Cross-Platform Analysis of Seller Performance and Churn for Ecommerce Using Artificial Intelligence

Anuj Batta, Indian Institute of Technology, Delhi, India
https://orcid.org/0000-0002-8583-5781
Arpan Kumar Kar, Indian Institute of Technology, Delhi, India*
Shyamali Satpathy, IILM University, Greater Noida, India

ABSTRACT

Suppliers and sellers play a crucial role in the ecommerce ecosystem. Sellers and ecommerce firms use social media to increase user engagement, visibility, and sales. Seller ratings are as important as the product ratings on ecommerce platforms to drive buying decisions. Based on sellers’ actions on social media, this study examines seller turnover and disengagement on e-commerce platforms. The study has been supported by the justice theory. Seller reviews and ratings from e-commerce platforms and conversations from social media platforms have been gathered. Using natural language processing, machine learning, partial least squares (PLS) path analysis, and statistical inferences, objectives of the study are met. The study offers recommendations for both practitioners and researchers. The sellers must focus more on interaction and communication than marketing. Through a longitudinal analysis, the study also establishes that ecommerce organizations can use seller social media performance as a predictor of future seller churn and disengagement so they can take the necessary remedial action.

KEYWORDS:
Cross platform Impact, E-Commerce, Machine learning, Natural language processing, Seller Churn, Seller disengagement, Seller Review Consensus, Social Media Content Type

INTRODUCTION

Almost half of the online purchase decisions are based on online reviews and ratings given by customers in the ecommerce context (Mosteller & Mathwick, 2016). Customer reviews and ratings are more trusted and considered highly credible in comparison to the advertisements and promotions by the brands or sellers (Thakur, 2018). It is not only the product and brand rating or reviews which are important, however, the seller reviews and ratings also have an impact on the purchase decision.
by customers (Chen et al., 2008). Thus, retailers, sellers, and suppliers continue to invest heavily in various techniques and methods, apart from providing the best possible customer experience, to get better online ratings and reviews. The available literature in this domain indicates that customer engagement is a primary driver which leads to several effective outcomes, including positive ratings and reviews leading to a positive buying decision by customers (Brodie et al., 2013; Kaur et al., 2020; Xiao & Li, 2019).

Sellers and suppliers use social media platforms as one of the primary channels for customer engagement and interaction. Millions of customer interactions occur every day on social media platforms such as Facebook, Twitter, YouTube, etc. (Appel et al., 2019; Gupta and Ramachandran, 2021). There is a robust interrelated phenomenon between customer relationship management, social media technologies, customer engagement, positive word of mouth and brand loyalty (Dewnarain et al., 2019). With the exponential growth in social media users and the trend reflecting a fundamental shift from primarily company-customer interactions to customer-customer interactions that impact company-customer relationships, it is imperative for sellers to have an effective social media customer engagement plan (Agnihotri, 2020). Social media appeared as, and continues to be a business phenomenon. Increasingly, current and prospective customers use social media to communicate about the products and services they plan to buy. At the same time, sellers or suppliers are also using the social media platform to express their products and services and resolve customer issues (Kaplan & Haenlein, 2010; Laroche et al., 2013).

Social media customer engagement for most companies began on external social media sites such as Facebook, Twitter, LinkedIn etc. Today, 90% of mid-size and large enterprises spent at least 11% of their total marketing expenses on these platforms (de Oliveira Santini et al., 2020). As this is a huge cost for companies, there has been an immense interest in evaluating the impact and return on investment of this spend, however, it is not an easy and a straightforward mechanism. There are several issues including limitation of the social media analysis tools to only include keywords, topics or sentiment analysis, difficulty in identifying the right questions, suggestions, desires etc., and most importantly the failure in considering the communication in the conversation context (Abbasi et al., 2018). Due to the high cost and ineffective ways to measure the impact, companies quickly realized multifold benefits of supporting customer communities on their own website. Customers can raise questions or complaints about the product either on a regular website from where they buy the product or on the social media platform. Also, companies are utilizing these platforms to communicate their brands and product to customers. Most companies that engage with customers online, do so on both of these avenues, i.e. social media platforms and their own websites (Connell et al., 2019; Kumar et al., 2022).

Looking at the literature in this domain, it is evident that there is a significant amount of work done in the area of customer engagement on social media, the impact of product or seller reviews, word of mouth, customer loyalty, buying intention, and customer retention (de Oliveira Santini et al., 2020; Rosado-Pinto & Loureiro, 2020). However most of the work has been unidimensional which looks at the impact of customer engagement independently for a particular objective and on a single platform (Gligor et al., 2019; Hollebeek et al., 2021; Yen et al., 2020). The current study explores a very interesting field of work which is at a nascent stage, where the interrelationship between customer engagement on social media platforms, and reviews or rating intentions on ecommerce platforms is looked at (Thakur, 2018).

This study contributes to the e-commerce literature and industry practice through a cross platform behavior analysis. It explores the impact of customer engagement by the sellers through a social media platform i.e. Twitter, on the reviews and rating intention of customers on an ecommerce platform, Amazon. Twitter has been chosen for this study because it is an excellent source of opinion data from customers in large numbers and in the form of microblogging, which can give effective and efficient insights needed for this study (Giachanou & Crestani, 2016). This study also establishes a model to predict sellers’ churn on the ecommerce platform. It explores the degree of engagement that
a seller demonstrates with the buyers on Twitter and the overall ratings given to the same seller on Amazon. It will also help to analyze the degree of consensus in the reviews provided by the buyers to these sellers, both for positive and negative reviews, along with the average review polarity. Most importantly, it establishes the relationship between the sellers’ behavior or engagement on Twitter and the sellers’ churn on Amazon. Using a combination of methods from big data analytics, natural language processing, machine learning, and statistical tests for inferences, the research objectives has been achieved. To summarize, following are the objectives of this study:

1. To determine the relationship between sellers’ engagement with the customers (including customer orientation) on Twitter, and the overall seller ratings on Amazon,
2. To identify the degree of consensus in the reviews given by customers for the sellers on Amazon,
3. To identify the relationship between the degree of consensus in the reviews and the overall sellers’ ratings,
4. To establish the relationship between sellers’ behaviour on Twitter, and seller disengagement, including churn on Amazon.

This study uses Justice theory as the theoretical model, which is shown to be useful in identifying antecedents to psychological processes and explaining individual’s reaction to conflict situation. In the literature, three types of justice, namely distributive, procedural, and interactional justice, are suggested as relevant to explain the psychological process of social interaction. In the context of seller’s engagement on social media with the customers, or in the buyer-supplier relationship performance, distribution justice can take the form of acknowledgement of inconvenience or issues and taking remedial actions (Liu et al., 2012). When customers feel being heard and perceive fairness in handling their complaints, customers tend to be satisfied, have a higher trust for the seller and the company. Thus, they tend to spread positive word of mouth about it and give a positive rating to the seller (Luca, 2017).

The rest of the article is structured as follows. Section 2 represents the contextual background and literature review, followed by section 3, which is the theoretical background and hypothesis development. Section 4 describes the research methodology and data collection. Analysis and findings are discussed in section 5 after which the results are summarized. Discussion on the findings and insights are in section 6. Finally, the article ends with a conclusion in section 7, and limitations and further opportunities in this area of research in section 7.

BACKGROUND

This section is divided into the following three subsections: customer engagement through social media; seller ratings and reviews on ecommerce platforms; customer intention of rating and reviews. The first subsection highlights the usage of social media platforms by the sellers for customer engagement. Next subsection introduces the sellers ratings and reviews on ecommerce platforms. The last subsection defines the importance and intention of customers for giving ratings or writing reviews for the sellers on ecommerce platforms.

Customer Engagement Through Social Media

Customer Engagement can be defined as the process of identifying, attracting, absorbing, motivating, and interacting with customers for achieving business objectives (Harrigan et al., 2017; Pansari & Kumar, 2017). Social media has become the pervasive channel for interactions and connections for customer engagement (Grover and Kar, 2020). These interactions are not only limited to customers but multiple stakeholders including organizations, stakeholders, and non-customers (Carlson et al., 2018; Shawky et al., 2020). Since every brand’s effort on social media involves substantial investments, they
are left with a rudimentary question: Is it worth investing into social media and does this investment translate into better customer acquisition, retention, and profits? However, most of the companies still struggle in evaluating the impact of maintaining their personalized brand pages or these investments on all the perceived benefits. They also fail to clearly see or calculate the return on investment (Lal et al., 2019; Maecker et al., 2016). Consequently, companies following a multichannel customer management approach (Neslin and Shankar, 2009; Neslin et al., 2006) along with investing heavily in social media by establishing brand fan pages on which companies convey brand-related content (i.e., brand posts) that users can like, comment on, or share. These open, social media-enabled interactions between a company and its customers help create a community that revolves around the brand and fosters the brand relationship (McAlexander et al., 2002; Muniz and O’Guinn, 2001).

Past research has examined and suggested possible benefits of customer engagement through social media. It includes a more profound and positive brand image (Quynh, 2019), brand value (France et al., 2016), better revenue (Tu et al., 2018) and process improvements (Kushwaha et al., 2021). Recent research has advocated the value of social media interactions for companies that propose conceptual frameworks to manage brands (Gensler et al., 2013) and customers (Malthouse et al., 2013), develop metrics, and assess the marketing potential in the social media context. However, there is a need for a deeper understanding of the implications of customer engagement through social media, and its utilization on the impact on company’s performance and customer satisfaction (Hennig-Thurau et al., 2013), based on empirical evidence. Without knowing how and how much social media influences customer relationships, companies struggle to assess the return on their investments in terms of revenue and goodwill.

Seller Ratings and Reviews on E-Commerce Platforms

The innovation in technology and the proliferation of e-commerce companies in recent past has revolutionized and transformed the retail industry worldwide, and are growing globally at an annual rate of 23%. However, the real success of these platforms, like Amazon, lies in the performance of the sellers who create great value for these platforms and the stakeholders by customer engagement, strategic initiatives, and operational excellence (Mu & Zhang, 2021).

In India, Amazon alone has done sales over 2.1 billion dollars in the financial year 2021 showing a growth rate of 49% from the previous year (ETech, 2022). With the huge opportunities that lie ahead for e-commerce companies, there is also a great challenge of competition along with managing the sellers on their platforms to get the competitive edge. There has been a lot of research that suggests the benefits of positive brand image, ratings and feedback of sellers on the e-commerce platforms on sales and revenue (Deshpande & Pendem, 2022; Mu & Zhang, 2021; Rodriguez et al., 2021).

It is evident from the above that seller ratings and positive reviews are as important as the product reviews if not more. For ecommerce companies, thus, it is critical to get the best possible ratings and reviews for the sellers. E-commerce companies, like Amazon, spent a lot of resources in managing the sellers and make sure that it works in the best of their interest. The current study tries to fill the literature gap and enriches it by looking at how Indian seller ratings and reviews on Amazon are affected by the customer engagement done by these sellers on a social media platform, like Twitter.

Customer Intention of Rating and Reviews

Majority of literature focuses on the customer intention for purchase and impact of the reviews or word of mouth on online purchase decisions (Liu & Tang, 2018; Meilatinova, 2021; Wang et al., 2018). As customer ratings and reviews for sellers are critical for assessing customer engagement and experiences, it is important to understand these reviews for the improvement of services (Kumar et al., 2021). It is also crucial to understand the reasons and drivers behind the customer intention to write a review or rate the sellers, which is an area, to best of our knowledge, where there is huge gap in the literature. There can be a lot of drivers which are internal and external, which includes a positive experience by the customers, customers receiving free samples, a negative experience,
incentives including rewards points and discounts, helping and guiding others, desire to help brand improve their products, and validation from other shoppers (Team, 2022).

In the last decade, the electronic word of mouth marketing (e-WOM) has gained prominence because of the disruptive growth in online users and customers becoming the contributors as well. As e-WOM is always available in the future for the stakeholders to use and impacts the buying decisions, it is critical for the industry practitioners to identify drivers which motivate the generation of e-WOM in the form of reviews and ratings for the sellers (Jain et al., 2022; Rouibah et al., 2021). While companies focus a lot on the above mentioned reasons for getting more reviews and ratings, however, there still is a lot of scope to explore new areas in this context. The current study focuses on one particular such aspect to fill this literature gap to find the relationship and impact of customer engagement on social media platforms done by sellers, on the intention and score of reviews and ratings by the customers for these sellers on e-commerce platforms.

THEORETICAL FRAMEWORK AND HYPOTHESES DEVELOPMENT

This study uses Justice Theory as the theoretical model, which is shown to be useful in identifying antecedents to psychological processes and explaining individual’s reaction to conflict situation. In the literature, three types of justice, namely distributive, procedural, and interactional justice, are suggested as relevant to explain the psychological process of social interaction. In the context of seller’s engagement on social media with the customers, distribution justice can take the form of acknowledgement of complaints or issues and taking remedial actions (Ye and Tripathi, 2016). When customers feel being heard and perceive fairness in handling their complaints, customers tend to be satisfied, have a higher trust for the seller and the company. Thus, they tend to spread positive word of mouth about it and give a positive rating to the seller and hence it leads to more seller engagement (Luca, 2017). The literature indicates that the role of justice perception is paramount in customer satisfaction and in the intention for word of mouth and reviews for the sellers. Distributive, Procedural and Interactional justice have positive impact on customer satisfaction and thus increases the likelihood of word of mouth (Fu et al., 2015).

In the past, researches have typically emphasized distributive justice in order to predict customer satisfaction (Lapidus & Pinkerton, 1995; Martínez-Tur et al., 2006). In purchase transactions, customers invest inputs, such as money, and obtain outcomes such as service quality. The degree to which the customer feels the transaction was unfair depends on how much it affected him/her after the transaction. The customer expects reciprocity in terms of tangible matters, such as relating quality to price. With respect to the distribution of outcomes, individuals seek to maximize gains and minimize losses. Thus, distributive justice reflects a more outcome-oriented and instrumental evaluation, given that a satisfactory outcome to input ratio is desired. From a service point of view and as a seller on ecommerce platform, distributive justice refers to the satisfaction of a customer and their experience in terms of the fairness with which they have been treated and provided the service (Mccoll-Kennedy & Sparks, 2003).

The current study extends this reasoning in cross platform ecosystem based on the justice theory, where the sellers who engage and address their customers’ concerns on social media platform, get better ratings and more reviews on ecommerce platforms. Further, it explores the effect on the sellers’ rating and degree of consensus on Amazon if the seller is active on social media platform like Twitter, with a particular way of communicating and handling customers. It expands the horizon of customer engagement by identifying the impact on seller disengagement or seller churn. Based on the literature review, and to fulfill the objectives of this study, the following conceptual model, shown in Figure 1, is proposed.

Reviews, Ratings, and Justice Theory

In the context of customer engagement and issue resolution, according to the available literature, justice is perceived as the evaluation done by the customers of fairness of customer complaint handling process
and it is divided into four categories: distributive justice, informational justice, interactional justice and procedural justice (Badawi et al., 2021). There has been a lot of research and available literature in this area of justice in complaints or issue handling (Ali et al., 2022; Karatepe, 2006; Muralidharan et al., 2019). From a service point of view and as a seller on ecommerce platform, distributive justice refers to the satisfaction of a customer and their experience in terms of the fairness with which they have been treated and provided the service (Mccoll-Kennedy & Sparks, 2003).

The literature indicates that the role of justice perception is paramount in customer satisfaction and in the intention for word of mouth and reviews for the sellers. Distributive, Procedural and Interactional justice have positive impact on customer satisfaction and thus increases the likelihood of word of mouth (Fu et al., 2015). In the context of ecommerce and sellers, a satisfied consumer is more likely to respond to the seller in a positive way. It has been empirically shown in the past researches that customer satisfaction has a significant impact on word of mouth in online business context (Manyanga et al., 2022; Nam et al., 2020). The current study extends this reasoning in cross platform ecosystem based on the justice theory, where the sellers who engage and address their customers’ concerns on social media platform get better ratings and more reviews on ecommerce platforms.

**Degree of Consensus in Reviews**

Consensus in the online reviews has received some attention from the market research, primarily for the objectives of identifying fake reviews and evaluating the genuineness of overall ratings. Jiménez & Mendoza (2013) defined the consensus in online reviews as the perceived agreement among the reviewers in their evaluation of the services of the sellers, that also includes the average ratings. Benedicktus et al. (2010) evaluated the impact of consensus through the rating score on overall purchase intention of the buyers. If a particular review, whether positive or negative, is in line with the other reviews for the seller, it is perceived as accurate and genuine review which leads to changing the perception and decision of the reader. Consensus in the reviews has been shown as a factor of trustworthiness (Pavlou & Dimoka, 2006). If the consensus is high in reviews for a seller, it indicates the genuineness of the reviews, which is a necessary pre-requisite for the present study (Munzel, 2016).

The hypotheses related to degree of consensus propose that there is a high degree of consensus in the reviews for a particular seller on Amazon.com, and the degree of consensus in positive reviews is directly proportional to the overall ratings of that seller.
H1a. There is a high degree of consensus in the reviews for the sellers on Amazon.com.

The hypothesis H1b looks at the consensus in the positive reviews in order to look at the positive ratings.

H1b. Degree of consensus in positive reviews is directly proportional to the overall ratings of that seller.

**Customer Orientation, Content Type, and Seller Ratings**

Customer orientation can be defined as a seller’s ability to identify, acknowledge, and address the needs of its target customers along with creating value for them (Mukerjee & Shaikh, 2018). There has been a lot of work done for finding the impact of customer orientation through digital marketing on buying intention, word of mouth, brand image, and perceiving high value of the product or the service. Customer focus and customer orientation by the companies lead to positive word of mouth and ratings (Izogo & Mpinganjira, 2020; Palazón et al., 2022). At the same time, researchers have found out that the type of the content and timing of posting are equally important for user engagement and influencing buyers’ behavior (Grover & Kar, 2018).

This study extends this phenomenon by exploring and identifying the impact of customer orientation and type of content on the overall customer rating across platforms. The tweets and conversation has been divided into marketing tweets and user centric communication tweets using machine learning, referred to as content type. Customer orientation on Twitter is captured by the ratio of user centric tweets to marketing tweets. For a company to be customer oriented, they must have more user centric tweets than marketing tweets. The hypotheses test the impact of the different types of content on the overall ratings for sellers.

H2a. There is no statistical difference between the number of marketing tweets vs. customer focused communication tweets posted by Amazon sellers on Twitter.

H2b. There is a positive impact of customer focused communication tweets on the overall ratings of the sellers.

**Seller Disengagement and Churn**

Seller disengagement and churn is a major focus area for huge ecommerce platforms like Amazon and even small ecommerce companies, because even 1% churn in sellers leads to 2% to 3% revenue losses for these platforms (Tandon, 2017). At the same time, sellers must compete with other sellers as tens of sellers are selling the same product on such platforms. Also, sellers have a big challenge managing the fast-changing protocols, policies, and algorithms by these ecommerce platforms. For example, many sellers were offering huge discounts or even free products in exchange of reviews from customers, which was banned by Amazon in the year 2016. A couple of years ago, Amazon stopped the option of personal identification for the sellers which created a huge challenge for sellers as the sellers were heavily dependent on this for their marketing campaigns (Rodriguez et al., 2021).

Even though seller churn is a critical area for ecommerce ecosystem, however, most of the academic research is focused on customer churn, and it is hard to find much work around sellers’ churn. The current study is a steppingstone in this area by exploring a small part of seller churn by seeing if there is any relationship between sellers churning from Amazon in India and their customer disengagement on social media platforms like Twitter. The hypothesis is that sellers who are churning on Amazon are disengaged with their customers on Twitter.

H3. There is a strong positive correlation between seller disengagement on social media and their churn on ecommerce platforms.
RESEARCH METHODOLOGY

The current study utilizes the textual data extracted from user generated content of two platforms, i.e. Amazon and Twitter, to achieve the objectives. Thus, a combination of social media analytics, natural language processing, machine learning and statistical testing has been used as the research methodology for this study for theory building (Eisenstein, 2019; Emmert-Streib & Dehmer, 2019; Kar and Dwivedi, 2020; Otter et al., 2021).

Out of the four types of social media analytics, this study uses descriptive and content analysis to help in understanding basic descriptive statistics about the textual data elements and to convert the textual qualitative data into more reliable and valid quantitative data (Krippendorff, 2004). Natural Language Processing (NLP) is the use of machine learning algorithms for computers to understand and use human languages for various applications. In the present study, various techniques including term frequency–inverse document frequency, more commonly known as, TF-IDF for NLP have been used to make sense of the textual data in the customer reviews and seller-customer communication (Chowdhary, 2020; Trstenjak et al. 2014).

Machine learning has been used for two primary tasks; first, a supervised classification technique has been used on labelled data for text categorization, and second, for sentiment, polarity, and consensus analysis (Aphinyanaphongs et al. 2014). Text cleaning, preparation and quantification of data was done before implementing the machine learning techniques. Partial Least Squares (PLS) path analysis has been used to find out the relationship between latent variables, and the impact on seller churn because of ratings on Amazon and the variables retrieved from Twitter, like number of tweets, tweet category, polarity etc. (Henseler et al., 2016; Sohaib, 2021). For statistical tests, multiple linear regression model has been used on the tabular data which is combined based on the sellers, between Amazon and Twitter (Dospinescu et al., 2021). Finally, for seller churn, longitudinal study was done to identify those sellers who have stopped doing business at Amazon India. Figure 2 provides the overall, high level methodology for the current study. This section is divided into three subsections. The first part explains the data collection process, the second part provides the details with data processing and implementation, and the last subsection is about the way data has been analysed for the study.

Data Collection

The data for this study has been collected through the following four steps:

Step 1: Randomly search sellers on Amazon India with at-least 50, pre-covid period, reviews. This was done to ensure there is no bias of the pandemic on the analysis, and second there should be sufficient transactions for a seller to be part of this study. From this list of sellers, only those sellers were considered who manage an active Twitter page with at least 100 Tweets. Only those sellers who fulfil these two conditions were chosen for this study. It is a very time consuming and critical task to identify the same seller on Twitter based on their presence on Amazon. 260 such sellers were identified and selected for the study from over 1000 initial sellers which were identified from Amazon. These sellers are from multiple locations across India covering presence across nation, and thus the study is not limited to a particular geography in the country. Similarly, the category of products is not limited to a specific subset of product types.

Step 2: For these 260 sellers, all the reviews given by the customers along with the ratings were captured for pre-covid pandemic time period. The data was stored in Excel files to be later read into Python environment for processing. A total of 12950 reviews were collected.

Step 3: Using Tweetter API and Tweepy in Python, the tweets were extracted from Twitter for these 260 sellers. A total of 256887 tweets were collected which included marketing and interaction tweets.
Example of marketing tweet: “1083306762227204096|2019-01-10 10:17:13|b'New Styles Added and Huge Discounts Dropped; what are you waiting for?

https://t.co/pawf3X44uW https://t.co/QivY4Wpct1”

Example of interaction tweet: “@5TXLLc8WdjcbMvN sorry about this. will ask the customer care to resolve.”

Step 4: After two years, the status of these 260 sellers was relooked at. Those sellers, who haven’t got any reviews on Amazon for more than recent six months were tagged as inactive and churned. 44 out of 260 sellers were tagged as churned.

Data Pre-Processing and Data Preparation

The textual data collected, i.e. the tweets collected from Twitter and Seller reviews from Amazon, suffers from numerous syntactical and semantic issues like spelling mistakes, punctuation, incorrect grammar, usage of slang, abbreviations, and emojis. The objective is to remove this noise from the data to ensure unbiased and accurate outcome. A series of steps have been followed during the pre-processing of the texts. The URLs and hashtags were deleted from the tweets and comments during the first data cleaning step. In the second step, all emoji were removed from the data, and all text was converted in lowercase. In the third step, slang from the tweets and messages were removed through data pre-processing. Lemmatization process has been used to convert multiple words with the same root or meaning into single words to different inflected forms of an expression. The next stage was to reveal and utilize the emotions present in the text. The feelings in the text were transformed; for example, ‘:)’ is replaced with ‘Happy’ and ‘:(' is replaced with ‘sad.’ Further, stop words were eradicated from texts as these word do not contribute in the analysis. All this was done using the Spacy library in Python along with regular expressions and inbuilt Python features.

Data Analysis

Various types of analysis has been used on the two sets of textual data such as content analysis, and social media analytics. Content analysis has been used to categorize the total number of 256887 tweets into communication and advertisement/marketing tweets using machine learning models. First, around 8000 tweets were manually labelled and then using machine learning, multiple models were trained for splitting the tweets into two categories. 25% of the tweets were kept for testing using k-fold validation. After trying multiple machine learning models, Logistic Regression resulted in the best accuracy of 99.20%. The tweets were converted into a numerical matrix using TF-IDF, and then a
model using Logistic Regression on 135 features was created. Using the trained model, the remaining
tweets were divided into two buckets, i.e. marketing and communication tweets.

Amazon reviews were clustered into three buckets to check the consensus on the reviews. Three
clusters were formed based on the ratings scores; first for rating 1 and 2, second for rating 3 and third
for rating 4 and 5. The Silhouette coefficient is calculated to analyse the consensus amongst reviews
of different ratings (Tang et al., 2020). The high value of coefficient indicates that the clusters are
well formed and the degree of consensus is high. In the next cycle, all the common words and stop
words were removed to ensure there is no bias due to the commonly used words in the reviews. The
Silhouette coefficient is still above the recommended threshold and confirmed high consensus in
the reviews for the sellers.

Finally, Partial Least Squares (PLS) path analysis was used to find out the relationship between
latent variables, and the impact on seller churn because of ratings on Amazon and the variables
retrieved from Twitter, like number of tweets, tweet category, polarity etc. It has helped as confirmatory
statistical analysis using iterative ordinary least squares regression for PLS. Linear regression was
also used for conforming the impact of customer engagement by sellers on overall rating scores. For
the seller churn with longitudinal data, Mann Whitney U test was conducted to find the statistically
significant dissimilarity between the two groups of sellers, i.e. sellers who are active vs. the sellers
who have churned.

FINdINGS

This section has been divided into three subsections. The first subsection covers the analysis on
reviews and ratings to find the consensus between reviewers’ feedback. Next subsection presents the
impact of customer engagement through Twitter on the seller ratings on Amazon followed by the last
subsection which covers the seller disengagement and churn.

Degree of Consensus in Reviews

Literature indicates a direct impact of consensus in the reviews of customers on the purchase intention.
Customers are more likely to buy from a seller who has received feedback with a consensus in positive
reviews, as it increases the trustworthiness for the seller (Benedicktus et al., 2010). There is a strong
positive correlation between agreement in the positive reviews for a seller and the purchase intention
from the seller in ecommerce (Jiménez & Mendoza, 2013). Hypothesis H1a proposes that the reviews
on Amazon are genuine and have a high consensus.

H1a. There is a high degree of consensus in the reviews for every seller on Amazon.com.

After pre-processing step, i.e. applying regular expression, cleaning noise and stop words, the
clean textual data is available for every sellers’ reviews on Amazon along with the rating given against
each review. After applying TF-IDF on this data, k-mean clustering was applied with the rating score
as one of the variables. Three clusters were formed; one for rating 1 and 2, second for rating 3, and
third for rating 4 and 5. To find out whether the words used are same and there is a consensus, the
Silhouette coefficient was calculated for each cluster. With the score of 0.82, it is evident that the
clusters are very well representing similarity in the words (Shahapure & Nicholas, 2020). Further,
more stop words and commonly used words were identified, e.g. those words which were showing
emotions, to eliminate any possible bias towards high Silhouette coefficient, and the process was
repeated to calculate Silhouette coefficient on the new clusters. It still gives a silhouette coefficient
of 0.74. It suggests that there is a high consensus in the reviews (Larose & Larose, 2015). Thus, this
data can be used for further processing and analysis.
As an example, here are three reviews from different customers which are in one cluster, showing high consensus:

1. “Packaging was safe n secure.”
2. “Nice packing”
3. “Nice Product and Packaging.”

It would be interesting to see the relationship and impact of consensus on overall rating of the seller. Hypothesis H1b proposes this relationship as follows:

**H1b.** Degree of consensus in positive reviews is directly proportional to the overall ratings of that seller.

For a seller, the consensus score for positive reviews was calculated based on the Silhouette coefficient for positive review clusters, i.e. for ratings 4 and 5 and compared against the overall rating score of the seller. For the 260 sellers, the correlation coefficient is 0.84, which indicates a very strong correlation between consensus in positive reviews and overall ratings of the seller. Thus, based on the previous researches and literature as mentioned previously, it can be concluded that higher consensus in positive reviews has higher overall ratings that eventually leads to positive buying intentions.

### Customer Orientation, Content Type, and Seller Ratings

Literature suggests that customer orientation has a positive impact on the buying intentions and is a critical attribute in e-commerce context (Hamidi & Moradi, 2017). With the help of dataset collected from Twitter for these 260 sellers, this particular analysis has been started with the following hypothesis to see if sellers focus more on marketing tweets on social media or they use it for customer orientation and engagement.

**H2a.** There is no statistical difference between the number of marketing tweets vs. customer focused communication tweets posted by Amazon sellers on Twitter.

At $\alpha = 0.05$ (assumption), $n$ number of observations at 260, and $k$ = number of sample at 2, degree of freedom is $(1, 259)$. According to the decision rule, if calculated value of $F$ is greater than the Statistical table value of $F$, the null hypothesis can be rejected. With the current sample, the $F$ value at 5% level of significance is 6.28, thus the null hypothesis is rejected, and there is a significant difference between marketing tweets and customer oriented tweets. With a higher number of marketing tweets in the sample, it can be concluded that sellers use social media platforms, like Twitter, for marketing significantly more than customer orientation. Figure 3, Word clouds show the common words used in interaction vs. marketing tweets.

**H2b.** There is a positive impact of customer focused communication tweets on the overall ratings of the sellers.

This is one of the most valued analysis of the current work. A linear regression was calculated to find the impact of customer focused communication tweets, representing customer orientation, on the ratings of the sellers. The combined data from Amazon and Twitter based on the sellers, with average rating as the dependent variable, was used for the regression model. A significant regression equation was found with these parameters: $F = 57.87$, $p < 0.000$, and $R^2 = 0.72$. Higher the number of customer focused tweets, better is the rating score of the sellers.

Here is the functional relationship of the model used:
Seller Rating Score on Amazon = \( f \) (Count of marketing tweets on Twitter, Count of communication tweets on Twitter, Subjectivity of the Tweets, Polarity of the Tweets, Count of Reviews, Business category of the sellers, Years of operations of sellers on Amazon, Place of business), where the last three dependent variables have been used as control variables.

**Seller Disengagement and Churn**

Seller churn has a major loss for e-commerce giants like Amazon. While there can be several reasons for seller churn, however, it would be interesting to see the impact and relationship between lack of customer engagement and being inactive on social media, on seller churn on e-commerce platform.

H3. There is a strong positive correlation between seller disengagement on social media and their churn on ecommerce platforms.

The null hypothesis is that there is no significant relationship between the number of customer-oriented tweets by a seller and the seller’s churn on e-commerce platform. For the 44 sellers who churned, against the 216 sellers who continue to do well in terms of sales and reviews, Mann Whitney U test is used to find if there is any statistically significant difference between the number of customer-oriented tweets. The test is used because the variance is not same between the two groups and it is not normally distributed, thus Mann Whitney U test is the best choice (Ghobadi et al., 2015; Zimmerman, 1987). At \( \alpha = 0.05 \) (assumption), the statistics is equal to 112, with \( p < .000 \), thus the null hypothesis is rejected. The average number of customer-oriented tweets by the churned seller during their active phase is 119, whereas the average number of customer-oriented tweets by the other sellers in the same period is 378, thus, it is evident that less customer orientation on social media indicates seller churn on ecommerce platforms.

Table 1 below summarizes the bootstrap algorithm result for path significance analysis. With the coefficient scores and t statistics, it is evident that the inferential model supports the contextual model introduced earlier through the paths and hypothesis testing discussed above (Henseler et al., 2016).
Sellers and suppliers are an integral part of the ecommerce ecosystem, and the disengagement or churn makes a huge negative impact on the profitability and competitive advantage for ecommerce companies (Qin et al., 2021). There are several areas in which the ecommerce companies work to ensure a healthy environment for the sellers. Customer engagement and orientation is an important need in every business, however, it is even more important, essential and critical in ecommerce, as the buying intentions are heavily driven by it.

The only factors that are available to a buyer are the way these sellers engage with them, and the reviews given to these sellers by other buyers. While both of these factors are independently useful and have been studied in detail, the present study has looked at the two factors together to see the impact of one on another across different online platforms. The study explores the unexplored area of the cross-platform relationship between the customer engagement by the sellers on social media and performance of these sellers on ecommerce. The study also paves a path for exploring the area of seller churn on ecommerce platforms further.

To enable the study, customer reviews and ratings from ecommerce website, Amazon were taken for 260 sellers who had an active presence on social media (Twitter) as well. The tweets and posts by these sellers on Twitter were also extracted. After using natural language processing and machine learning, the action and focus of these sellers on social media were analysed for validating the statistical models.

Hypothesis H1a and H1b are important to ensure the data is valid and useful for testing the most important findings of this study. H1a and H1b reveal that the data is trustworthy because of high consensus and the higher consensus in positive ratings also has a positive impact on higher ratings for the sellers. Sellers and ecommerce platforms must continue their focus on the consensus in reviews and must drive these based on their strategic and operational excellence.

H2a and H2b focused on the relationship between sellers’ behaviour on social media and their ratings score and reviews on ecommerce platform. The tests suggest that the sellers are more focused on marketing in general rather than engaging the customers and proving support to them. However, it is clearly shown by the data, that more customer engagement and customer-oriented behaviour the seller demonstrates, better reviews and ratings they receive. As this is a lot easier to manage in comparison to exceeding the expectation of customers in their service, the sellers must focus more on taking care of their customers on social media. Ecommerce companies must drive this behaviour and can offer schemes for these sellers if not already doing it to promote this behaviour.

This is the first study which has tried to explore this cross platform impact of social media engagement on seller disengagement and churn. Hypothesis H3, which is the highlight of the present study and uses longitudinal data to find the impact of involvement and activities by the sellers through social media engagement on the disengagement or churn on ecommerce. As mentioned above, seller churn significantly impacts the profitability of ecommerce companies in a negative way. Ecommerce

<table>
<thead>
<tr>
<th>Path</th>
<th>Path coefficient (original sample)</th>
<th>Mean of path coefficients (bootstrap samples)</th>
<th>Standard Error</th>
<th>t-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer orientation -&gt; Subjectivity</td>
<td>0.529</td>
<td>0.531</td>
<td>0.081</td>
<td>5.696</td>
<td>0.000</td>
</tr>
<tr>
<td>Customer orientation -&gt; Polarity</td>
<td>0.488</td>
<td>0.489</td>
<td>0.111</td>
<td>5.115</td>
<td>0.000</td>
</tr>
<tr>
<td>Subjectivity -&gt; Seller Ratings</td>
<td>0.315</td>
<td>0.312</td>
<td>0.016</td>
<td>2.963</td>
<td>0.003</td>
</tr>
<tr>
<td>Polarity -&gt; Seller Ratings</td>
<td>0.331</td>
<td>0.353</td>
<td>0.091</td>
<td>3.988</td>
<td>0.002</td>
</tr>
<tr>
<td>Seller Disengagement -&gt; Seller Ratings</td>
<td>0.281</td>
<td>0.285</td>
<td>0.021</td>
<td>2.234</td>
<td>0.025</td>
</tr>
<tr>
<td>Seller Disengagement -&gt; Seller Churn</td>
<td>0.642</td>
<td>0.649</td>
<td>0.106</td>
<td>6.642</td>
<td>0.000</td>
</tr>
<tr>
<td>Seller Ratings -&gt; Seller Churn</td>
<td>0.411</td>
<td>0.402</td>
<td>0.053</td>
<td>3.043</td>
<td>0.002</td>
</tr>
</tbody>
</table>

DISCUSSION

Sellers and suppliers are an integral part of the ecommerce ecosystem, and the disengagement or churn makes a huge negative impact on the profitability and competitive advantage for ecommerce companies (Qin et al., 2021). There are several areas in which the ecommerce companies work to ensure a healthy environment for the sellers. Customer engagement and orientation is an important need in every business, however, it is even more important, essential and critical in ecommerce, as the buying intentions are heavily driven by it.

The only factors that are available to a buyer are the way these sellers engage with them, and the reviews given to these sellers by other buyers. While both of these factors are independently useful and have been studied in detail, the present study has looked at the two factors together to see the impact of one on another across different online platforms. The study explores the unexplored area of the cross-platform relationship between the customer engagement by the sellers on social media and performance of these sellers on ecommerce. The study also paves a path for exploring the area of seller churn on ecommerce platforms further.

To enable the study, customer reviews and ratings from ecommerce website, Amazon were taken for 260 sellers who had an active presence on social media (Twitter) as well. The tweets and posts by these sellers on Twitter were also extracted. After using natural language processing and machine learning, the action and focus of these sellers on social media were analysed for validating the statistical models.

Hypothesis H1a and H1b are important to ensure the data is valid and useful for testing the most important findings of this study. H1a and H1b reveal that the data is trustworthy because of high consensus and the higher consensus in positive ratings also has a positive impact on higher ratings for the sellers. Sellers and ecommerce platforms must continue their focus on the consensus in reviews and must drive these based on their strategic and operational excellence.

H2a and H2b focused on the relationship between sellers’ behaviour on social media and their ratings score and reviews on ecommerce platform. The tests suggest that the sellers are more focused on marketing in general rather than engaging the customers and proving support to them. However, it is clearly shown by the data, that more customer engagement and customer-oriented behaviour the seller demonstrates, better reviews and ratings they receive. As this is a lot easier to manage in comparison to exceeding the expectation of customers in their service, the sellers must focus more on taking care of their customers on social media. Ecommerce companies must drive this behaviour and can offer schemes for these sellers if not already doing it to promote this behaviour.

This is the first study which has tried to explore this cross platform impact of social media engagement on seller disengagement and churn. Hypothesis H3, which is the highlight of the present study and uses longitudinal data to find the impact of involvement and activities by the sellers through social media engagement on the disengagement or churn on ecommerce. As mentioned above, seller churn significantly impacts the profitability of ecommerce companies in a negative way. Ecommerce
companies can monitor the sellers’ social media behaviour, and take appropriate actions to avoid the churn. The result suggests that lesser customer focus and involvement on social media by a seller indicates the seller churn in near future.

**Theoretical Contributions**

The present study tries to address various gaps in this area and in this process, it makes important theoretical contributions to both customer engagement and customer experience literature (Kumar et al., 2022). This study has two primary theoretical contributions. First, the study extends the limited research on the understanding of the categories of communication done by sellers on social media and its impact on the reviews and ratings these sellers receive on ecommerce platforms. This study is one of the first studies to consider social media performance of a seller as an important antecedent of the ratings for the sellers on ecommerce platform.

Second, the study looks at the critical factor of seller churn on ecommerce platform by analyzing the impact of the sellers’ social media engagement and disengagement. As this is a novel way of looking at seller churn, the study will open a completely new area of research and will contribute significantly to the literature in the future. No previous studies to the best of authors’ knowledge and thorough search have empirically explored the relationship between sellers’ social media performance and seller churn on ecommerce in an academic setting. However, literature and research show the impact of sellers’ customer-orientation on the buying intentions, the positive impact of consensus on the trustworthiness of the reviews, and the negative impact of seller churn on the profitability of ecommerce companies (Chakravarty et al., 2014; Chang et al., 2003; Clemons, 2007; Cuong, 2020; Goad & Jaramillo, 2014; Liu et al., 2020).

The theoretical lens for this study is Distribution Justice part of the Justice Theory that supports the explanation of proposed relationship between social media engagement by the seller and the factors like ratings, reviews and seller churn on ecommerce. The study would add to the theoretical development by integrating cross platform relationship between the variables chosen on ecommerce and social media platform to foster the seller engagement.

**Implications for Practice**

Literature suggests that users have four primary needs to use social media, viz. cognitive, social, personal, and effective need. This study proposes better presence and more focus by the sellers on customer engagement and communication posts on social media to fulfil the needs of the existing and prospective customers. Hypothesis H2a, H2b and H3 indicates the benefit of the sellers’ appropriate usage and engagement on Twitter for both sellers and ecommerce platforms. Thus, this study suggests all the ecommerce sellers and ecommerce companies to consistently track, motivate and invest in social media engagement with customer orientation for better sales, positive word of mouth, positive reviews, and better ratings for the sellers. It also helps ecommerce companies to retain the sellers and will reduce seller churn.

The companies, both sellers and ecommerce companies, can formulate strategies to look at the timing, frequency, and impact of seasonality in the social media engagement. The study suggests that both type of these firms must focus on (a) using social media to fulfil the needs of the customers, (b) have high focus on customer orientation while posting on social media, and (c) track and formulate strategies to have the maximum impact of social media engagement on the revenue, ratings, and seller retention. As this study is specifically focused on ecommerce platforms like Amazon and the sellers on these platforms, anyone from both these types of companies can use the findings of the study for proactive strategies and actions to reduce seller churn, and to have effective customer engagement for better ratings and profits.
CONCLUSION

This study tries to explore the cross-platform impact of customer-oriented seller engagement through social media on the reviews and ratings for the sellers on ecommerce platform, along with the indication of seller churn. The focus of sellers on social media has been analyzed by looking at the content type. The impact has been analyzed from both user perspective and sellers’ perspective. This study introduces the conceptual model and uses machine learning, natural language processing and statistics to validate the model and assumptions. It has been revealed that the firms, i.e., sellers and ecommerce platforms, must gauge, focus, and encourage customer-oriented posts on social media and continuously fulfil the needs of the users by addressing their concerns and communicating with them. While social media platforms are very helpful marketing channels, however, the sellers must increase the communication and interaction posts, and must try to have it more than the marketing messages. This is the first study to explore this cross-platform relationship and unexplored impact of social media behavior by sellers on their ratings and reviews on ecommerce platform. For ecommerce companies, the sellers’ social media behavior can be used as a key performance indicator to evaluate the performance, and ecommerce companies should set thresholds and targets for the sellers to have a minimum number of customer-oriented tweets or posts on social media platforms. Same methodology, which is used in this study, can be used, and extended for further research and studies in this area.

Limitations and Future Work

Although this study makes a substantial contribution to the understanding of seller engagement, disengagement, and churn because of sellers’ use of social media, the utilization of exclusively seller-generated content severely restricts the reach of this study. Future studies should examine and assess user-generated content to understand its function and significance. There is a possibility that some sellers were not identified on Twitter out of 1000 sellers from Amazon as it was searched manually on Twitter through the name of the sellers despite doing multiple rounds of search. Also, a deeper contextual analysis can be done by bringing in the seller type in the study. Since Twitter was the source of the data at a certain period for this study, it can be extended to include data from other social media sites by quantifying the relationship between sellers’ roles on those sites and seller ratings. This study relies only on social media data and the customer reviews. In future, customer surveys can be administered to build and validate theoretical models.

ACKNOWLEDGMENT

We would like to thank Quanolytics LLP, India for supporting this research.

COMPETING INTERESTS

All authors of this article declare there are no competing interest.

FUNDING AGENCY

This research and publication were supported by Quanolytics LLP, India.
REFERENCES


**Anuj** is the founder and CEO of Quanolytics, which is a fast growing data science consulting company, headquartered in India. Before Quanolytics, he has worked for companies like Adobe, Sapient, HCL technologies etc. in leadership roles in data science and strategy. He also teaches at premier universities and institutions as a visiting and part time professor in India, Europe and USA. Anuj has double post-graduation and is an IIM-Calcutta alumnus. He’s doing his research at IIT Delhi.

**Arpan Kar** is Chair Professor in Data Science and Decision Science in Indian Institute of Technology Delhi, India. Within IIT Delhi, he shares a joint appointment in the Department of Management Studies and Yardi School of Artificial Intelligence. Administratively he is the Chair of Corporate Relations and is a member of Board of Academic Programme and Institute of Eminence Committee, among others. He also supports multiple ministries and other universities also in different advisory boards/roles. He works in the space of data science, digital platforms and apps, social media and ICT-based public policy. He has authored over 175 peer reviewed articles and edited 10 books and has over 8000 citations with an H-Index of 47. Over 60 publications are in ABDC A, ABS 3 or WoS Q1 journals including 12 publications in ABDC A* journals. He is the Editor in Chief of International Journal of Information Management Data Insights, published by Elsevier, which is ranked as a Q1 journal in Scopus. He has been a Guest Editor for ABDC A*/A journals like Decision Support Systems, Industrial Marketing Management, International Journal of Information Management, Information Systems Frontiers, Australasian Journal of Information Systems, etc. He has received over 20 prestigious awards for his research.

**Satpathy** has over 16 years of mixed experience in corporate and academics. Pursuing her passion in teaching, she is competent to bring together wide array of information into the classroom that is effective at driving performance. She holds a PhD in Management and Masters in Business Administration. Se has presented papers in national and international conferences and published research in various journals of high repute.