**Trains and Twitter: Firm Generated Content and Framing Effects**

**Abstract.** Twitter, a micro-blogging website, is one of the fastest growing social networks on the Internet. The high variance of the information that propagates through large user communities makes this network a significant player in service-oriented markets.In this paper, we examine the impacts of Twitter content on train users’ performance risk and how train providers’ message framing moderates these relationships. Framing regards the way in which messages are worded concerning a particular object. In many consumer markets such as train journeys, firms frame messages in both a positive and negative light to persuade individuals to make purchase decisions. We thus go beyond the literature’s current focus on consumer generated content (CGC), and bring into contention the important role that marketer generated content (MGC) plays in shaping the contour of effective social media communication. Specifically, we analyze commuter tweets about 14 train operators, along with the companies’ Twitter feeds. The findings, obtained using sentiment analysis tools, suggest that both sentiments and frames play important roles in impacting consumers’ risk performance.

**Keywords:** Twitter; Consumer sentiment; Framing effect

**Conference Track:** Consumer Behavior

**1. Introduction**

Commuters are often hit with travel misery when a series of rush-hour problems strike their transport network. For example, signaling failures may see a part suspension on line 1 of the network, while line 2 maybe held up at a station due to a safety alarm. Another signaling problem may lead to delays on line 3, and a faulty train at a different station may add to the problems, causing severe delays on line 4. In a situation like this, crowds of commuters could be seen queuing for buses, with scores of others choosing to join the long wait for taxis. Rail staff could be seen busily directing travellers as the queues often snake along the pavements and into the road. Commuters may also endure these miserable journeys to and from work when a dispute between rail workers and managers disrupts the whole or part of the network. Overcrowding is also a common occurrence on commuter trains, which can be a source of great inconvenience for passengers, and, in some situations, it can be potentially highly dangerous.

Commuters now have an open forum to provide feedback, voice dissatisfaction with service and make complaints. In recent developments, rail passengers widely use Twitter to alert rail providers to the challenges and problems experienced on running services. In this research, we examine how online buzz and attention is created for different rail operators and how that changes over time. Specifically, we focus on how positive and negative opinions propagate via Twitter and how train providers’ message frames influence consumers’ risk performance. Previously, consumer engagement on social media often lacked marketers’ input, severely limiting the scope of social media as a tool for brand building and marketing communication. In a recent development, many companies have set up their own social media accounts that open the way for direct engagement with consumers.Although a few studies have already examined the scope of marketer generated content (MGC) along with withCGC (Godes and Mayzlin 2009;Trusov et al. 2009), they, nonetheless, focus on limited textual aspects of MGC content. In this study,we fill this gap and investigate the marketer generated content using the framing effect. Our main source of data is chatter from Twitter, as previous studies have found that the chatter of a community can be used to make more informed and stable predictions (Trusov et al. 2009). Using sentiment analysis from data obtained for 2013 and 2014, we find that a potent mix of delays, crowding and cancellations creates deeply uncomfortable situations for passengers. This results in often strong reactions on services as reflected through Twitter. We further examine the impact of user sentiments on their risk performance in terms of PPM (Public Performance Measure). We find a significant impact of negative sentiments on PPM. However, we also find a positive moderating role of companies’ positive frames on both positive and negative sentiments. This suggests that firms’ own engagement with passengers via Twitter plays a vital role in the commuter-train provider relationships.

**2. Literature Review**

The Internet has spawned a great variety of social networking platforms that have deeply affected how people communicate with one another (Albuquerque et al. 2012). These varieties can be seen from blogs, reviews and posts on Internet forums and websites as used by people who are mostly unknown to each other. When people buy a product or access product related company information, it may lead them to express opinion or raise concern on a social media platform (Godes and Mayzlin 2009). They may exhibit favorable attitudes and sentiments if their experience of the product or attachment to the company’s policies has been satisfying. The converse may also be true if they have found something troubling or felt unpleasant when consuming the product. These general evaluations of a company or its products and services can be referred to as valence embedded in CGC (Trusov et al. 2009). One may argue that positive valence of CGC should drive consumer purchases. Indeed, a great deal of marketing research shows that positive valence results in increases in consumer purchases (Godes and Mayzlin 2009). However, valance also refers to negative product or company reviews. People may weigh negative reviews more than positive reviews, as theorized by classic prospect theory (Tversky and Kahneman 1981). Further research shows that negative information affects consumer decisions more than positive information because negativity stays in people’s minds more than positivity does (Trusov et al. 2009). As the environment is generally more positive so consumers notice when negative details arise, and these may have a significant effect on decision-making. In the context of train journeys, this means that customers’ tarveling decisions may be affected by both positive eWOM and negative eWOM. We therefore present our first two hypotheses.

*Hypothesis 1. Positive Twitter sentiments have a significant effect on a commuter’s risk performance.*

*Hypothesis 2. Negative Twitter sentiments have a significant effect on a commuter’s risk performance.*

*2.1 Moderating effects*

Advertising and other non-paid communication targeted at customers are an important aspect of all businesses’ marketing strategies. It involves getting the right message across to the target audience, in order to acquire customers. Thus, it is important to understand the way in which to word a particular message, in order to persuade customers to make a purchase decision. The concept of framing has begun to contribute to this understanding by looking into the framing of products, either positively or negatively, against others in order to affect choice. Framing involves creating ‘frames’ (in other words, messages about a certain object) that emphasize gains or losses depending on whether the message is framed positively or negatively, respectively. For example, a cheaper product may be framed positively to emphasize that an individual would benefit (in monetary terms) from buying the cheaper product over the more expensive one. In contrast the more expensive could be framed negatively in terms of an individual losing money by buying the more expensive product over the cheaper one. In both cases the frames are intended to persuade the customer toward purchasing the cheaper product.

Framing effects have been of significant importance since Tversky and Kahneman (1981) put forward their contribution of preference reversal, based on the framing of messages in both positive and negative forms. Specifically, consumers are likely to be more risk averse in order to secure potential gains when a message is framed positively than when a message is framed negatively. Consequently, consumers will conduct more thorough searches to refine their market knowledge. Garner (1986) has earlier shown that consumers who are more risk averse conduct a more thorough analysis of the available information prior to decision making. We can therefore predict that the effect of tangible ques such as train breakdowns on performance risk will be less (greater) for consumers who are exposed to a positively (negatively) framed message. This is because those consumers would have done extensive searches before embarking on a train journey. Furthermore, it is well understood that whenever consumers attribute reporting or knowledge bias to the source, the persuasive impact of the message is typically lessened. It is thus important to determine a spokesperson’s (or train operator’s) credibility, which is generally seen as the source’s expertise, trustworthiness and reputation (Garner 1986). When source credibility is low, it is likely that consumers will discount the arguments in a message. We now present our next sets of hypotheses.

*Hypothesis 3a. A positively framed message and source credibility moderate the Twitter positive sentiments – risk performance relationship.*

*Hypothesis 3b. A positively framed message and source credibility moderate the Twitter negative sentiments – risk performance relationship.*

*Hypothesis 4a. A negatively framed message and source credibility moderate the Twitter positive sentiments – risk performance relationship.*

*Hypothesis 4b. A negatively framed message and source credibility moderate the Twitter negative sentiments – risk performance relationship.*

**3. Methodology**

We specifically examine the volume of tweets directed at the 14 rail providers that bring people into London, during the period of 2013 and 2014**.** Companies especially use Twitter handles to measureand respond to the severity of incidents that trigger the negative sentiment.We sourced tweets using the Twitter Search Api, which included information about the author, timestamp and tweet text. We extracted 2.47 million tweets referring to 14 train operators over a period of twenty-four months. To harness the data into a form that allows for specific predictions about particular outcomes, we employ sentiment methodology that enables us to probe key points relevant to the research. In this respect, weuse text classifiers to distinguish positively oriented tweets from negative ones. We included three variables to capture the effect of Twitter tweets: (1) the share of commuter Tweets who traveled on a weekday and sent a positive tweet about it within 24 hours (hereafter, PTWEET share); (2) the share of commuter Tweets who traveled on a weekday and sent a negative tweet about it within 24 hours (hereafter, NTWEET share); and (3) the ratio of positive to negative TWEETS (TWEETRATIO). We also included the total number of tweets sent within 24 hours of a weekday train journey as a measure of TWEETVOL. To eliminate confounding effects and to rule out any alternative explanations, we include in our analysis a number of control variables. Using the open-source data mining software WEKA, we divided the remaining tweets into positive and negative CGC using sentiment analysis. The analysis was executed simultaneously for all tweets in our data.

The varied nature of rail operations means that different countries have adopted a variety of performance measurement systems. A performance measure used when services are predominantly for short commuter trips will be different from when there is a mix of longer distance and commuter journeys. In the UK, the Public Performance Measure (PPM) measures the arrival punctuality of individual trains at their final destination against their planned timetable. More specifically, the performance of individual trains advertised as passenger services is measured against their planned timetable as agreed between the operator and Network Rail at 22:00 the night before. In other words, PPM is the percentage of trains ‘on time’ compared to the total number of trains planned. It combines figures for punctuality and reliability into a single performance record. This means that services that are cancelled or fail to operate their entire route, calling at every station, count as a PPM failure. Its driving principle is that by allowing the public to compare the performance of train operating companies against one another, operators can be held to account on rail performance. We therefore use PPM as an indicator of a commuter’s performance risk. We use company Twitter feeds (TFEED) as a moderator variable that influences the relationship between consumers’ use of Twitter and rail providers’ performance. For instance, in the following example those manning the Twitter feeds for individual tube lines seem to be doing a decent job of answering commuter queries about which parts of the line are open. The update presented in Figure 1 is a little baffling as information given is not as clear cut as it could be. PFEED are tweets that include positive words or statements like on-time, punctuality, customer service, etc whereas NFEED includes words or statements like breakdown, delay, damage etc. In the rail industry, the number of stars a rail provider receives denotes the quality of service and hospitality experienced by passengers. We use rail operator star rating as a proxy for their reputation and source credibility (CRED). Which? gives star rating based on consumer satisfaction and, therefore, it indicates the level of service a rail operator offers and ranges from 1 to 5 (Which? 2015). We also include a number of controls including train operator size (SIZE), train operator age (AGE), and NETEXP (the railway network’s annual spending on rail network in US$.).

**4. Results**

In Table 1, we present the descriptive statistics and correlations among variables in the study. Table 2 presents the findings of our empirical analysis when the dependent variable is PPM. The baseline model, Model 1 in Table 2, shows the effects of the controls and the moderators on PPM. Notably, size of train operators is positively related to PPM in such a way that smaller operators are less able to meet the PPM standard than bigger operators. The parameters for the other controls mostly behave as expected. For instance, improvements in performance are affected by network expenditures, suggesting that a rail network’s performance is inextricably linked to investment in physical infrastructure. Twitter feeds is positively related to PPM in that the higher the volume of Twitter feeds, the higher its influence on meeting the PPM standard. Furthermore, Age and TWEETVOL are insignificant, explaining no variance above and beyond the UGC related measures. Although the correlation between positive sentiments and PPM was highly significant (*r* = .357, *p* < .01, from Table 1), the regression coefficient for positive sentiments in Model 2 of Table 2 is only marginally significant (β = .761, *p* < .10). In contrast, negative sentiments NTWEET is significantly related to the rail industry’s PPM standard (β = -36.539, *p* < .05). Model 3 examines the moderating influence of Twitter feed on consumers’ positive (and negative) sentiments - PPM relationship. As shown in the table, the effects of the interaction terms of positive Twitter feed, negative (and positive) sentiments and source credibility (Model 3) are significant, but the effects of the interaction terms of negative Twitter feed, negative (and positive) sentiments and source credibility are not significant. These results suggest that the effect of a positive Twitter frame (PFEEF) on PPM is stronger for both positive and negative sentiments than a negative Twitter frame (NFEED). This implies an important role for message frames in how a train operator can enhance the commuter experience by focusing on giving positive messages (e.g. transmitting more novel and creative information). We thus find a confirmation of Hypothesis 3a/b and 4a/b. Model 4 presents the fixed effect regression results as there may be a concern that factors other than “Twitter Effect” are in play. However, we do not find any significant difference from the results reported in earlier Models.

**Table 1: Descriptive statistics and correlations**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Mean** | **SD** | **1** | **2** | **3** | **4** | **5** |
| 1. *LN*(NETEXP) | 9.784 | 8.362 |  |  |  |  |  |
| 2. TFEED | .0261 | .0252 | −.312 |  |  |  |  |
| 3 CRED | 3.372 | 0.736 | .301\*\* | .192\*\* |  |  |  |
| 4. AGE | 17.563 | 15.924 | .082 | .013 | .065 |  |  |
| 5. SIZE | 5,727.830 | 3,873.242 | .253 | .392\*\*\* | .332\*\*\* | .124 |  |
| 6.TWEETRATIO | 8.064 | 7.382 | −.078 | .382 | .276 | .093 | .083 |
| 7. TWEETVOL | .000 | 3,583.473 | .432 | .416\*\* | .074 | .088 | .256\*\* |
| 8. PTWEET | .051 | .0453 | .103 | .325\*\*\* | .282\*\* | .123 | .152\*\* |
| 9. NTWEET | .357 | .326 | .186 | .123\*\* | .133 | .255 | .237\* |
| 10. PPM | 0.932 | 0.917 | .132\*\* | .112 | .197 | .176 | .236\*\* |
| 11. *LN*(FPER) | 3.456 | 3.382 | .157\*\* | .282 | .119 | .143 | .247\* |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **6** | **7** | **8** | **9** | **10** |
| 1. *LN*(NETEXP) |  |  |  |  |  |
| 2. TFEED |  |  |  |  |  |
| 3. CRED |  |  |  |  |  |
| 4. AGE |  |  |  |  |  |
| 5. SIZE |  |  |  |  |  |
| 6. TWEETRATIO |  |  |  |  |  |
| 7. TWEETVOL | .094 |  |  |  |  |
| 8. PTWEET | .173\*\* | .175\*\* |  |  |  |
| 9. NTWEET  | .143\* | .248\*\*\* | .215\*\*\* |  |  |
| 10. PPM | .124\* | .226\* | .357\*\*\* | .364\*\*\* |  |
| 11. *LN*(FPER) | .187\* | .135\* | .184\*\* | .254\*\* | .334\* |

Notes. \*\*\* *p*<.01, \*\* *p*<.05, \* *p*<.10.

**Table 2. Estimation results for PPM**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Model 1** | **Model 2** | **Model 3** | **Model 4** |
|  | Coef. (Std. Err) | Coef. (Std. Err) | Coef. (Std. Err) | Coef. (Std. Err) |
| Constant | 51.598\*\*\*(5.863) | 51.683\*\*\* (5.468) | 52.274\*\*\*(5.681) | 51.389\*\*\*(5.372) |
| *LN*(NETEXP) | 3.648\*\*(1.757) | 6.896\*\*(1.608) | 4.517\*(1.723) | 3.331\*\*(1.613) |
| TFEED | 4.473\*\*\*(.546) | 4.389\*\*\* (.384) | 3.298\*\*\*(.542) | 3.273\*\*\*(.550) |
| CRED | .132(.392) | −1.914\*\*\*(.362) | .027(.022) | .063(.051) |
| AGE | -.253(.737) | -.218(.552) | -.172(.532) | -.012(.357) |
| SIZE | 2.463\*\*(.783) | 2.158\*\*(.082) | 3.105\*\*(.102) | 2.319\*\*(.023) |
| TWEETRATIO | 1.894\*\*(.739) | 1.843\*\* (.592) | 1.193\*\*(.533) | 1.593\*\*(.774) |
| TWEETVOL | 2.783(1.648) | -1.843(1.273) | -1.284(1.054) | 1.432(1.252) |
| PTWEET |  | .761\*(.337) | .490\*(.609) | .052\*(.142) |
| NTWEET |  | -36.539\*\*\*(11.372) | -28.161\*\*\*(12.009) | -27.256\*\*(11.492) |
| PTWEET x PFEED x CRED |  |  | 4.331\*\*\*(.889) | 8.491\*\*\*(.711) |
| PTWEET x NFEED x CRED |  |  | .112(.015) | .134(.017) |
| NTWEET x PFEED x CRED |  |  | 3.886\*\*(.576) | 3.810\*\*(.434) |
| NTWEET x NFEED x CRED |  |  | .167(.075) | .047(.010) |
| R2 = Adjusted R2 =  | .514.493 | .557.538 | .593.542 | .548.537 |

Notes: PPP*i* = β0 + β1NETEXP*i* + β2TFEED*i* + β3CRED*i* + β4AGE + β5SIZE + β6TWEETRATIO*i* + β7TWEETVOL*i* + β8PTWEET*i* + β9NTWEET*i* + β10PTWEET*i* x PFEED*i* x CRED*i* + β11PTWEET*i* x NFEED*i* x CRED*i* + β12NTWEET*i* x PFEED*i* x CRED*i* + β13NTWEET*i* x NFEED*i* x CRED*i* +ε*i*. \*\*\* *p*<.01, \*\* *p*<.05, \* *p*<.10.

**5. Conclusion**

In this study, we empirically test the “Twitter effect” of user sentiments on their risk performance. Findings have alluded to numerous insights concerning the use of Twitter to manage commuter relationships. Amongst the most significant of findings is the role that a train operator’s message frames play in mitigating the unfavorable impact of delays and cancellations. Train operators use their Twitter feed to engage with commuters in real time and provide up to date information, the impact of which can be seen in how positive messages enhance consumers’ risk performance by allowing them to do more searches and minimizing their risks. We thus find strong support for the moderating role of message frames and source credibility.

**References**

Albuquerque P, Pavlidis P, Chatow U, Chen KY, Jamal Z (2012) Evaluating promotional activities in an online two-sided market of user-generated content. *Marketing Science*. 31(3): 406–432.

Garner, S. J. (1986), Perceived Risk and Information Sources in Services Purchasing, *Mid-Atlantic Journal of Business,* 24 (Summer), 49-58.

Godes D, Mayzlin D (2009) Firm-created word-of-mouth communication: Evidence from a field test. *Marketing Science* 28(4): 721–739.

Trusov M, Bucklin RE, Pauwels K (2009) Effects of word-of-mouth versus traditional marketing: Findings from an Internet social networking site. *Journal of Marketing* 73:90-102.

Tversky A., Kahneman D. (1981) The framing of decisions and the psychology of choice, *Science,* 211(4481): 453-458.

Which? (2015) *Train satisfaction survey for 2015*. <http://www.which.co.uk/home-and-garden/leisure/reviews-ns/best-and-worst-uk-train-companies/best-train-companies-overall/>. Accessed 12 March, 2015.