in the following periods $t+k$, conditioned on voicing at time t.

Table A8. Effect of Brand Reply on Customer Complaint Propensity (cutoff = 5)

Time	Chronic		Non-chronic		All	
	Without Reply	With Reply	Without Reply	With Reply	Without Reply	With Reply
$t+1$	49.69%	41.90%	6.50%	6.53%	7.63%	7.29%
$t+2$	43.87%	40.11%	5.54%	5.42%	6.60%	6.19%
$t+3$	42.15%	37.53%	5.07%	5.10%	6.10%	5.84%
$t+4$	40.01%	35.13%	4.76%	4.89%	5.75%	5.60%
$t+5$	39.98%	33.79%	4.66%	4.74%	5.63%	5.44%
$t+6$	36.99%	33.62%	4.51%	4.55%	5.48%	5.26%
$t+7$	36.05%	32.66%	4.32%	4.50%	5.26%	5.19%
$t+8$	34.63%	30.31%	4.31%	4.38%	5.20%	5.04%
$t+9$	33.41%	29.74%	4.14%	4.24%	5.02%	4.90%
$t+10$	30.00%	28.89%	4.11%	4.23%	4.94%	4.89%
$t+11$	28.97%	28.13%	4.04%	4.15%	4.85%	4.80%
$t+12$	29.15%	27.96%	4.23%	4.29%	5.01%	4.94%

Note. Table reports customers' complaint propensities in the following periods $t+k$, conditioned on voicing at time t.

Table A9. Test on the Difference in Customers' Complaint Propensities with (without) Brand Reply

Cutoff	Chronic			Non-chronic		
	Without Reply	With Reply	*p*-value	Without Reply	With Reply	*p*-value
1	8.38%	7.99%	0.565	2.41%	2.53%	0.539
3	20.10%	18.47%	0.368	3.76%	3.88%	0.676
5	34.40%	31.34%	0.318	4.44%	4.51%	0.833

Note. This table reports the t-tests on average customers' complaint propensities with or without brand replies, given different cutoffs (i.e., 1, 3, 5) in defining chronic complainers.

Table A10. Model Selection of the Memory Attrition Parameter

δ	Log-Likelihood	AIC	BIC
0.1	-6,256.77	12,583.53	12,840.92
0.2	-6,251.95	12,573.90	12,831.29
0.3	-6,246.07	12,562.14	12,819.53
0.4	-6,240.34	12,550.68	12,808.07
0.5	-6,235.87	12,541.75	12,799.13
0.6	**-6,233.75**	**12,537.49**	**12,794.88**
0.7	-6,235.08	12,540.16	12,797.55
0.8	-6,241.43	12,552.86	12,810.24
0.9	-6,254.82	12,579.63	12,837.02

Note. This table reports the model statistics for different choices of the memory attrition parameter δ, where $\delta = 0.6$ corresponds to the lowest AIC, BIC and the highest log-likelihood.

APPENDIX B. SENTIMENT ANALYSIS

Based on a reading of about 2,000 random tweets, Gunarathne et al. [20] developed a lexicon that contains 326 n-grams for complaint tweets and 354 n-grams for compliment tweets. A customer tweet is classified as a complaint if it matched at least one term in the complaint lexicon and none in the compliment lexicon. Gunarathne et al. also selected a random sample of 8,700 tweets from the predicted complaints and identified 7,354 tweets that are indeed complaints. In other words, they reported 84.5% precision for their lexicon-based complaint classifier.

In this study, we adopted the lexicon proposed by Gunarathne et al. [20] and implemented a bag-of-words approach to categorize customer-initiated tweets into complaints, compliments, and informational posts. More specifically, if a customer-initiated tweet matched at least one term in the complaint lexicon and none in the compliment lexicon, we considered it as a *complaint.* Similarly, if a customer-initiated tweet matched at least one term in the compliment lexicon and none in the complaint lexicon, we considered it as a *compliment.* For customer-initiated tweets that did not fall into any of these two categories, we consider those tweets as neutral or informational voices.

APPENDIX C – TEXT CLASSIFICATION FOR RESOLUTION

We hired an annotator to carefully read through 3,258 customer service-related conversations from Southwest Airlines and determine whether there was a resolution of the service intervention based on the following criteria.

- If a consumer expressed satisfaction or gratitude in the end, the outcome of the conversation was labeled as a *resolution*. Taking the 1st conversation listed in Table C1 as an example, the consumer explicitly appreciated the agent for updating the flight information, and thus the outcome of the conversation was a resolution. In contrast, the consumer in the second example conversation still thought the company was incompetent and unaccountable after multiple interactions with the agent, which indicated a service failure.
- If a consumer sent private messages to the airline, the outcome of the conversation was labeled as a *resolution*. The rationale is that private communications can reduce the risk of a negative externality when firms fail to address complaints publicly [21]. Therefore, a focal customer's agreement to communicate via private messages at least signals a "public" resolution to bystanders. The 3rd conversation listed in Table C1 is an example.
- If a consumer didn't explicitly express gratitude but the agent successfully addressed the question or complaint, then the outcome of the conversation was labeled as a *resolution*. Taking the 4th conversation listed in Table C1 for instance, the agent provided detailed information to answer the customer's question of why the flight was canceled.

For a small set of conversations that did not fall into any of the circumstances, the annotator followed two steps: made his judgment based on factors, such as consumers' sentiment changes and agents' efforts and reasoning; discussed with authors and determine the label according to all readers' votes. Given the labeled data set, we extract a list of binary features based on the preprocessed texts and then build a SVM classifier to identify if resolution was reached. The performance of the classifier is reported in Table C2.

Table C1. Sample Conversations for Labeling Criteria

ID	Role	Content	Resolution
1-1	Consumer	@SouthwestAir what's going on with Flight 40 out of Dallas to Chicago?	✓
1-2	Agent	@consumer Hey there, Courtney. It looks like inclement weather and associated Air Traffic Control delays have held your aircraft up, but we are working to get your flight to Chicago another aircraft ASAP. We hope to have you on your way soon. -Adrienne	
1-3	Consumer	@SouthwestAir On another aircraft? So, there's something wrong with the plane as well? I'm seeing 7:20 departure time	
1-4	Agent	@consumer Nope! Nothing wrong, just that it has been held up with delays, and we are hoping to get you an alternative plane to get you out sooner. We are currently anticipating an ETD of about 7:20pm. -Adrienne	
1-5	Consumer	@SouthwestAir Thanks for the update	
2-1	Consumer	@SouthwestAir you leave me stranded on St. Patricks day at an airport with nothing open. No offer of anything except the payment of my time. Sweet. Fun. Awesome	✗
2-2	Agent	@consumer We regret the disappointment today, Brent. We know that your time is precious, and we appreciate your patience today! -Nicole	
2-3	Consumer	@SouthwestAir Yeah that's not it. I'll use a different airline for now on. Thanks for your patience	
2-4	Consumer	@SouthwestAir Are you going to cover my costs for the extra $ it will cost me for this delay?	
2-5	Agent	@consumer While we don't cover interim expenses, you're welcome to reach out to us once your flight is completed so we can assist at that time. -Nicole	
2-6	Consumer	@SouthwestAir Airlines. The only industry that can waste your time, cost you $ and there is no accountability	
3-1	Consumer	@SouthwestAir I was charged for wifi on my phone, but I still can't get a connection for wifi on my phone. What should I do?	✓
3-2	Agent	@consumer Hey, Mister. Our apologies for the inconvenience. Please DM (direct message) the email address you used when you purchased the WiFi, and we'll follow-up. ^KD	
3-3	Consumer	@SouthwestAir I've sent the message	
4-1	Consumer	@SouthwestAir what gives? Cancelled flight in line at the airport and on infinite hold with customer service. #nothappy	✓
4-2	Agent	@consumer It's not our intention to disappoint you, Kim. Please speak to an Agent to get rebooked. We can refund the unused portion of your flight if you are able to find alternate transportation. Please DM if you need further assistance. ^SJ https://t.co/mQmfkXW4oV	
4-3	Consumer	@SouthwestAir Wondering why the delta flight at the same time is on time?	
4-4	Agent	@consumer Mind sending your flight details so we can look into this? ^SJ	
4-5	Consumer	@SouthwestAir 224 from ATL to MCI	
4-6	Agent	@consumer Thank for the additional information. Our records indicate Flight #224 is canceled due to weather. We cannot speak to other carriers. Our number one concern is operating a safe flight for our Customers and Employees. I am sorry for frustration this has caused. ^SJ	

Table C2. SVM Classifier Performance on Conversation Resolution Using 10-Fold Cross Validation

Fold #	Precision (0)	Recall (0)	F1(0)	Precision (1)	Recall (1)	F1 (1)
1	0.74	0.73	0.73	0.82	0.82	0.82
2	0.74	0.72	0.73	0.80	0.81	0.80
3	0.73	0.70	0.72	0.80	0.82	0.81
4	0.77	0.72	0.74	0.81	0.85	0.83
5	0.76	0.73	0.75	0.81	0.84	0.82
6	0.77	0.73	0.75	0.80	0.83	0.81
7	0.73	0.70	0.72	0.78	0.81	0.79
8	0.73	0.76	0.74	0.82	0.81	0.81
9	0.74	0.72	0.73	0.80	0.82	0.81
10	0.74	0.69	0.72	0.78	0.82	0.80
Average	**0.75**	**0.72**	**0.73**	**0.80**	**0.82**	**0.81**

Note. (0) represents conversations ended without resolution, and (1) represents conversations ended with a resolution.