Journal of Marketing Analytics (2023) 11:149–159
https://doi.org/10.1057/s41270-022-00168-5

ORIGINAL ARTICLE

Impact of delivery performance on online review ratings: the role of temporal distance of ratings

Prashanth Ravula[1]

Revised: 27 July 2021 / Accepted: 23 April 2022 / Published online: 18 May 2022

Abstract
Customers are increasingly using online reviews in their purchase decision-making processes. As sellers benefit from displaying several reviews with favorable ratings, many sellers solicit reviews from customers. When a customer places an order on an e-commerce platform, the seller gets a notification to fulfill the order, and the customer is notified of the estimated delivery date. Some customers receive their products on time, while others receive their orders either earlier or later than the notified delivery date. After customers receive their products, the sellers often solicit reviews. This research focuses on the impact of delivery performance on review ratings. Specifically, this study addresses two questions: (1) Do customers reward sellers for early delivery in the same way they penalize them for late deliveries? (2) What is the role of the temporal distance of rating in online ratings in the context of delivery performance? The study estimates ordinal logit models in the Bayesian framework. Findings of the study indicate that customers give much lower (a little higher) ratings to orders delivered late (early) than to orders delivered on time. Further, the findings indicate that temporal distance is positively associated with ratings for late deliveries. The study discusses the theoretical and managerial implications of these results.

Keywords Online reviews · e-commerce · Delivery performance · Bayesian models

Introduction

With the increased diffusion of smartphones and the internet, e-commerce retail sales have seen significant growth in the past decade (Statista 2021). Given this phenomenon, many retailers have been adding online channels to reach more customers and increase sales revenue. Further, e-commerce platforms have created new opportunities for resource-constrained small retailers, allowing them to sell merchandise through the platforms. On the other hand, the number of consumers shopping online has also seen tremendous growth. According to an industry estimate, over two billion consumers have purchased goods or services online in 2020 (Statista 2022).

The recent global coronavirus (Covid-19) pandemic has led to a surge in e-commerce and accelerated these trends. A recent report shows the strong uptake of e-commerce sales across the world, with many consumers, especially in emerging economies, making the greatest shift to online shopping (UNCTAD 2022). For example, a study by Appsflyer (2021) reveals that the number of users who downloaded an e-commerce app on smartphones has increased by 48% in 2021. In addition, many consumers expanded the type of products they purchase online. As consumers are getting used to the convenience of online shopping, they are more likely to stay with that habit of online shopping (SBT 2022).

Perhaps, reflecting the above trends, industry estimates (eMarketer 2022) indicate that e-commerce retail sales are expected to increase to $7.39 trillion by 2025, accounting for 23.6% of global retail sales. This growth in e-commerce retail sales and competition among sellers require sellers to provide fast and convenient delivery, as after-sales service can impact customer satisfaction. As such several sellers collect customer satisfaction ratings after orders were delivered.

Online reviews have increasingly become an alternative to traditional customer satisfaction surveys. When a customer places an order on an e-commerce platform (e.g., Alibaba, Amazon, Flipkart, Olist Store), the seller gets a notification to fulfill the order, and the customer is notified of the *estimated delivery date*. Some customers receive products on

✉ Prashanth Ravula
pravula@njcu.edu

[1] NJCU School of Business, New Jersey City University, Jersey City, NJ 07311, USA

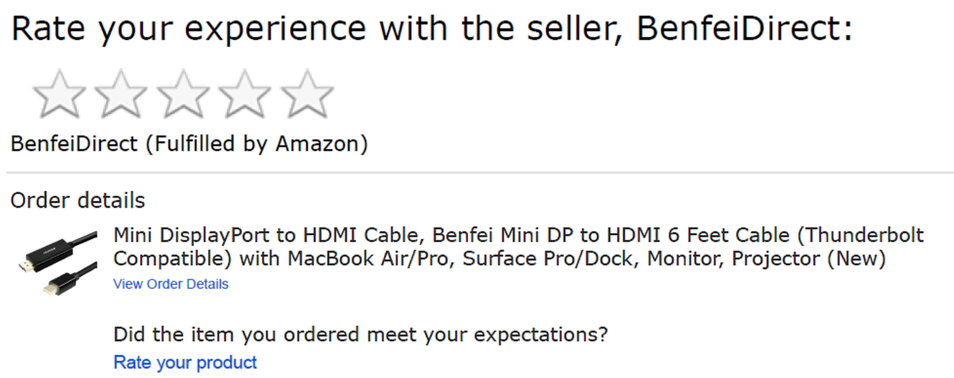

Fig. 1 Example of solicitation for rating

time, while other customers receive their orders either *earlier* or *later* than the *notified delivery date.* After customers receive their products, the sellers often solicit reviews from customers. Specifically, as shown in Fig. 1, sellers ask customers to rate the experience with the seller. While consumers tend to comprehend purchase experience with the seller on an e-commerce platform in a holistic perspective, it can be assumed that delivery performance will affect the overall customer satisfaction reflected in rating because these reviews are solicited right after delivery of orders. As such, we use online review ratings to evaluate the relationship between delivery performance and customer satisfaction.

Further extant research suggests that prospective customers increasingly use online reviews in their purchase decisions (Changchit and Klaus 2020; Berger et al. 2010; Chevalier and Mayzlin 2006; Zhu and Zhang 2010). For example, a recent survey (Carter 2022) found that around 89% of consumers agree that online reviews are an essential part of the purchase process. However, despite the growing importance of online reviews in e-commerce retail, surprisingly little is known about the relative effects of delivery performance: early and late vs. same-day delivery on online review ratings. On the other hand, some customers post their reviews, rating the sellers, on the same day of delivery, and some do it later. Extant research suggests that temporal distance[1] of rating influences online review ratings (Huang et al. 2016). However, there is no published research on the role of the temporal distance of ratings in the context of e-commerce delivery performance to the best of our knowledge. Filling the above gaps in the literature, this study, specifically, investigates two key questions in the context of after-sales service in the e-commerce retail sector: (1) Do customers reward sellers for early delivery in the same way they penalize them for late deliveries? (2) What is the role of the temporal distance in online review ratings in the context of delivery performance?

For the empirical analysis, e-commerce platform data were used from an emerging market. Data consist of late, early, and same-day deliveries. However, these incidences were not randomized. The possibility also existed that early and late deliveries were systematically different from same-day deliveries. A propensity score matching (PSM) algorithm was used to generate matched samples to address these issues. We then estimated ordinal logistic regression models on the matched samples. Specifically, two models were estimated: early vs. same-day deliveries and late vs. same-day deliveries. Insights from this study are managerially relevant. As e-commerce retail growth has become more prevalent in emerging markets than in developed countries (Kuhn and Petzer 2018), these findings assist retailers in devising effective review solicitation strategies. As such, this research represents the first study that underscores the importance of review solicitation time in the context of late deliveries, contributing to different streams of literatures: prospect theory, construal level theory, online reviews, and retail.

Theoretical background

As the e-commerce sector of retailing is rapidly growing, many firms are transforming their distribution channels to reengineer their relationships with customers. Several new features or drivers of customer satisfaction were identified in this online setting. The first set of studies focused on features related to the internet, such as ease of use, trust, etc. (Bhatnagar et al. 2000; Zeithaml et al. 2002). As purchase experience with e-commerce firms can be evaluated according to a range of features, the second set of studies identified several other features related to personalization, product range, prices, checkout, shipping, etc. as drivers of customer satisfaction (Dholakia and Zhao 2010; Jin and Park 2006). While several studies focused on the pre-delivery process, Jiang and Rosenbloom (2005) investigated and showed that customer satisfaction can vary between e-commerce checkout and after delivery, indicating delivery performance is a

[1] Temporal distance is measured as the difference between the actual delivery date and the review posting date in the number of days.

critical touchpoint in consumers' overall satisfaction with the sellers. Later research provided ample evidence, recognizing delivery performance (such as online time delivery, total delivery time) as a significant determinant of overall customer satisfaction in the e-commerce sector of retailing (Blut 2016; Collier and Bienstock 2006; Dholakia and Zhao 2010; Jain et al. 2015; Thirumalai and Sinha 2005; Vakulenko et al. 2019).

Traditionally, practitioners and researchers used surveys to understand the impact of delivery performance on customer satisfaction. However, as an alternative to customer satisfaction surveys, online reviews have become more popular during the past two decades (Rese et al. 2014). As discussed earlier, after delivering the orders, e-commerce platforms ask customers to rate their experience with sellers. As these reviews are solicited after delivery of the order, it is reasonable to expect that review ratings may reflect customer satisfaction due to delivery performance. As such, for evaluating the relationship between delivery performance and customer satisfaction, we use online review ratings proxy for customer satisfaction.

Further, extant research suggests that prospective consumers are increasingly relying on these online reviews for their purchase decisions (Changchit and Klaus 2020). Perhaps reflecting this, online reviews, especially the review ratings (Luca 2011), have impacted sales (e.g., Berger et al. 2010; Chevalier and Mayzlin 2006; Zhu and Zhang 2010). As such, several studies focused on determinants of online review ratings. However, most of these studies in online reviews literature were focused on reviews related to consumption experience with products.

A few studies have investigated the drivers behind review ratings in the context of e-commerce delivery. Qu et al. (2008) investigated the driving factors behind the review ratings on Yahoo! Merchants and found that the review ratings increased with on-time delivery. Park et al. (2012) examined the differential effects of pre- and post-transaction performance such as the fulfilled delivery, order tracking, and customer support on online review ratings using data from BizRate.com. Wu et al. (2021) focused on the relative effects of coupons vs. free shipping on review ratings for an e-commerce website. Li and Wang (2021), using reviews from Amazon, found that more retailers using free shipping increased product review ratings.

A few other studies focused on temporal distance in different contexts on review ratings (Huang et al. 2016; Li et al. 2019; Stamolampros and Korfiatis 2018; Yang et al. 2018; Wu et al. 2021). For example, in the context of restaurants, Huang et al. (2016) found that the temporal distance of review and consumption has a positive effect on review rating. Wu et al. (2021), in the context of e-commerce purchases, found that coupons increase review ratings through perceptions of monetary savings when the temporal distance of purchase and review is close but decrease review ratings through low perceived product quality when temporal distance is far.

The above literature suggests that factors related to e-commerce delivery and temporal distance of ratings have a role in review ratings. However, there are a few insights into questions like (1) Do customers reward sellers for early delivery similarly as they penalize late deliveries? (2) What is the role of the temporal distance in online review ratings in the context of delivery performance? These are the gaps that we address in this research.

Conceptual framework

This section will discuss the conceptual framework used for this study and propose corresponding hypotheses.

Consumer evaluation of early and late vs. same-day deliveries

For e-commerce orders, consumers expect to be notified when they receive their orders. As discussed earlier, sellers notify consumers with promised delivery dates after processing orders. In some cases, sellers deliver the orders earlier than notified delivery date. Some consumers receive on the same day of the notified delivery date, while some other consumers receive later. Against this backdrop, drawing from prospect theory, we argue that delayed deliveries are experienced with greater psychological force than early deliveries of similar magnitude, affecting review ratings (Chan et al. 2018; Gal and Rucker 2018; Kahneman and Tversky 2013; Thaler 2000).

Prospect theory (PT), particularly, asserts that consumers are more attuned to differences (relative to a reference point) and inclined to place greater weights on losses than gains of an equal magnitude (Kahneman and Tversky 2013; Tversky and Kahneman 1992). In these lines, Thaler (2000, p.137) argued that "losses hurt about twice as much as gains make us feel good." Extant research applied prospect theory to several different contexts. For example, PT is applied to the sustainable operation of transport infrastructure projects under government regulation (Ma et al. 2021), to pilot weather-related decision-making in an uncertain situation involving monetary gains and risk-seeking (Walmsley and Gilbey 2020), to explain investment strategies under uncertainly (Frazzini 2006), to product pricing strategy (Koszegi and Rabin 2006), to explain the effectiveness of promotional prices of leisure services (Crompton 2016), and to travelers' behavior in situations involving travel time uncertainty (Ramos et al. 2014). Similarly, asymmetric disconfirmation in satisfaction literature suggests that consumers' negative consumption experiences have a greater influence on their judgment than positive experiences (Anderson and Sullivan

1993; Darke et al. 2010; Mittal et al. 1998). A few studies focused on online review ratings have confirmed this phenomenon (Chan et al. 2018; Hao et al. 2010; Moe et al. 2011). The above literature suggests a steeper value function in the loss region as compared to gains. In other words, individuals evaluate potential losses differently from potential gains with equal magnitude.

Extending the above discussion to the present context of e-commerce retail transactions related to online reviews, this study argues that consumers evaluate early and late deliveries differently; specifically, that consumers use same-day delivery as a reference and compare early and late deliveries with the same magnitude (e.g., a day late/early), giving more weight to the late deliveries than the early deliveries. Based on the above information, it is predicted that consumers will penalize sellers more for late deliveries (e.g., one day late) than reward sellers for early deliveries with similar magnitudes (e.g., one day early). Therefore, the following hypothesis is proposed:

H1 The negative effect of a late delivery on a review rating will be larger than the positive effect from early delivery.

Role of temporal distance

The temporal distance can be defined as the time between two events. In the present context, temporal distance refers to the difference (in days) between an order delivery date and the review posting date. Extant literature shows that temporal distance is one of the key psychological distances that shape consumers' judgments (Adler and Sarstedt 2021; Mishra et al. 2020; Trope and Liberman 2010). Specifically, the construal level theory (CLT) suggests that consumers' memories related to events (e.g., consumption experience) are inconsistent with their perception of those events at the time they happen (Trope and Liberman 2010). When consumers recall their experiences with those events for making judgments, they tend to use different mental representations (construals) depending on their perceived psychological distance from those events (Trope and Liberman 2010; Yudkin et al. 2020; Wu et al. 2021). For instance, if consumers perceive the greater psychological distance between themselves and those events, they process events at higher levels of construal, abstractly. In contrast, if consumers perceive less psychological distance, they process the event at lower levels of construal, thoroughly or concretely. The above discussion suggests that events occurring at relatively closer temporal proximities are processed differently than events occurring in more distant proximities. Such differences in the processing of events are expected to influence consumer evaluations of those events (Adler and Sarstedt 2021; Mishra et al. 2020; Trope and Liberman 2010).

A recent study (Wang and Lin 2021) investigated the effect of temporal distance on consumer price evaluation. Specifically, the authors show that when the temporal distance is near, a nine-ending price may be perceived as larger than a price that is actually one dollar higher. However, the perceived magnitude of difference due to the left-digit effect has diminished when the temporal distance is distant. Another study (Liu et al. 2020), focused on temporal distance in consumer evaluation of online promotion activities and purchase behavior, found that temporal distance has a positive (negative) impact on purchase decision of high (low) involvement products. Choi et al. (2019) study found that consumers perceive partitioned pricing as more attractive than combined pricing for a temporally distant event. Su et al. (2022) found that consumers evaluate travel items into more superordinate (subordinate) categories when it comes to distance-future (near-future) trips. The above literature suggests that consumers, while evaluating different events/objects/decisions, use different mental representations (abstract vs. concrete) depending on the temporal distance.

A few studies, more relevant to the present study, have focused on the effect of temporal distance on review ratings given to consumption experiences (Huang et al. 2016; Li et al. 2019; Stamolampros and Korfiatis 2018; Yang et al. 2018; Wu et al. 2021). All these studies found that temporal distance to consumption experience is positively associated with review rating. The above literature suggests that temporal distance influences consumer evaluations.

In the context of the e-commerce sector of retailing transactions related to online reviews, it is expected that temporal distance (psychological distance) will affect how consumers construe their purchase experiences with sellers. Specifically, on the same day of delivery, the event (i.e., delivery performance) is psychologically very close and can be processed in a detailed, concrete manner. As the days pass, however, the event will be construed more abstractly, and detailed aspects of the delivery performance will gradually fade (Kim et al. 2008). In such situations, consumers are more likely to rely on their overall experiences with the sellers. Research also suggests that positive aspects of an event are more salient when processing abstractly, and pros are easier to think of when considering temporally distance versus close events (William et al. 2014; Huang et al. 2016). Therefore, the more time between the delivery date and the date on which they post their reviews, the greater the temporal distance. In this instance, consumers are more likely to make more positive review ratings. Thus, the following hypothesis is proposed:

H2 The temporal distance between the delivery date and the review posting date is positively associated with the review rating.

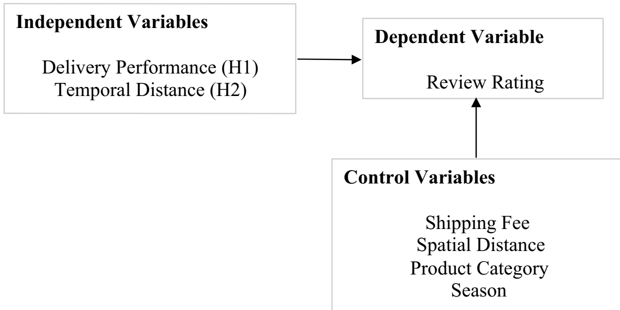

Fig. 2 Conceptual framework

Following literature, we have included several control variables for accounting observed heterogeneity: shipping fee (Kim and Cheon 2020; Kulkarni 2020; Ma 2017; Wu et al. 2021), spatial distance (Arentze and Timmermans 2001; Blut et al. 2018); product category (Kim 2020; Kim and Cheon 2020); and season (Lee et al. 2018). We present the conceptual framework in Fig. 2.

Methodology

Data

In order to test these hypotheses, data from e-commerce company in an emerging market were used for this study. After a customer purchases a product from the e-commerce store, the seller gets notified to fulfill that order. After the seller ships the product, customer is notified of the delivery date. Once the customer receives the product, sellers solicit reviews. The dataset primarily consists of order information, customer zip code, seller zip code, and respective online review related information. For this study, only orders with a single product from a single seller were used.

Variables

$rating_{io}$ is an ordinal variable used to indicate the customer evaluation of his/her purchase experience with the seller. Specifically, a review rating refers to the number of stars (1 to 5) allocated by customer i when indicating his/her assessment of his/her purchase experience with the seller for order o.

del_days_{io} is a continuous variable that represents the relative deliver days measured as the deviation of the *actual delivery date* from the *notified delivery date*. For example, if a retailer promised a two-day delivery and the package arrived after (within) one day, then the del_days_{io} would be equal to one day late (one day early).

$temp\_dist_{io}$ is a continuous variable that represents the temporal distance of the review rating by customer i for order o. It was measured as the difference between the *actual delivery date* and the *review posting date* in number of days.

$pcat\_conv_{io}$ is an indicator variable that represents whether the product purchased was from a convenience product category. Following the literature (Nguyen et al. 2019; Thirumalai and Sinha 2005), products were classified into three categories: convenience goods, shopping goods, and specialty goods. This classification is based on the volume and unit value of the products purchased. For example, consumers tend to buy convenience goods in large volumes and at low unit costs. $pcat\_conv_{io}$ was set to 1 if the product purchased by customer i in order o was from a convenience product category, 0 otherwise.

$pcat\_splt_{io}$ is an indicator variable that represents whether the product purchased was from a specialty product category. Similar to the convenience product category, $pcat\_splt_{io}$ was set to 1 if the product purchased by customer i in order o was from a specialty product category, 0 otherwise.

$spat\_dist_{io}$ is a continuous variable indicating the distance between the seller's city and customer's city. Using the latitude and longitude of the city zip codes, the spatial distance between the seller and customer was computed in kilometers. For example, if the customer and seller are from the same city (i.e., same zip codes), then the spatial distance between the seller's city and customer's city was equal to zero.

$shipfee_{io}$ is a continuous variable measured relative to order value. For example, if the shipping costs were $30 for an order worth $30, then the respective $shipfee_{io}$ was 1. Similarly, if shipping costs were $30 for an order worth $100, then the respective $shipfee_{io}$ was 0.3.

$season_{io}$ is an indicator variable that measured whether the purchase was completed during a holiday season (i.e., November and December). Therefore, if customer i placed order o in either November or December, then $season_{io}$ was set to 1, 0 otherwise.

Model

The dependent variable, review rating, is an ordinal variable consisting of five ordered categories. Therefore, the Ordinal Logit Model, the best-fitting statistical model for handling an ordered outcome (McCullagh 1980), was used to investigate the effect of delivery performance on a customer's review rating. Assuming that the utility of customer i is represented by an unobservable latent variable U_i, then customer i gives a certain rating $rating_{io}$ between 1 and 5 on the basis of U_i.

$$U_i = x'_i\beta + \varepsilon_i \tag{1}$$

$$\text{Rating}_{io} = j \text{ if } \theta_{j-1} < U_i \leq \theta_j, \tag{2}$$

Table 1 Descriptive statistics

Variable		Early delivery	Same-day delivery	Late delivery	Full sample
Rating	$rating_{io}$	4.411	4.156	3.452	4.006
Shipping fee	$shipfee_{io}$	0.260	0.263	0.267	0.263
Spatial distance	$spat\_dist_{io}$	488.956	507.113	471.895	489.321
Convenience goods	$pcat\_conv_{io}$	0.236	0.290	0.271	0.265
Shopping goods	$pcat\_shop_{io}$	0.650	0.586	0.576	0.604
Specialty goods	$pcat\_splt_{io}$	0.115	0.124	0.083	0.131
Season	$season_{io}$	0.070	0.064	0.083	0.072
Relative delivery days	del_days_{io}	11.175	0	1.920	NA
Temporal distance	$temp\_dist_{io}$	1.000	0.990	0.678	0.890
Number of observations		314	314	314	942

where x'_i is matrix of independent variables related to customer i, β is a vector of the coefficients, $\{\theta\}$ are thresholds, and ε_i follows a logistic distribution. The cumulative distribution of ε_i is

$$F(z) = \frac{e^z}{1 + e^z} \tag{3}$$

then the probability that customer i will give rating j is

$$p\left(\text{Rating}_i = j\right) = p\left(\theta_{j-1} < U_i \le \theta_j\right) = F\left(\theta_j - x'_i\beta\right) - F(\theta_{j-1} - x'_i\beta). \tag{4}$$

Matched sample

The challenge in using secondary data is that the incidences in the same-day delivery, early delivery, and late delivery groups are not randomly selected. The matched sample approach essentially attempts to address this issue by creating a pseudo-random sample. In recent years, the *propensity score matching* approach has gained popularity because it allows a refined matching process along multiple characteristics (Dehejia and Wahba 1999). In essence, this approach attempts to correct for the non-random treatment effect by matching a treated incidence (early or late delivery) to an untreated incidence (same-day delivery) that has similar observed characteristics. Although the results from the matched samples do not establish a causal relationship, they provide evidence that the observed relationship is related to the delivery performance and temporal distance rather than other observed characteristics.

The R package TriMatch (Bryer 2013) was used for this study as it estimates propensity scores and finds the best matched triplets with replacement. Shipping fee, spatial distance, product type, and season were used for the selection of the matched samples. The details for the propensity score matching procedure are provided in Appendix. The resulting matched sample consisted of 942 incidences (314 per each group).

Model estimation

For ease of comparison, two models were estimated. First, a model for early deliveries vs. same-day deliveries was estimated. In other words, a model (Eqs. 1–4) was estimated for the data consisting of 314 early deliveries and 314 matched same-day deliveries. Similarly, another model was estimated for late deliveries vs. same-day deliveries. For this model, data consisting of 628 observations (314 late deliveries and 314 matched same-day deliveries) were utilized.

For inference regarding the parameters, the PROC MCMC method in SAS with highly diffuse priors for the model parameters was used. Specifically, the Gaussian prior distribution for all parameters in $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ was used. For each model (early deliveries and late deliveries), an MCMC chain with 50,000 samples was simulated and the first 10,000 samples were discarded as burn-in. From the remaining samples, every 10th iteration was selected, allowing for a retention of 4000 samples for posterior inference of means and standard deviations of the parameter estimates.

Results

Descriptive statistics and correlations are provided in Tables 1 and 2. For the early delivery sample, the average review rating was 4.411. While same-day deliveries received a lower average rating (4.156) than early deliveries, late deliveries had the lowest average rating (3.452). One of the key variables was relative delivery days (i.e., the number of days before (after) the notified delivery date for early (late) deliveries). For same-day deliveries, the relative delivery days was zero, which was the reference point for consumers to compare their deliveries against. Early deliveries were delivered, on average, 11.175 days before this date and late deliveries were delivered, on average, 1.92 days after the notified delivery date.

The most interesting insight from this analysis was related to the temporal distance of the review rating. Consumers

Table 2 Correlations

Variable	1	2	3	4	5	6	7	8
Rating	1							
Shipping fee	– 0.050	1						
Spatial distance	0.007	0.212	1					
Convenience goods	0.012	0.090	– 0.057	1				
Shopping goods	0.025	– 0.242	0.027	– 0.742	1			
Specialty goods	– 0.052	0.233	0.036	– 0.233	– 0.479	1		
Season	– 0.042	– 0.044	– 0.075	0.009	– 0.017	0.014	1	
Temporal distance	0.159	– 0.007	0.000	– 0.009	0.040	– 0.047	0.036	1
Number of observations	942							

Correlations > \|0.090\| are significant at 0.05

Table 3 Results for early delivery model

Variable	Mean	SD	Mean	SD
Shipping fee	– 0.527	0.427	– 0.532	0.418
Log (spatial distance)	– 0.016	0.067	– 0.018	0.066
Convenience goods	– 0.164	0.191	– 0.156	0.186
Specialty goods	0.100	0.297	0.107	0.292
Season	– 0.506*	0.292	– 0.523*	0.306
Relative delivery days	0.245***	0.066	0.248***	0.066
Temporal distance			0.055	0.492
Number of observations	628			
– Log Likelihood	694		694	

Sig. ***0.01, **0.05, *0.10

Table 4 Results for late delivery model

Variable	Mean	SD	Mean	SD
Shipping fee	– 0.079	0.367	– 0.084	0.354
Log (spatial distance)	– 0.031	0.061	– 0.037	0.063
Convenience goods	– 0.150	0.172	– 0.143	0.174
Specialty goods	0.023	0.243	0.043	0.231
Season	– 0.314	0.285	– 0.312	0.297
Relative delivery days	– 0.864***	0.136	– 0.610***	0.150
Temporal distance			1.109***	0.286
Number of observations	628			
– Log likelihood	856		849	

Sig. ***0.01

who received early deliveries posted their reviews, on average, a day (1.000) after they received their deliveries. However, consumers who received late deliveries posted their reviews, on average, 0.678 days after they received their deliveries (i.e., less than a day). Four variables (i.e., shipping fee, spatial distance, type of good, season) were used to select the matched samples (see Appendix). Reflecting on this information, perhaps the matched samples had similar descriptive statistics for these four variables across the three samples.

The correlations between different variables are presented in Table 2. An interesting correlation was between the review rating and temporal distance. As temporal distance increased, so did the rating.

Results for the early deliveries model

To estimate the effect of early deliveries on review ratings, Eqs. 1–4 were utilized for the sample of 314 early-day deliveries and 314 same-day deliveries. The parameter estimates for the early deliveries model are displayed in Table 3. Consistent with this study's expectations, the coefficients of rdd_{io} had the expected sign and significance (i.e., 95% posterior distribution of the difference of means excluding zero). The coefficient of rdd_{io} was statistically significant and positive (0.119, sig. = 0.05), implying that early deliveries were positively associated with the review ratings. The coefficient of $temo\_dist_{io}$ was positive, but not statistically significant (0.055).

Results for the late deliveries model

To estimate the effect of late deliveries on review ratings, Eqs. 1–4 were utilized for the sample of 314 late-day deliveries and 314 same-day deliveries. The parameter estimates for the late deliveries model are displayed in Table 4. Consistent with this study's expectations, the coefficients of rdd_{io} and $temp\_dist_{io}$ had the expected signs and significance (i.e., 95% posterior distribution of the difference of means excluding zero). The coefficient of rdd_{io} was statistically significant and negative (– 0.610, sig. = 0.05), implying that late deliveries were negatively associated with the review ratings. In other words, the longer the delivery was past the expected date, the lower the review rating. The coefficient of $temo\_dist_{io}$ was positive and statistically significant (1.109, sig. = 0.05), implying that the temporal distance of the review rating was positively associated with the review rating.

The above results supported this study's hypotheses. First, the comparison of the parameters for relative delivery days in the early (0.119, sig. = 0.05) and late delivery models (− 0.610, sig. = 0.05) indicated that the negative effect of late deliveries on consumers' review ratings was larger than the positive effect of early deliveries with similar magnitudes. This result was consistent with the prediction from the Prospect theory, which suggests that the effect from the perceived loss was stronger than the effect from the perceived gains (Chan et al. 2018; Kahneman and Tversky 2013). This finding supported the H1.

Second, the parameters for temporal distance in early (0.055) and late delivery models (1.109 sig. = 0.05) indicated that the temporal distance of the review rating was positively associated with the review rating. This result was consistent with the predictions from the Construal Level theory (Huang et al. 2016; Trope and Liberman 2010; Yudkin et al. 2020; Wu et al. 2021), which indicates that consumers with the greater distance between the event and themselves, more favorable rating. However, the effect of temporal distance for early deliveries was not significant. These findings support H2 partially.

Conclusions

Consumers are increasingly using online reviews in their purchase decisions. As such, retailers are soliciting reviews from their customers after delivering the products. In the context of e-commerce delivery performance, this study aimed to answer two key research questions: (i) do customers reward sellers for early delivery in the same way they penalize them for late deliveries? and (ii) what is the role of the temporal distance in online review ratings in the context of delivery performance? E-commerce data from a particular company were used within an emerging economy for this study. To establish the causal relationship, a propensity score-matched sample was used. Two models: early deliveries and late deliveries were estimated. The empirical results supported the proposed hypotheses. Specifically, we hypothesized (in H1) that the negative effect of a late delivery on a review rating will be larger than the positive effect from early delivery. The study findings indicated that customers gave much lower (a little higher) ratings to orders delivered late (early) than to orders delivered on time. These results supported H1 and answered the first research question of the study. Further, we hypothesized (in H2) that the temporal distance between the delivery date and the review posting date is positively associated with the review rating. Consistent with our prediction, the study's findings indicated that temporal distance was positively associated with review ratings. This result supported H2 and answered the second research question of the study.

Overall, the main contributions of this study to the existing body of literature are threefold. Specifically, the research highlights the value of investigating relative delivery days and the temporal distance of review ratings. It also contributes to the literature on prospect theory, construal level theory, and online reviews.

First, the findings extend an array of prior studies using prospect theory (Chan et al. 2018; Kahneman and Tversky 2013) and indicated that customers give much lower (a little higher) ratings to orders delivered late (early) than to orders delivered on time, consistent with the predictions of loss aversion from prospect theory. Second, the findings extend an array of prior studies using Construal Level theory (e.g., Huang et al. 2016; Trope and Liberman, 2010), which has been researched in wide range of disciplines, such as marketing, organizational study, psychology, and education. Third, this study extended the online review literature by applying prospect theory and construal level theory in a context of significant relevance to practitioners and academics in a growing sector of retail: e-commerce. The majority of prior studies on online reviews has focused on understanding the consequences of review ratings, such as sales (Chevalier and Mayzlin 2006; Zhu and Zhang 2010). This study, in contrast, examined the antecedents of online review ratings (e.g., Chen and Kirmani 2015; Huang et al. 2016).

Managerial implications

E-commerce delivery performance is viewed as one of the core features of customer satisfaction (Vakulenko et al. 2019) reflected in online reviews. As prospective consumers increasingly use reviews in their purchase decisions (Chevalier and Mayzlin 2006), sellers would benefit from having more favorable ratings. As such, sellers solicit reviews after delivery of orders. Lemon and Verhoef (2016) suggest sellers can influence customer satisfaction at touchpoints owned by sellers, such as providing a seamless shopping experience at e-commerce platforms, obtaining favorable ratings. However, customer satisfaction in the case of late deliveries is beyond sellers' control; therefore, sellers require some degree of management and a strategy for soliciting ratings. The study's findings will help firms devise when to solicit online reviews in cases of late deliveries.

In general, sellers on e-commerce sites solicit online reviews immediately after a shipment has been delivered (i.e., on the same day as the delivery). The findings of this study indicated that consumers who received late deliveries posted their ratings, on average, on the same day and gave much lower ratings. Therefore, the findings indicated that the temporal distance of the review ratings was positively associated with the review ratings. In other words, if customers post their reviews a day after receiving their deliveries, they may give a higher rating.

Therefore, sellers may avoid getting penalized for late deliveries by soliciting online reviews later rather soliciting online reviews on the same day of delivery. Second, the findings suggested that retailers do not gain much benefit in terms of favorable ratings for early deliveries. Therefore, sellers can avoid promising longer delivery dates and delivering early.

Limitations and future research

Although this study has provided a better understanding of the effects of delivery performance and temporal distance on review ratings in the context of e-commerce, some limitations exist that could be addressed in future studies. First, this study used secondary data. Although a propensity score-matched sample was employed to investigate causality, the data are still considered weak in regard to discovering the underlying reasons for the phenomenon. Future studies should employ experimental methods that provide more detailed reasoning about the relationships identified in the study. Second, other issues related to delivery performance, such as the quality of the shipment (e.g., damages to the product during shipment) and the customer's relationship duration with the seller, could be added to the model in future studies to achieve greater explanatory power in regard to consumer satisfaction ratings. Third, as this study used data from an emerging economy, generalizations related to the findings should be made with caution. In the future, other researchers should conduct cross-country comparison studies so that the results can be more easily generalized.

Appendix: propensity score matching

For selecting a matched sample, we use R package TriMatch, which estimates the propensity scores and finds best matching triplets. We employ different variables as matching variables: Shipping fee, Spatial distance, Type of product (e.g., convenience, shopping, specialty, and season. Further, we compute the Standardized Bias (SB), a widely used techniques to ensure the balance of the samples (Harmeling et al. 2015). A SB score below 0.1 indicates the PSM is effective in balancing the distributions of the covariates. We report the SB Scores in Table 5.

$$SB_{\text{match}} = \frac{M_1(X_k) - M_0(X_k)}{\sqrt{0.5(V_1(X_k) - V_0(X_k)}}, \tag{5}$$

Table 5 Standardized bias scores for matched sample

Variable	Early delivery to be more Same-day delivery Standardized bias	Late delivery to be more Same-day delivery Standardized bias
Shipping fee	− 0.01	0.02
Log (spatial distance)	− 0.04	− 0.08
Convenience goods	− 0.12	− 0.04
Specialty goods	− 0.03	0.08
Season	0.03	0.07

where $M_1^A(X_k)[V_1(X_k)]$ and $M_0^A(X_k)[V_0(X_k)]$ are the means [variances] of the observable k for the treated group and the matched control group.

References

Adler, S., and M. Sarstedt. 2021. Mapping the jungle: A bibliometric analysis of research into construal level theory. *Psychology & Marketing* 38 (9): 1367–1383.

Anderson, E.W., and M.W. Sullivan. 1993. The antecedents and consequences of customer satisfaction for firms. *Marketing Science* 12 (2): 125–143.

Appsfluer. 2021. The state of ecommerce app marketing. Retrieved April 10, 2022, https://www.appsflyer.com/infograms/state-of-ecommerce-2021/

Arentze, T.A., and H.J. Timmermans. 2001. Deriving performance indicators from models of multipurpose shopping behavior. *Journal of Retailing and Consumer Services* 8 (6): 325–334.

Berger, J., A.T. Sorensen, and S.J. Rasmussen. 2010. Positive effects of negative publicity: When negative reviews increase sales. *Marketing Science* 29 (5): 815–827.

Bhatnagar, A., S. Misra, and H.R. Rao. 2000. On risk, convenience, and internet shopping behavior. *Communications of the ACM* 43 (11): 98–105.

Blut, M. 2016. E-service quality: Development of a hierarchical model. *Journal of Retailing* 92 (4): 500–517.

Blut, M., C. Teller, and A. Floh. 2018. Testing retail marketing-mix effects on patronage: A meta-analysis. *Journal of Retailing* 94 (2): 113–135.

Bryer, J.M. 2013, May. TriMatch: An R package for propensity score matching of non-binary treatments. In *The R user conference, useR* (pp. 10–12).

Carter, R. 2022. The ultimate list of online review statistics for 2022. https://findstack.com/online-review-statistics/. Accessed March 04, 2022.

Chan, T., Z. Liu, & W. Zhang. 2018. Delivery service, customer satisfaction and repurchase: Evidence from an online retail platform. *Customer satisfaction and repurchase: Evidence from an online retail platform (September 15, 2018)*.

Changchit, C., and T. Klaus. 2020. Determinants and impact of online reviews on product satisfaction. *Journal of Internet Commerce* 19 (1): 82–102.

Chen, Y.J., and A. Kirmani. 2015. Posting strategically: The consumer as an online media planner. *Journal of Consumer Psychology* 25 (4): 609–621.

Chevalier, J.A., and D. Mayzlin. 2006. The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research* 43 (3): 345–354.

Choi, J., D.E. Bolton, and M. Grishin. 2019. The moderating effect of temporal distance on partitioned vs combined pricing. *Journal of Consumer Marketing* 36 (5): 529–538.

Collier, J.E., and C.C. Bienstock. 2006. Measuring service quality in e-retailing. *Journal of Service Research* 8 (3): 260–275.

Crompton, J.L. 2016. Implications of prospect theory for the pricing of leisure services. *Leisure Sciences* 38 (4): 315–337.

Darke, P.R., L. Ashworth, and K.J. Main. 2010. Great expectations and broken promises: Misleading claims, product failure, expectancy disconfirmation and consumer distrust. *Journal of the Academy of Marketing Science* 38 (3): 347–362.

Dehejia, R.H., and S. Wahba. 1999. Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association* 94 (448): 1053–1062.

Dholakia, R.R., and M. Zhao. 2010. Effects of online store attributes on customer satisfaction and repurchase intentions. *International Journal of Retail & Distribution Management* 38 (7): 482–496.

eMarketer. 2022. Global ecommerce forecast 2022. https://www.emarketer.com/content/global-ecommerce-forecast-2022. Accessed 11 April 2022.

Frazzini, A. 2006. The disposition effect and underreaction to news. *Journal of Finance* 61 (4): 2017–2046.

Gal, D., and D.D. Rucker. 2018. The loss of loss aversion: Will it loom larger than its gain? *Journal of Consumer Psychology* 28 (3): 497–516.

Hao, Y., Q. Ye, Y. Li, Z. Cheng. 2010, January. How does the valence of online consumer reviews matter in consumer decision making? Differences between search goods and experience goods. In *2010 43rd Hawaii international conference on system sciences*. IEEE, pp. 1–10

Harmeling, C.M., R.W. Palmatier, M.B. Houston, M.J. Arnold, and S.A. Samaha. 2015. Transformational relationship events. *Journal of Marketing* 79 (5): 39–62.

Huang, N., G. Burtch, Y. Hong, and E. Polman. 2016. Effects of multiple psychological distances on construal and consumer evaluation: A field study of online reviews. *Journal of Consumer Psychology* 26 (4): 474–482.

Jain, N.K., H. Gajjar, B.J. Shah, and A. Sadh. 2015. A conceptual framework for measuring e-fulfillment dimensions: A consumer perspective. *Journal of Internet Commerce* 14 (3): 363–383.

Jiang, P., and B. Rosenbloom. 2005. Customer intention to return online: Price perception, attribute-level performance, and satisfaction unfolding over time. *European Journal of Marketing* 39 (1/2): 150–174.

Jin, B., and J.Y. Park. 2006. The moderating effect of online purchase experience on the evaluation of online store attributes and the subsequent impact on market response outcomes. *ACR North American Advances* 33: 203–211.

Kahneman, D., & A. Tversky. 2013. Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*, pp. 99–127.

Kim, R.Y. 2020. The influx of skeptics: An investigation of the diffusion cycle effect on online review. *Electronic Markets* 30 (4): 821–835.

Kim, O., and H.J. Cheon. 2020. Product category and shopping options of logistic service quality. *The Journal of Distribution Science* 18 (8): 113–125.

Kim, K., M. Zhang, and X. Li. 2008. Effects of temporal and social distance on consumer evaluations. *Journal of Consumer Research* 35 (4): 706–713.

Koszegi, B., and M. Rabin. 2006. A model of reference-dependent preferences. *Quarterly Journal of Economic* 121 (4): 1133–1165.

Kulkarni, A.A. 2020. No shipping fees or free shipping? Impact of temporal proximity on the relative effectiveness of promotional framing. *Journal of Promotion Management* 26 (1): 50–74.

Kuhn, S.W., and D.J. Petzer. 2018. Fostering purchase intentions toward online retailer websites in an emerging market: An SOR perspective. *Journal of Internet Commerce* 17 (3): 255–282.

Lee, S., J. Park, H. Hyun, S. Back, S. Bryan Lee, F. Gunn, and J. Ahn. 2018. Seasonality of consumers' third-party online complaining behavior. *Social Behavior and Personality: An International Journal* 46 (3): 459–470.

Lemon, K.N., and P.C. Verhoef. 2016. Understanding customer experience throughout the customer journey. *Journal of Marketing* 80 (6): 69–96.

Li, D., and H. Wang. 2021. Effects of retailers' free shipping promotions on manufacturers' product sales and product review ratings in multichannel retailing. *Managerial and Decision Economics* -. https://doi.org/10.1002/mde.3499.

Li, H., Z. Zhang, F. Meng, and Z. Zhang. 2019. "When you write review" matters: The interactive effect of prior online reviews and review temporal distance on consumers' restaurant evaluation. *International Journal of Contemporary Hospitality Management* 31 (3): 1273–1291.

Liu, Q., X. Zhang, S. Huang, L. Zhang, and Y. Zhao. 2020. Exploring consumers' buying behavior in a large online promotion activity: The role of psychological distance and involvement. *Journal of Theoretical and Applied Electronic Commerce Research* 15 (1): 66–80.

Luca, Michael. 2011. Reviews, reputation, and revenue: The case of yelp.com. https://www.hbs.edu/ris/Publication%20Files/12-016_a7e4a5a2-03f9-490d-b093-8f951238dba2.pdf. Accessed 07 June 2021.

Ma, S. 2017. Fast or free shipping options in online and Omni-channel retail? The mediating role of uncertainty on satisfaction and purchase intentions. *The International Journal of Logistics Management* 28 (4): 1099–1122.

Ma, C., Y. Chen, and Y. Zhang. 2021. Simulation analysis of the evolution of sustainable operation of transport infrastructure projects under government regulation based on prospect theory and BP neural network. *Scientific Programming*. https://doi.org/10.1155/2021/6868487.

McCullagh, P. 1980. Regression models for ordinal data. *Journal of the Royal Statistical Society: Series B (methodological)* 42 (2): 109–127.

Mishra, A.N., A. Raj, and A.K. Pani. 2020. Construal level research in decision making: Analysis and pushing forward the debate using bibliometric review and thematic analysis. *American Business Review* 23 (1): 8.

Mittal, V., W.T. Ross Jr., and P.M. Baldasare. 1998. The asymmetric impact of negative and positive attribute-level performance on overall satisfaction and repurchase intentions. *Journal of Marketing* 62 (1): 33–47.

Moe, W.W., D.A. Schweidel, and M. Trusov. 2011. What influences customers' online comments. *MIT Sloan Management Review* 53 (1): 14.

Nguyen, D.H., S. de Leeuw, W. Dullaert, and B.P. Foubert. 2019. What is the right delivery option for you? Consumer preferences for delivery attributes in online retailing. *Journal of Business Logistics* 40 (4): 299–321.

Park, I., J. Cho, and H.R. Rao. 2012. The effect of pre-and post-service performance on consumer evaluation of online retailers. *Decision Support Systems* 52 (2): 415–426.

Qu, Z., H. Zhang, and H. Li. 2008. Determinants of online merchant rating: Content analysis of consumer comments about Yahoo merchants. *Decision Support Systems* 46 (1): 440–449.

Ramos, G.M., W. Daamen, and S. Hoogendoorn. 2014. A state-of-the-art review: Developments in utility theory, prospect theory and regret theory to investigate travellers' behaviour in situations involving travel time uncertainty. *Transport Reviews* 34 (1): 46–67.

Rese, A., S. Schreiber, and D. Baier. 2014. Technology acceptance modeling of augmented reality at the point of sale: Can surveys be replaced by an analysis of online reviews? *Journal of Retailing and Consumer Services* 21 (5): 869–876.

Smith Brain Trust (SBT). 2022. The pandemic, the inflation and the way wed shop now. https://www.rhsmith.umd.edu/research/pandemic-inflation-and-way-we-shop-now. Accessed 11 April 2022.

Stamolampros, P., and N. Korfiatis. 2018. Exploring the behavioral drivers of review valence. *International Journal of Contemporary Hospitality Management* 30 (10): 3083–3099.

Statista. 2021. Retail e-commerce sales worldwide. https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/. Accessed 5 April 2021.

Statista. 2022. E-commerce worldwide—Statistics & facts. https://www.statista.com/topics/871/online-shopping/. Accessed 10 April 2022.

Su, L., X. Yang, and Y. Huang. 2022. Tourists' goal-directed behaviors: The influences of goal disclosure, goal commitment, and temporal distance. *Journal of Travel Research* 61 (4): 940–960.

Thaler, R.H. 2000. From Homo economicus to Homo sapiens. *The Journal of Economic Perspectives* 14: 133–141.

Thirumalai, S., and K.K. Sinha. 2005. Customer satisfaction with order fulfillment in retail supply chains: Implications of product type in electronic B2C transactions. *Journal of Operations Management* 23 (3–4): 291–303.

Trope, Y., and N. Liberman. 2010. Construal-level theory of psychological distance. *Psychological Review* 117 (2): 440.

Tversky, A., and D. Kahneman. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5 (4): 297–323.

UNCTAD. 2022. COVID-19 and e-commerce a global review. https://unctad.org/webflyer/covid-19-and-e-commerce-global-review. Accessed 10 April 2022.

Vakulenko, Y., P. Shams, D. Hellström, and K. Hjort. 2019. Online retail experience and customer satisfaction: The mediating role of last mile delivery. *The International Review of Retail, Distribution and Consumer Research* 29 (3): 306–320.

Walmsley, S., and A. Gilbey. 2020. Applying prospect theory to pilot weather-related decision-making: The impact of monetary and time considerations on risk taking behaviour. *Applied Cognitive Psychology* 34 (3): 685–698.

Wang, J.W., and C.H. Lin. 2021. Temporal distance and motivation matter: Effects of psychological distance and left-digit in price evaluation. *Current Psychology*. https://doi.org/10.1007/s12144-021-02359-2.

Williams, L.E., R. Stein, and L. Galguera. 2014. The distinct affective consequences of psychological distance and construal level. *Journal of Consumer Research* 40 (6): 1123–1138.

Wu, J., H. Zhao, and H. Chen. 2021. Coupons or free shipping? Effects of price promotion strategies on online review ratings. *Information Systems Research* 32 (2): 633–652.

Yang, Y., L. Wu, and W. Yang. 2018. Does time dull the pain? The impact of temporal contiguity on review extremity in the hotel context. *International Journal of Hospitality Management* 75: 119–130.

Yudkin, D.A., N. Liberman, C. Wakslak, and Y. Trope. 2020. Better off and far away: Reactions to others' outcomes depends on their distance. *Organizational Behavior and Human Decision Processes* 156: 13–23.

Zeithaml, V.A., A. Parasuraman, and A. Malhotra. 2002. Service quality delivery through web sites: A critical review of extant knowledge. *Journal of the Academy of Marketing Science* 30 (4): 362.

Zhu, F., and X. Zhang. 2010. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing* 74 (2): 133–148.

Dr. Prashanth Ravula is an Assistant Professor of Marketing at the NJCU School of Business, New Jersey City University. He received his Ph.D. (Marketing) from the University of Wisconsin - Milwaukee. Dr. Ravula's research interest is in Digital Marketing and has published his research in the *International Journal of Research in Marketing* and *Research in Transportation Economics*.