



Operationalizing the Last Mile: A Sentiment Analysis Approach to Delivery Quality in Indonesian Mobile Commerce

Sulwani Husna Afrizal^{1*}, Desty Hapsari Kirana², Irsyad Kamal³

Universitas Padjadjaran

Corresponding Author: Sulwani Husna Afrizal sulwani@unpad.ac.id

ARTICLE INFO

Keywords : Last-Mile Delivery, Sentiment Analysis, User-Generated Reviews, Logistics Service Quality, Shopee Indonesia, Ramadan, E-Commerce

Received : 3 June

Revised : 20 July

Accepted: 13 August

©2026 Afrizal, Kirana, Kamal : This is an open-access article distributed under the terms of the Creative Commons Atribusi 4.0 Internasional.



ABSTRACT

Last-mile delivery is a critical determinant of customer satisfaction in platform-based e-commerce. However, specific delivery failures that generate negative customer sentiment mechanisms remain insufficiently understood. Especially in rapidly expanding Southeast Asian markets, which are characterized by island geography and seasonal demand fluctuations. This study examines last-mile delivery failures and their effects on customer sentiment. With analyzing 67,600 Google Play Store reviews on Shopee Indonesia Apps collected from January to April 2026. Using keyword extraction, 12,797 delivery-related reviews were identified, of which 54.6% expressed negative sentiment. These reviews were classified into nine sub-themes of delivery failure. The results showed that the highest scores were for delivery delay (41.6%), courier misconduct (32.7%), and SPX Express infrastructure failure (32.70%), which were the most prevalent. Temporal analysis reveals that the Ramadan shopping season significantly increased negative delivery reviews, resulting in a sixfold rise between January and March 2026 and an increase in the delivery negativity rate from 42.7% to 58.3%. Drawing on the Logistics Service Quality framework, the Expectation-Disconfirmation Model, and Service Failure and Recovery Theory, the findings indicate that delivery failures are structurally compounding. Courier misconduct leads to delays, delays result in inaccurate estimated times of arrival, and these inaccuracies intensify customer dissatisfaction before recovery measures are implemented. This study extends the LSQ framework to platform-mediated logistics by introducing the concept of attribution asymmetry. Namely, by validating the use of user-generated reviews as a source of real-time data and documents the forms of delivery failures

INTRODUCTION

Indonesia's e-commerce market has experienced rapid expansion and is expected to continue growing, with projections estimating that the number of e-commerce users will reach 99.1 million by 2029. This upward trajectory is supported not only by increasing consumer participation in digital commerce, but also by structural enablers such as the growing adoption of cashless transactions and the integration of artificial intelligence (AI) into platform operations, which enhance personalization, transaction efficiency, and overall user experience (Hentihu et al., 2025). As the largest e-commerce market in Southeast Asia, Indonesia offers particularly strong conditions for platform-based growth. This position is underpinned by its large and digitally connected population, widespread smartphone penetration, and the presence of major marketplace operators supported by fintech-enabled payment systems that simplify and accelerate online transactions (Ha & Chuah, 2023). These developments reinforce Indonesia's strategic importance in the regional digital economy and intensify competition among domestic and international marketplace platforms, while also drawing regulatory attention toward the development of a more coherent and supportive e-commerce ecosystem (Moha et al., 2023). Within this rapidly growing environment, logistics performance has become a crucial determinant of platform competitiveness, particularly at the last-mile stage. Prior literature consistently shows that customers' evaluations of service quality, trust, and repurchase intention are strongly shaped by last-mile delivery performance, especially in terms of timeliness, reliability, and the condition of items received (Mangiaracina et al., 2019). The last mile is therefore widely recognized as the final and most customer-facing stage of the delivery process, one that not only represents a substantial share of logistics costs but also directly influences customer satisfaction, brand perception, and the overall credibility of the platform.

These broader developments become especially important in the context of Shopee Indonesia, one of the country's major e-commerce platforms, where high transaction intensity makes delivery performance highly visible to consumers. The dataset analyzed in this study provides strong preliminary evidence of such urgency. Out of 67,600 Google Play Store reviews collected from January to April 2026, 12,797 reviews explicitly mentioned delivery-related issues. Strikingly, the negativity rate among delivery-related reviews reached 54.6%, far exceeding the overall platform negativity rate of 25.2%. These figures suggest that delivery-related failures are not peripheral complaints, but a dominant source of negative customer evaluations, thereby reinforcing the strategic importance of last-mile delivery in platform-based e-commerce.

Although last-mile delivery (LMD) is widely recognized as the most customer-facing and variable stage of e-commerce logistics, much of the logistics service quality (LSQ) literature still relies on surveys, experiments, and internal KPI measurements (Mangiaracina et al., 2019; Ha et al., 2022; Argyropoulou et al., 2022). As a result, unsolicited user-generated reviews and social data remain underutilized as real-time evidence of service failure, even though they can provide more immediate insights into problems such as delivery delays,

damaged parcels, and incorrect items (Jurburg et al., 2023; Vrhovac et al., 2024). This gap is especially relevant in Southeast Asia and Indonesia, where archipelagic geography, urban-rural disparities, and infrastructure limitations make last-mile delivery more complex and more variable (Zhu et al., 2021). Moreover, studies that integrate delivery failure types, customer sentiment, and seasonal demand surge within a single model are still limited, as these dimensions are often examined separately (D'agostini, 2023). Therefore, an important research gap remains in developing an integrated model that explains how specific delivery failures shape customer sentiment under peak-demand conditions, particularly in Indonesian e-commerce platforms such as Shopee Indonesia.

In response to this research gap, this study examines last-mile delivery failures on Shopee Indonesia by focusing on four main objectives. First, it seeks to identify the dominant forms of delivery failure reflected in user-generated reviews. Second, it aims to assess how each failure type influences customer sentiment outcomes, thereby clarifying which delivery problems generate the strongest negative reactions. Third, the study examines whether the Ramadan demand surge acts as a moderating condition that strengthens the relationship between delivery failures and customer sentiment during peak shopping periods. Finally, it interprets the strategic implications of these findings for e-commerce platforms, particularly in relation to logistics governance, service quality management, and the coordination of delivery partners in high-volume digital marketplaces.

LITERATURE RIVIEW

E-Commerce Growth and the Indonesian Digital Economy

The digital economy in Indonesia is growing rapidly and as a result is altering consumer expectations as well as the way e-commerce companies compete. Ha and Chuah (2023) point out that Indonesia is unique in Southeast Asia due to the high level of smartphone usage, the size of its online population, and the development of fintech payment systems which make it easier to shop online. With an increasing number of people going online, the volume of transactions goes up, causing businesses to enhance their services and user experience. Hentihu and Sandi (2025) also note that artificial intelligence is becoming more widely used in Indonesian e-commerce, enabling the platforms to provide more personalised services and to operate more efficiently. This indicates that embracing new technology has now become a major way for these platforms to differentiate themselves.

Moha and others in 2023 state that the e-commerce sector in Indonesia is attracting greater attention from regulators, who are aiming to establish a well-defined and supportive legal framework for online business. They point out that the regulations in Indonesia are different from those in more developed digital markets such as Singapore, which presents both challenges and opportunities for e-commerce companies. This regulatory environment helps to account for the pressures experienced by platforms such as Shopee; since an increasing number of people are shopping online and expect a higher standard of service, the rules

now require companies to take responsibility for each aspect of the delivery process, including the final leg to the customer.

In a market that is growing rapidly, the way that companies manage logistics has now become a major factor in their competitiveness. Research indicates that customers' assessments of the quality of service, their level of trust, and their intention to make a further purchase are mostly determined by the last-mile delivery, particularly by how fast, reliably, and safely their orders reach them (Mangiaracina et al., 2019). The last mile represents the final and most noticeable stage of the delivery process and has a direct impact on customers' opinions of a brand and on the trust they place in the platform. It is important to understand how problems occurring at this stage result in negative customer feedback, both from a research perspective and for the purpose of shaping business strategy in Indonesia. The Expectation-Disconfirmation Model (Oliver, 1980) provides an explanation for this: when customers make a purchase, they form expectations regarding the delivery on the basis of the estimated delivery date displayed on the platform; their level of satisfaction then depends on whether the actual delivery meets or fails to meet these expectations. Since the data from this study shows that deliveries often fall short, negative feedback accumulates and shows up in user reviews.

Last-Mile Delivery

Last-mile delivery (LMD) is known as the most challenging and noticeable part of e-commerce fulfillment. According to Mangiaracina et al. (2019), the last mile accounts for a significant portion of logistics costs, often ranging from 28% to 53% of total delivery expenses, and has the biggest impact on customer satisfaction. Despite large investments in technology such as drones, parcel lockers, and autonomous vehicles, LMD continues to struggle with service quality problems, particularly in busy cities and fragmented markets like Indonesia.

Ha, Akbari, and Au (2022) reviewed more than 1,000 studies on LMD and found that timeliness, reliability, and communication are the main topics related to service quality. They also highlight a lack of research on Southeast Asia, where issues like infrastructure, island geography, and dependence on motorcycle couriers are common. This gap matters for this study, which examines how these factors cause service failures for consumers in Indonesia.

Mentzer, Flint, and Hult (2001) created the logistics service quality (LSQ) framework to explain what customers value in delivery services. The LSQ model includes nine quality factors, but timeliness, how order problems are handled, and staff interactions are most relevant to the complaints in this study. While LSQ was first used for business-to-business logistics, this study finds that e-commerce customers expect the platform to take responsibility for service quality throughout the entire process, including the platform, carrier, and consumer, regardless of who caused the issue. Vargo and Lusch (2016) support this view with their idea of service ecosystem responsibility, which makes the main platform responsible for the overall customer experience, even if problems begin with the carrier.

Argyropoulou et al. (2022) found that sudden increases in demand show the limits of carrier capacity, which carriers cannot manage alone unless they cooperate. Zhu, Lan, and Zhang (2021) point out that being flexible and located near supply chain hubs is important for effective logistics in Indonesia's fragmented market. D'agostini (2023), studying Vietnam, found that customers choose last-mile providers mainly for reliability and damage-free delivery, underlining the importance of timeliness and service quality. When these standards are not met, Bitner, Booms, and Tetreault's (1990) service failure model explains that human errors, such as couriers marking deliveries as failed without attempting delivery, are unplanned actions that make the situation worse. Smith, Bolton, and Wagner (1999) also show that customer reactions depend on the type and size of the failure, and that recovery efforts must fit the problem to restore satisfaction, which generic customer service responses often fail to do.

Studying LMD failures through user-generated reviews (UGR) is an important methodological choice. Jurburg et al. (2023) point out that social and digital data are underused in LMD research and recommend using real-time consumer feedback. Unlike surveys, UGR from app stores is unsolicited, covers longer periods, and is not limited by sample size. These reviews capture both everyday complaints and major failures, and peer endorsements show which issues are most important. Vrhovac et al. (2024) found that reliability and clear communication are the best predictors of last-mile satisfaction, which matches the main themes in this study's UGR analysis. Rao, Griffis, and Goldsby (2011) show that late deliveries greatly reduce the chance of future purchases, especially among loyal customers, supporting the reputational impact found in this study.

Despite improvements in last-mile delivery and logistics service quality research, three significant gaps remain. First, the Logistics Service Quality (LSQ) framework has not been comprehensively adapted for platform-based logistics. A central issue is attribution asymmetry, in which customers attribute problems caused by the carrier to the platform. Most previous studies conceptualize platforms, providers, and consumers as components of an integrated service system. This results in the platform often being held accountable for the entire delivery experience (Vargo & Lusch, 2016). Second, the moderating effect of seasonal religious demand events such as Ramadan on the failure-to-dissatisfaction relationship remains underexplored in the Southeast Asian context. These events introduce predictable yet intense capacity challenges that differ from those documented in Western and East Asian logistics contexts (Argyropoulou et al., 2022; Zhu et al., 2021). Third, user-generated reviews remain underutilized as a primary data source in last-mile delivery research. User reviews offer concrete evidence of service failures without being prompted, while surveys often overlook these conditions, especially for issues that rarely occur. (Jurburg et al., 2023; Ha et al., 2022). This study addresses these gaps by integrating sentiment analysis of user reviews with the LSQ and Expectation Disconfirmation Model (EDM), which focus on the Indonesian e-commerce sector during Ramadan.

METHODOLOGY

The study uses a qualitative-descriptive methodology and draws on secondary textual data from user-generated reviews (UGR) of the Shopee Indonesia app as listed on the Google Play Store. Since UGR offers genuine, unsolicited accounts of consumer experiences, it is especially appropriate for detecting naturally occurring patterns of delivery failure rather than those that are triggered by survey questions. This approach differs from the survey and KPI-based methods that are commonly used in research into logistics service quality (Mangiaracina et al., 2019; Ha et al., 2022; Argyropoulou et al., 2022; Jurburg et al., 2023; Vrhovac et al., 2024) and directly responds to the calls for incorporating real-time digital feedback into the analysis of last-mile delivery.

The dataset includes 67,600 reviews of the Shopee Indonesia app on the Google Play Store that were submitted from 24 January to 13 April 2026. This time period was deliberately chosen in order to cover the Ramadan-Lebaran peak season (February to March 2026), a time when demand in the Indonesian e-commerce sector reaches its highest level of intensity each year. By including this period, the study is able to look at how logistics capacity performs when faced with increased demand, a situation which the literature has identified as a major factor in causing last-mile service failures (Argyropoulou et al., 2022; Zhu et al., 2021).

The data analysis was carried out in two stages, each of which was guided by the theoretical frameworks mentioned earlier. In the first stage, keyword-based extraction was employed in order to find reviews that directly dealt with delivery problems. The keywords were chosen from both informal and formal Indonesian, including terms concerning delivery, courier, package, expedition, shipping cost, transit status, delay, and return. As a result, 12,797 reviews relating to delivery (18.9% of the total dataset) were identified. Of these, the reviews which had ratings of 1 to 2 stars were regarded as negative delivery reviews (numbering 6,981) and were taken as the main unit of analysis, in accordance with the EDM's definition of negative disconfirmation as a factor preceding negative sentiment outcomes (Oliver, 1980).

In the second stage, the negative delivery reviews were divided into complaint sub-themes by means of keyword-based pattern matching. The classification follows the typology of service quality dimensions as set out in the LSQ framework (Mentzer et al., 2001), this framework identifying timeliness, the way discrepancies with orders are handled, and the quality of contact with personnel as the most important factors in customers' assessments of delivery performance. Since individual reviews can mention more than one type of failure, the sub-themes that result are not mutually exclusive. The nine sub-themes that were found were: delivery delays, courier misconduct, SPX Express failures, returns without notification, inaccuracies in the estimated time of arrival, high shipping costs, packages showing no movement, wrong or missing deliveries, and complaints about JNT's removal.

The sentiment results were evaluated by means of two indicators. The star rating, on a scale from 1 to 5, was used as the main measure of customers' evaluation, and rating acted as an indicator validated by the community to show

the salience of complaints and the social visibility of each kind of failure. Reviews that specifically mentioned competitor platforms were also identified as possible signs of customer churn. These indicators implement the dependent constructs of the service failure and recovery framework (Bitner et al., 1990; Smith et al., 1999), according to which the severity and extent of the customer response are seen as the outcomes of the failure experience.

To examine the moderating effect of the Ramadan demand surge, the monthly distribution of negative delivery reviews was studied by comparing the rates of negativity in January, February, March, and April 2026. The analysis looks at whether the conditions of increased demand strengthen the relationship between the types of delivery failure and negative customer sentiment. This is in line with capacity constraint theory in the context of seasonal logistics (Argyropoulou et al., 2022) and the behavioural effects of delivery failures as described by Rao, Griffis, and Goldsby (2011).

RESULT

Descriptive Overview

The dataset analyzed in this study comprises 67,600 Google Play Store reviews of Shopee Indonesia collected between January 24 and April 13, 2026. Overall, user evaluations were predominantly favorable, with approximately 66.6% of reviews assigning a 5-star rating, while 1–2-star ratings represented about one quarter of the total observations. A markedly different pattern emerged, however, when attention was restricted to reviews concerning delivery. Keyword-based filtering identified 12,797 delivery-related reviews, equivalent to 18.9% of the dataset. Within this subset, 54.6% expressed unfavorable evaluations, more than twice the platform-wide rate of 25.2%. This contrast indicates that last-mile delivery represents a disproportionately important source of customer dissatisfaction and suggests that delivery performance plays a central role in shaping users' overall service experience. At the same time, Shopee displayed a high level of responsiveness, replying to 62,748 reviews, or 92.8% of the total dataset. Despite this extensive response activity, the persistently high level of dissatisfaction within delivery-focused reviews suggests that responsiveness alone may be insufficient to compensate for shortcomings in actual delivery performance.

Delivery Complaint Typology

To identify the main forms of customer dissatisfaction with delivery services, negative delivery-related reviews were further classified into several complaint sub-themes. This classification was conducted to reveal the most frequently occurring problem patterns and to distinguish the severity of each complaint type based on its impact on order fulfillment and customer experience. Because a single review may contain more than one issue, the categories are not mutually exclusive. Table 1 summarizes the typology of negative delivery complaints identified in the data, together with their frequency, proportion of total negative delivery reviews, and severity level.

Table 1. Typology of Negative Delivery Complaints

Sub-theme	Count	% of negative delivery reviews	Severity
Delivery delay / slow	2,903	41.6%	High
Courier misconduct	2,282	32.7%	High
SPX Express failures	1,255	18.0%	High
Return without notice	832	11.9%	Medium
Inaccurate delivery estimate (ETA)	803	11.5%	Medium
Shipping cost too high	583	8.4%	Medium
Package stuck / no movement	479	6.9%	High
Wrong or missing delivery	331	4.7%	High
JNT removal backlash	148	2.1%	Medium

The dataset analyzed in this study consists of 67,600 Google Play Store reviews of Shopee Indonesia collected between January 24 and April 13, 2026. Overall, the composition of the dataset indicates that user evaluations of the platform tend to be positive, with approximately 66.6% of the reviews assigning a 5-star rating, while negative reviews rated 1–2 stars account for roughly one quarter of the total reviews. However, this pattern changes markedly when the analysis is narrowed to reviews discussing delivery-related issues. Based on keyword-based filtering, 12,797 reviews (18.9%) were explicitly related to delivery, and within this group the proportion of negative reviews reached 54.6%, substantially higher than the overall platform-wide negativity rate of 25.2%. These findings indicate that dissatisfaction associated with delivery is considerably greater than general dissatisfaction with the platform, suggesting that last-mile delivery may be regarded as a primary source of negative customer experience. Interestingly, Shopee also demonstrates a very high level of responsiveness, with a reply rate of 92.8%, or 62,748 out of 67,600 reviews, indicating that customer complaints are being actively addressed. Nevertheless, the high proportion of negative sentiment within delivery-related reviews suggests that platform responsiveness alone has not been sufficient to mitigate the negative perceptions arising from delivery failures.

Monthly Trend and Ramadhan Effect

To examine how delivery-related dissatisfaction changed over the observation period, the next stage of analysis maps the monthly distribution of negative delivery reviews. This analysis is important for determining whether customer complaints remained relatively stable or instead intensified during particular periods characterized by higher demand pressure. By comparing the number of negative delivery reviews, the total number of delivery reviews, and the monthly negativity rate, Table 2 provides an overview of the seasonal pattern of delivery complaints, particularly during the Ramadan–Eid period as the peak transaction season.

Table 2. Monthly Trend of Negative Delivery Reviews

Month	Negative delivery reviews	Total delivery reviews	Negative rate	Context
January 2026	405	948	42.7%	Baseline
February 2026	2,487	4,702	52.9%	Ramadan peak
March 2026	2,891	4,959	58.3%	Ramadan / Eid
April 2026*	1,198	2,188	54.8%	Post-Eid

*April is a partial observation period

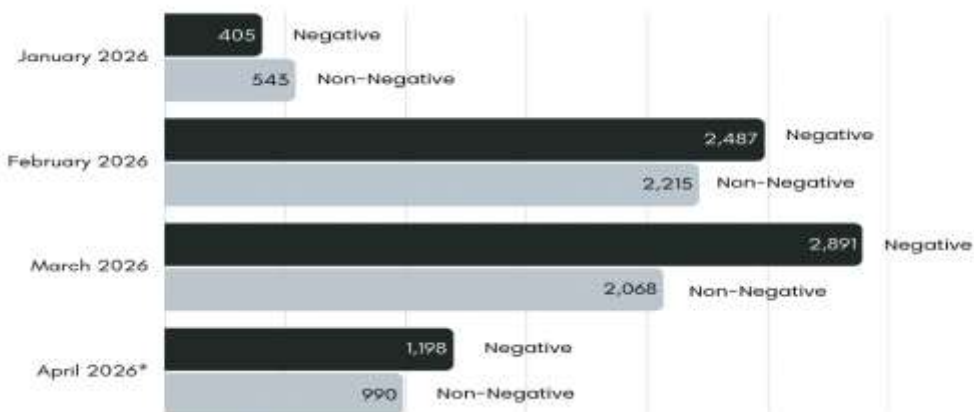


Figure 1. Monthly Trend of Negative & Non-Negative Delivery Reviews

The monthly analysis shows a clear surge in delivery-related dissatisfaction from January 2026 to February - March 2026. Negative delivery reviews increased from 405 cases in January (42.7%) to 2,487 in February (52.9%) and peaked at 2,891 in March (58.3%), representing an increase of more than sixfold during the Ramadan-Lebaran period. This pattern suggests that January serves as the baseline month, whereas February and March reflect the platform's peak season, when higher transaction volumes likely exacerbated delivery failures. Although April 2026 represents only a partial post-Lebaran period, the negativity rate remained high at 54.8%, indicating that delivery-related dissatisfaction persisted even after the main holiday period.

DISCUSSION

The findings indicate that delivery problems are not merely a minor operational issue, but a major source of negative customer sentiment on the Shopee Indonesia platform. The dominance of complaints concerning delivery delays, courier misconduct, and SPX service failures suggests that the last-mile delivery experience plays a decisive role in shaping customers' overall evaluation of the platform.

Delivery delay as the most visible service failure

The analysis shows that delivery delays are the most common complaint in this dataset, making up 41.6% of all negative delivery-related reviews (n = 2,903 out of 6,981). This result matches established research in logistics service quality, where timeliness, or delivering within the promised or expected

window, is seen as the most important part of customer-perceived delivery performance (Mentzer, Flint, & Hult, 2001). The high number of delay-related complaints here suggests that, for Indonesian e-commerce customers, late deliveries are more than just an inconvenience, they represent a real failure to meet service expectations. It supports Mangiaracina et al. (2019)'s view that last-mile performance is the main factor in customer satisfaction for B2C e-commerce. The Expectation-Disconfirmation Model (EDM) (Oliver, 1980) helps explain this pattern. According to this model, customer satisfaction depends on the difference between what people expect before buying and what they experience after. For last-mile delivery, Shopee's platform clearly shows an estimated delivery date when customers make a purchase. This date sets their expectations. If the order arrives later than promised, customers feel disappointed and their satisfaction drops. One review with 237 upvotes describes this process very clearly:

"Shopee deliveries have been slower lately; items that have been ordered are arriving three days later than the estimated delivery date.." – 1-star review, 237 upvote

Packages consistently arrive three days later than the estimated delivery date. This indicates two main reasons. First, the delay happens regularly, which points to a deeper issue between Shopee's ETA algorithm and its actual logistics capacity. This kind of infrastructure scaling problem matches what Argyropoulou et al. (2022) found. The study said individual carriers often face capacity limits that hurt delivery performance when demand is high. Second, many people have reported the same problem, showing that it is common and recognized by others. According to service failure research (Bitner, Booms, & Tetreault, 1990), when a failure is both visible and widespread, it can seriously affect a company's reputation. In this dataset, delivery delays meet both of these conditions and are the main issue around which other problems seem to gather.

Courier misconduct and the accountability problem

The second most common complaint is courier misconduct, which makes up 32.7% of negative delivery reviews (n = 2,282). This issue is different from delays, which often result from systemic capacity problems described by Argyropoulou et al. (2022) and Mangiaracina et al. (2019). Courier misconduct, on the other hand, involves intentional or careless actions by the courier. The most frequent example is when couriers mark deliveries as "failed attempt" or cancel shipments without actually visiting the customer. This practice directly breaks the unwritten service agreement between the platform, seller, and buyer. This type of behavior matches the third category in Bitner, Booms, and Tetreault's (1990) service failure typology: unprompted and unsolicited employee actions that make a problem worse instead of fixing it. The courier is the last person from the platform that the customer interacts with, so customers blame the platform for any misconduct at this stage, no matter how the courier is employed. This is an example of what Vargo and Lusch (2016) call service ecosystem responsibility, where the platform is seen as responsible for the whole delivery process. One review with 579 peer endorsements shows this point clearly :

“I urge Shopee to improve the courier delivery system; lately, many couriers fail to deliver orders to the actual home address and instead cancel the delivery themselves, even though I am at home.” — 1-star review, 579 upvotes

The size of this problem matches Vrhovac et al.’s (2024) finding that courier reliability and clear communication are the most important factors for last-mile user satisfaction. Both are often ignored when couriers cancel deliveries without contacting customers. In addition, a review with 1,173 upvotes asking Shopee to provide courier WhatsApp contact numbers shows there is a gap in the platform’s accountability system. Without a direct way to escalate issues, customers have to find their own solutions. Smith, Bolton, and Wagner (1999) show that recovery efforts need to fit the type and seriousness of the failure to restore satisfaction. The platform’s high reply rate (92.8%) but still high negativity rate (54.6%) in delivery reviews suggests that standard customer service responses are not enough.

SPX failures and the governance risk of platform-owned logistics

One of the most important patterns is that many complaints focus on SPX Express, Shopee’s own logistics service. SPX appears in 18.0% of negative delivery reviews ($n = 1,255$) and is linked to the most upvoted complaint in the dataset (2,133 peer endorsements). This trend stands out because SPX became a main target for complaints just as Shopee reportedly stopped working with third-party logistics partners. As a result, customers could no longer choose a different carrier at checkout.

This change shows Shopee’s choice to manage last-mile logistics themselves, which the broader LMD literature says can help control costs and improve process visibility (Mangiaracina et al., 2019). However, the reviews reveal a risk: when one company handles all logistics, any service problems directly affect the platform’s reputation. Zhu, Lan, and Zhang (2021) found that adaptability is crucial in Indonesia’s complex logistics market, and removing carrier choice makes it harder for platforms to stay flexible. The most-upvoted review reflects how consumers feel about this decision :

“J&T is no longer an option; you have to use SPX. Shipping with SPX is truly terrible — packages can get stuck for days with no resolution. I ordered two more items via SPX, and now they’ve vanished into thin air..” — 1-star review, 2,133 upvotes

Many reviewers say that removing Other last-mile provider is the main reason for their dissatisfaction, showing that losing the option to choose a carrier feels like a loss of control for customers. Complaints about SPX often mention packages stuck at distribution centers without updates or routing mistakes between islands. These issues suggest problems with infrastructure scaling rather than just mistakes by individual couriers. This matches D’agostini’s (2023) finding in Vietnam that reliability and damage-free delivery are the most important factors for choosing last-mile providers, and that large-scale infrastructure failures hurt reputations more than single incidents. Jurburg et al. (2023) also point out that platform choices have lasting effects on customer experience, which often show up in digital feedback like app store review

Ramadan demand surge as an amplifier of delivery dissatisfaction

A key insight comes from the timing of delivery complaints. Negative delivery reviews rose sharply from 405 in January 2026 to 2,487 in February and 2,891 in March, a six-fold increase that lines up with the Ramadan shopping season. The percentage of negative delivery reviews also went up, from 42.7% in January to 58.3% in March. This suggests that the decline is due to a real drop in service quality, not just an increase in the number of reviews. This supports Argyropoulou et al.'s (2022) finding that sudden spikes in demand reveal capacity limits that logistics partnerships alone cannot solve, especially when a platform relies on a single main carrier.

In Indonesia, Ramadan brings a demand surge that is both predictable and very large, with order spikes caused by promotions, gift buying, and pre-Lebaran shopping. Zhu, Lan, and Zhang (2021) point out that Indonesia's many islands and uneven infrastructure make logistics especially vulnerable to seasonal pressure. Distribution chains that work well most of the year can break down when demand is high. Review data shows this strain through mentions of bottlenecks at distribution centers like 'Kosambi DC,' overloaded couriers, and inaccurate ETAs, which show that the system cannot adjust to sudden capacity limits. D'agostini (2023) also notes that in Southeast Asia, planning for peak demand is key for last-mile delivery success, which this study's data supports.

From a theoretical perspective, the Ramadan surge acts as a natural factor that affects the link between delivery failures and customer feelings. It does not change the types of failures, but it makes them happen more often and with greater impact, which strengthens the connection between failures and customer dissatisfaction described in service recovery theory (Smith, Bolton, & Wagner, 1999). The high negativity rates in April 2026 (54.8%), even after the Lebaran peak, suggest that dissatisfaction from the surge has a lasting effect on reputation. This matches Rao, Griffis, and Goldsby's (2011) finding that delivery failures can reduce future purchases, especially among loyal customers. This makes Ramadan capacity planning a major challenge for Indonesian e-commerce platforms. Although this is an annual KPI metric often missed, it is clear in long-term UGR data (Jurburg et al., 2023).

Theoretical Contribution

This study contributes to operations management and logistics research in three main ways. First, it extends the LSQ framework (Mentzer et al., 2001) by applying it to platform-based logistics in a developing market. The original LSQ framework focused on B2B logistics, and later research adapted it for B2C contexts (Vrhovac et al., 2024). However, these versions have not fully addressed cases where a platform coordinates sellers, logistics partners, and customers together. Our findings show that in these situations, customers still hold the platform accountable for LSQ outcomes, even when issues begin with the carrier. This suggests that LSQ frameworks should consider how responsibility is assigned. This insight also builds on Vargo and Lusch's (2016) concept of service ecosystem responsibility, offering a clearer view of how accountability moves through complex service systems.

Second, this study highlights the benefits of using user-generated reviews (UGR) as a data source in operations management research. As Jurburg et al.

(2023) and Ha, Akbari, and Au (2022) point out, most last-mile delivery studies rely on surveys and experiments, which often miss the detailed, real-time information found in digital user content. We analyzed 12,797 delivery-related reviews that were unsolicited, spanned a long period, and were not limited in number. These reviews include both common complaints and rare, serious failures. The peer endorsement feature also provides a community-based way to see which complaints are most important, something surveys cannot offer. This approach supports Vrhovac et al.'s (2024) call for more consumer-focused data in last-mile research and shows that UGR can match or even improve on survey-based results.

Third, this study contributes to service failure and recovery theory (Bitner et al., 1990; Smith et al., 1999) by showing how different types of failures can build on each other in platform logistics. For instance, if a courier behaves improperly, it can cause delays. These delays can then lead to inaccurate delivery estimates, which frustrate customers even more, especially if the platform does not address the main problem. This chain of issues suggests that current service recovery models, which usually see failure and recovery as a simple two-way process, do not reflect the complexity of platform e-commerce, where several types of failures can occur in a single order. Evidence from the Ramadan period also shows that these chains of failures become more intense during high-demand times, pointing to a gap in current LSQ and service failure models, as noted by D'agostini (2023) and Ha et al. (2022).

Managerial Implications

These findings have clear takeaways for platform operators, logistics managers, and policymakers working in Indonesia's e-commerce sector. Ramadan capacity planning is essential. The six-fold increase in negative delivery reviews from January to March 2026 shows that Ramadan brings not just more orders, but also puts service quality to the test. Following Argyropoulou et al.'s (2022) advice on handling unusual demand in last-mile delivery, platform operators should set up extra logistics capacity, place inventory at distribution centers ahead of time, and add more couriers at least four weeks before the busiest period. Capacity planning should use past UGR sentiment as well as order volume to judge if the infrastructure is ready, since the evidence shows that a drop in sentiment during Ramadan often comes before customers leave (Rao et al., 2011).

Courier governance should move beyond just holding individuals accountable and focus on better system design. With courier misconduct at 32.7%, managing performance at the individual level is not enough. Vrhovac et al. (2024) found that clear communication is key to last-mile satisfaction. Platform operators should add GPS-verified delivery confirmation, real-time courier tracking, and ways for customers to contact couriers directly. These changes help fix not only the misconduct but also the lack of accountability that makes customers look for informal solutions, as shown by the 1,173 upvotes for courier WhatsApp numbers. Smith, Bolton, and Wagner (1999) show that recovery efforts should fit the type of problem; for courier issues, proactive updates and direct ways to escalate are the best solutions.

Accurate delivery estimates help build trust. When ETAs are wrong, as seen in 11.5% of negative reviews, customers are disappointed even before a delivery fails, according to Oliver's (1980) EDM. Platforms should use dynamic ETA systems that factor in real-time carrier capacity, seasonal changes, and past delivery data for each route. If delays are expected, letting customers know in advance can greatly reduce dissatisfaction (Rao, Griffis, & Goldsby, 2011). However, current evidence suggests this is still not done enough on the Shopee platform.

Carrier diversification is an important way to manage risk. Many complaints about SPX and the strong reaction to JNT's removal show that relying on a single, platform-owned logistics provider creates risks that customers feel directly. Zhu, Lan, and Zhang (2021) also argue that adaptability, including maintaining a mix of supply chain partners, is vital in Indonesia's complex market. Allowing customers to choose from multiple carriers at checkout spreads out the risk, gives customers more control, and provides useful data for future decisions. D'Agostini (2023) also found that reliability is the main factor in choosing a carrier in Southeast Asia, so it makes sense to keep trusted third-party carriers alongside in-house options.

CONCLUSIONS AND RECOMMENDATIONS

This study looks at 67,600 reviews from the Shopee Indonesia Google Play Store to see how last-mile delivery affects customer sentiment. Out of 12,797 delivery-related reviews, 54.6% were negative, which is more than twice the platform's overall negative rate. This shows that last-mile delivery is a major cause of customer dissatisfaction. Customers mainly complained about four delivery issues: delivery delays (41.6%), irresponsible courier behavior (32.7%), lack of infrastructure (18.0%), and inaccurate ETAs (11.5%). Deliveries increased during Ramadan, which led to up to six times more complaints between January and March 2026. During this time, the rate of negative delivery-related complaints went up from 42.7% to 58.3%. These results show that delivery failures build on each other: courier behavior can cause delays, delays lead to inaccurate ETAs, and wrong estimates make customers more dissatisfied even before any recovery efforts. This chain reaction is not fully explained by the integrated service recovery model (Bitner et al., 1990; Smith et al., 1999).

The findings of this study are based on three theoretical frameworks. Expectation-Disconfirmation Model (Oliver, 1980) explains how this study uses three main theories. The Expectation-Disconfirmation Model (Oliver, 1980) shows that the ETA given at checkout sets customer expectations, and breaking this promise leads to negative feelings. The Service Failure and Recovery Theory (Bitner et al., 1990) explains how irresponsible courier behavior can make problems worse and why customers blame the platform. The LSQ framework (Mentzer et al., 2001) situates these failures within a broader set of logistics service quality issues. Vargo and Lusch's (2016) idea of service ecosystem responsibility explains why customers see the platform as responsible for the whole delivery process, reducing the concept of attribution asymmetry, where customers hold the platform accountable for failures at the courier level regardless of its operational structure. Second, validating user-generated reviews

as a source of real-time, unsolicited, and scale-independent data for operations management research. This data will capture failure patterns and demand-amplification dynamics that are not visible through conventional survey instruments (Jurburg et al., 2023; Ha et al., 2022). Third, documenting the cascade of demand-amplified failures during Ramadan that have not yet been formally integrated into existing LSQ models and service recovery offers a foundation for developing future models at the intersection of seasonal operations and platform logistics governance.

From the findings, this study suggests four priorities from a management perspective. First, plan capacity for Ramadan based on potential order numbers and consider negative sentiment. Second, improve courier oversight with GPS-verified delivery and real-time tracking. Third, use dynamic ETA algorithms that adjust for season and route. Lastly, bring back multiple courier options to spread risk and give customers more choice. The spike in failures at SPX Express after JNT was removed is a clear warning: while vertical integration in logistics can save costs, it also increases reputational risk during busy periods.

This study has some limitations. The data only covers one platform over six months, and the keyword search method does not use advanced NLP classification. Future research should use sentiment scoring with other method, compare more platforms, and formally test how Ramadan affects results using regression with interaction terms. Analyzing the content of customer service responses to different types of failures would also help improve service recovery theory for high-volume digital commerce.

FURTHER STUDY

This study still has limitations; therefore, further research on the topic "Operationalizing the Last Mile: A Sentiment Analysis Approach to Delivery Quality in Indonesian Mobile Commerce" is needed to refine this study and broaden the insights of both the author and the readers

REFERENCES

- Argyropoulou, M., Zissis, D., Korfiatis, N., & Zampou, E. (2022). Horizontal collaboration in the last mile distribution: Gauging managerial response to disruption and abnormal demand. *Benchmarking: An International Journal*, 30(2), 460–474. <https://doi.org/10.1108/bij-06-2021-0328>
- Bitner, M. J., Booms, B. H., & Tetreault, M. S. (1990). The service encounter: Diagnosing favorable and unfavorable incidents. *Journal of Marketing*, 54(1), 71–84. <https://doi.org/10.1177/002224299005400105>
- D'agostini, E. (2023). Evaluating the selection factors for Vietnamese last-mile delivery service providers using Best Worst Method. *International Journal of Supply Chain Management*, 12(2), 23–32. <https://doi.org/10.59160/ijscm.v12i2.6166>
- Ha, H., & Chuah, C. K. P. (2023). Digital economy in Southeast Asia: Challenges, opportunities and future development. *Southeast Asia: A Multidisciplinary Journal*, 23(1), 19–35. <https://doi.org/10.1108/seamj-02-2023-0023>

- Ha, N., Akbari, M., & Au, B. (2022). Last mile delivery in logistics and supply chain management: A bibliometric analysis and future directions. *Benchmarking: An International Journal*, 30(4), 1137–1170. <https://doi.org/10.1108/bij-07-2021-0409>
- Hentihu, M., & Sandi, R. (2025). The role of consumer trust in adopting AI technologies on Indonesian e-commerce platforms. *2025 International Conference on ICT for Management, Technology and Human Sciences (ICIMTech)*, 638–643. <https://doi.org/10.1109/ICIMTech67074.2025.11265193>
- Jurburg, D., López, A., Carli, I., Chong, M., de Oliveira, L. K., Dablanc, L., Tanco, M., & de Sousa, P. R. (2023). Understanding the challenges facing decarbonization in the e-commerce logistics sector in Latin America. *Sustainability*, 15(22), 15718. <https://doi.org/10.3390/su152215718>
- Mangiaracina, R., Perego, A., Seghezzi, A., & Tumino, A. (2019). Innovative solutions to increase last-mile delivery efficiency in B2C e-commerce: A literature review. *International Journal of Physical Distribution & Logistics Management*, 49(9), 901–920. <https://doi.org/10.1108/ijpdlm-02-2019-0048>
- Mentzer, J. T., Flint, D. J., & Hult, G. T. M. (2001). Logistics service quality as a segment-customized process. *Journal of Marketing*, 65(4), 82–104. <https://doi.org/10.1509/jmkg.65.4.82.18390>
- Moha, M. R., Ahmad, A., Harun, A. A., & Elfikri, N. F. (2023). The comparative law study: E-commerce regulation in Indonesia and Singapore. *Jurnal Legalitas*, 16(2), 248–259. <https://doi.org/10.33756/jelta.v16i2.20463>
- Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, 17(4), 460–469. <https://doi.org/10.1177/002224378001700405>
- Rao, S., Griffis, S. E., & Goldsby, T. J. (2011). Failure to deliver? Linking online order fulfillment glitches with future purchase behavior. *Journal of Operations Management*, 29(7–8), 692–703. <https://doi.org/10.1016/j.jom.2011.04.001>
- Smith, A. K., Bolton, R. N., & Wagner, J. (1999). A model of customer satisfaction with service encounters involving failure and recovery. *Journal of Marketing Research*, 36(3), 356–372. <https://doi.org/10.1177/002224379903600305>
- Vargo, S. L., & Lusch, R. F. (2016). Institutions and axioms: An extension and update of service-dominant logic. *Journal of the Academy of Marketing Science*, 44(1), 5–23. <https://doi.org/10.1007/s11747-015-0456-3>
- Vrhovac, V., Dakić, D., Milisavljević, S., Čelić, Đ., Stefanović, D., & Janković, M. (2024). The factors influencing user satisfaction in last-mile delivery: The structural equation modeling approach. *Mathematics*, 12(12), 1857. <https://doi.org/10.3390/math12121857>
- Zhu, J., Lan, W., & Zhang, X. (2021). Geographic proximity, supply chain and organizational globalized survival: China's e-commerce investments in Indonesia. *PLOS ONE*, 16(9), e0256837. <https://doi.org/10.1371/journal.pone.0256837>