

Differences of Consumer Reviews and Consumer Satisfaction on Cross-Border E-Commerce Platforms: A Text Mining Analysis Based on Cross-Cultural Perspective

Ruo-Xin Zhou ¹
Yu-Ting Zhu ^{1,2}
Run-Qi Zheng ³
Ji-Lei Zhou ^{4,*}

*Corresponding author

¹ School of Information Technology & Management, University of International Business and Economics, Beijing, China, rxzhou@uibe.edu.cn

² China Unicom (Xiamen Branch), Xiamen, Fujian, China, zhuyt63@chinaunicom.cn

³ Krieger School of Art & Sciences, Johns Hopkins University, Baltimore, MD U.S.A, rzheng17@jh.edu

⁴ School of Information, Renmin University of China, Beijing, China, zhoujilei@ruc.edu.cn

ABSTRACT

Cross-border e-commerce (CBEC) has accelerated the globalization of enterprises and enabled them to reach consumers worldwide. To meet the needs of consumers in different markets, CBEC platforms and enterprises must adopt differentiated strategies across regional markets by analyzing consumer behavior patterns and potential demand. Based on the cross-cultural theory, this study investigates consumer preferences in CBEC. A text mining analysis of 513,338 consumer reviews on the AliExpress platform was used to explore the cultural differences in consumer reviews and satisfaction. The U.S. (i.e., individualism) and Russia (i.e., collectivism) were selected for the cross-cultural study. The study found that individualistic consumers focused more on services and logistics while they paid less attention to products. In terms of satisfaction, U.S. consumers are more satisfied with logistics, while Russian consumers are more satisfied with products and services. The findings could help CBEC platforms and enterprises satisfy diverse market demands and improve consumer satisfaction.

Keywords: Cross-border e-commerce (CBEC), online review, consumer satisfaction, cross-cultural differences, latent Dirichlet allocation (LDA).

INTRODUCTION

Cross-border e-commerce (CBEC) has emerged as a pattern of international trade and commerce, integrating domestic and international markets. CBEC can provide convenience and opportunities for small and medium-sized enterprises (SMEs) to optimize and integrate their supply and distribution channels, positively impacting the transformation and modernization of the international trade industry (Li, 2021). However, opportunities are always accompanied by challenges. Compared with domestic e-commerce, export markets for CBEC are more heterogeneous. The “distance” (e.g., cultural distance, administrative or political distance, geographic distance, and economic distance) between different target markets and the country where an enterprise is located can dramatically change the enterprise’s picture of its strategic options (Ghemawat, 2001). In this regard, CBEC enterprises have to employ differentiation strategies to improve their performance (Knight, Madsen, & Servais, 2004). The prerequisite of adopting differentiated strategies is to analyze the differences in consumer preferences between countries and regions and to comprehend the factors influencing consumer choices and satisfaction. It is a challenging task since information asymmetry is prominent in CBEC.

To accurately understand different markets, CBEC platform and enterprises rely on online reviews to get consumer feedback. By analyzing consumer reviews, enterprises could uncover consumer demands, take control of the transaction process, optimize products and services (Clemons & Gao, 2008; Rose, Clark, Samouel, & Hair, 2012), improve consumer WOM and corporate image (Mou, Ren, Qin, & Kurcz, 2019), and thus enhance consumer brand loyalty (Maurer & Schaich, 2011). Instead of relying exclusively on numerical data like ratings, consumers are more inclined to comprehend and trust the textual content of online reviews. Unlike structured online ratings, textual reviews provide a more detailed and comprehensive portrait of a consumer’s experience and perception. Therefore, the textual content of online reviews is valuable in identifying consumer preferences (Ludwig *et al.*, 2013; Zhu, Ye, & Chang, 2017).

However, in the CBEC literature, little research leverages the textual content of online reviews to examine consumer heterogeneity and its cultural influencing factors. Using text analytics, this study reveals the similarities and differences in the preferences of online consumers across countries, thereby filling the gap. Specifically, we will address the following research questions:

RQ1: Are there differences in the textual topics of online reviews between cross-cultural countries?

RQ2: Do the impacts of these review topics on consumer satisfaction vary across countries?

To answer these questions, this study employed the Latent Dirichlet Allocation (LDA) document topic modeling of online reviews and compared the differences in the textual topics of online reviews between the U.S. and Russian consumers. Then, to examine the impact of review topics on consumer satisfaction in both countries, an empirical regression model evaluating the effect of topic probabilities on consumer satisfaction is utilized. Results show that, first, collectivists are more concerned with services and logistics, while individualists focus more on products. Second, consumers in the U.S. and Russia are pleased with the products, services, and logistics on the AliExpress platform. Comparatively, U.S. consumers are more satisfied with logistics, whereas Russian consumers are more satisfied with products and services. Russian consumers also rate products significantly higher than U.S. consumers. These findings could help CBEC exporters to differentiate their strategies for different target markets.

The rest of this paper is organized as follows. Section 2 discusses previous studies related to online reviews in CBEC and the application of cross-cultural theory to CBEC research. Then, the proposed hypothesis of cross-cultural consumer differences is illustrated in Section 3. In Section 4, the empirical results and analysis of our experiments are reported. The last section presents the concluding remarks and the future directions.

LITERATURE REVIEW

Online Consumer Reviews in CBEC

CBEC has become an important channel to help SMEs enter the international market. And its rapid growth has simultaneously led to the accumulation of valuable textual data, such as online consumer reviews (Mou et al., 2019). As a result, CBEC research on online consumer reviews has received an increasing amount of attention. From the perspective of an enterprise's entry into international markets, researchers have examined online reviews for consumer feedback on the internationalization process of enterprises. Besides recommending improvements for products and services, they also establish evaluation criteria for CBEC. For example, the most popular topics in CBEC enterprise and consumer reviews were examined by Mou et al.(2019) using the LDA model. It enables enterprises to utilize consumer feedback to improve their products and services.

Despite the achievements of previous text mining studies, the underlying factors resulting in these consumer differences have received little attention. To demonstrate the differences in consumer demand between countries and to reveal the underlying reasons, this study adopted Hofstede's cultural dimension theory to explore and compare the factors influencing consumer satisfaction and consumer preferences in different overseas markets from a cross-cultural perspective.

Cross-Cultural Study of Online Reviews for CBEC

Cultural factors are one of the barriers hindering the cross-border business process (Tian & Lan, 2009). And cultural backgrounds are the main cause of the differences in target market consumer preferences (de Mooij & Hofstede, 2002). Therefore, the development of CBEC requires an understanding of the cultural background of the target market and the implementation of corresponding strategies.

Based on Hofstede's cultural dimension theory, prior studies have analyzed cross-cultural differences in online reviews in terms of consumer ratings and review text (Jia, 2020; Tsiotsou, 2022) as well as review intentions (Li, Shen, & Destech Publicat, 2018). Tsiotsou (2022) analyzed online ratings and reviews of luxury hotels posted on TripAdvisor by consumers from four European areas and examined their cross-cultural differences. Jia (2020) explored cultural differences between consumers in China and the U.S. by analyzing online ratings and text reviews of restaurants to evaluate and compare their motivations and levels of satisfaction. Li et al. (2018) investigated the factors influencing Chinese and the U.S. consumers' online review intentions, and the group analysis results revealed ethnocultural differences in consumer review intentions.

However, CBEC consumer satisfaction and its underlying cross-cultural reasons have been underexplored. To bridge this gap, this study explored the impact of review topics on consumer satisfaction across countries and further analyzed the role of cultural factors in this impact. In addition, this study examined cross-cultural issues from the perspective of CBEC export, and the findings could help develop targeted and differentiated CBEC export strategies.

HYPOTHESES THAT LINK CULTURAL DIMENSIONS TO CBEC REVIEW TOPICS

This study investigates the differences in consumer behavior across different target markets in CBEC by comparing the consumer reviews of a specific product category (i.e., women's fashion). Individualism and collectivism, one of the Hofstede's cultural dimensions, is used to explain the differences between consumer reviews posted by people with different cultural backgrounds. To compare the differences between consumer reviews, this study analyzes and discusses the factors that affect the satisfaction of CBEC platforms from the perspective of products and services.

Products are one of the fundamental elements in CBEC transactions. Researchers found that no matter search or experience products, collectivists pay greater attention to this dimension than individualists do. For example, Faqih and Jaradat (2015) found that collectivists are more concerned with the usefulness of mobile commerce technology than individualists when deciding to

adopt such technology. In a cross-cultural study of the hotel experience, Leon (2019) discovered that collectivists inspect and recall rooms and meals more precisely than individualists, both of which are hotel functionalities. Consumers' interest in the product itself will result in an increase in product-related reviews (Xiao, 2018). Thus, the following hypothesis is proposed.

H1. Individualists pay less attention to products than collectivists, and they post less product-related reviews.

As the extended value of products, service has become a crucial instrument for attracting consumers, and a differentiation service strategy should be implemented for many e-commerce enterprises, especially for the B2C e-commerce enterprise (Gao & Li, 2014). Services can be divided into seller services and logistics services. Seller services mainly include responding to consumers' questions, presenting details of relevant products, and recommending products based on consumers' descriptions, such as the right size or color. Logistics services are one of the most expensive operations in CBEC and play a key role in facilitating online purchases (Ding, Gao, Huang, Shu, & Yang, 2018; Qin, Liu, & Tian, 2020). Logistics has become an obstacle impeding the expansion of CBEC's rapid growth (Wu, 2016). In addition, unlike the services provided by the sellers themselves, logistics services are delivered by a third-party logistics service provider, of which the seller has limited control. We therefore discuss logistics separately. In terms of the effect of cultural differences, individualism has a positive moderating effect on the relationship between service quality and satisfaction (Lee, Kang, & Kang, 2019). This implies that individualists will have higher expectations of services. Further, Punel, Hassan, & Ermagun (2019) indicated that higher levels of consumer expectations for services are closely related to their increased attention to the service. Consequently, individualistic consumers are more attentive to services. And we put forward the following hypotheses.

H2. Individualists pay more attention to services than collectivists, and they post more service-related reviews.

H3. Individualists pay more attention to logistics than collectivists, and they post more logistics-related reviews.

EMPIRICAL ANALYSIS

Data

Data Collection and Pre-processing

Based on the textual online reviews retrieved from AliExpress, the largest B2C export e-commerce platform in China, we examine the differences in consumer behavior across different target markets in CBEC. Women's fashion was the designated category for data collection and analysis. All data was crawled from the AliExpress website using the REQUEST and SELENIUM libraries in Python. A total of 2,241,862 pieces of review data, each containing the textual online consumer review, the country of the review, the rating and control variables related to reviews, products, and shops, have been retrieved from aliexpress.us. The time frame for the review data is October 2, 2021 to December April 3, 2022.

Data Pre-processing

For the convenience of subsequent analysis, English was chosen as the review language in this study. The AliExpress platform provides a review translation function that could unify the reviews of consumers from different countries into a specific language. After examining the translated reviews, we found that the majority had been successfully translated, but a few were still in other languages, requiring additional translation work.

First, following the general steps of data pre-processing, empty and duplicate reviews were removed from the data, and the non-English reviews were translated using Python and Baidu translation API. After translation, the non-English characters (e.g., emoticons and Chinese punctuation marks) and untranslated sentences in the reviews were removed. Then, following the common text mining procedure, the NLTK library in Python was used to remove the numbers and punctuation marks from the reviews, and the stop words (e.g., "is", "the" and "which") were removed using the stopwords list to make the topic modeling results more accurate. To avoid the impact of too short reviews and insufficient information on topic modeling, products with fewer than 10 valid reviews and reviews with fewer than 5 valid words after processing were removed (Wang, Yang, Tso, & Li, 2019), and a total of 513,338 review data were finally obtained.

Descriptive Analysis

After filtering the above data, we calculated and normalized the non-text review data to facilitate further comparison and regression analysis. First, certain products have varying prices for their different specifications, which are displayed as price intervals. To measure the value of the products, the average of the price intervals was used as the original and current prices of such products. Second, the expression of discounts differs for different products. This study used equation (1) to calculate the value of discounts, where $Discount_i$, o_price_i and c_price_i indicate the discount, original price and current price of review i respectively.

$$Discount_i = (c_price_i - o_price_i) / o_price_i \quad (1)$$

The descriptions of review-related characteristics and the descriptive statistics are shown in Table 1 and Table 2 respectively.

Table 1: Description of variables.

Variables	Description
$country_i$	Country of the consumer who posted review i
$rating_i$	Individual consumer rating of product to which review i belongs

$shippingMethod_i$	The dummy variable for whether the delivery method of the product to which review i belongs is standard shipping. The value is 1 for AliExpress Standard Shipping, otherwise it is 0
$reviewTime_i$	The number of days between the date review i was published and the date the earliest review was published
$reviewNum_i$	Total number of reviews for the product to which review i belongs
$shelfTime_i$	Length of time for the earliest review of the product to which review i belongs
$averageRating_i$	Average of all ratings for the product to which review i belongs
$originalPrice_i$	The original purchase price of the product to which review i belongs
$discount_i$	The actual discount of the product to which review i belongs
$productid_i$	The unique number for the product to which review i belongs
$productTypeDummy_p$	Dummy variables for product types (p denotes a specific product type)
$shopFollowers_i$	Number of followers of the shop where the product to which review i belongs is located

Table 2: Variable statistics.

Variables	N	Mean	Std. Dev	Min	P50	Max
$rating_i$	513,338	4.334	1.180	1	5	5
$shippingMethod_i$	513,338	0.197	0.398	0	0	1
$reviewTime_i$	513,338	100.6	48.09	0	106	183
$reviewNum_i$	513,338	465.6	815.4	14	201	8405
$shelfTime_i$	513,338	173.0	20.19	18	180	183
$averageRating_i$	513,338	4.617	0.227	1.600	4.700	5
$originalPrice_i$	513,338	29.75	27.15	1.070	23.99	545
$discount_i$	513,338	-0.445	0.182	-0.990	-0.450	0
$shopFollowers_i$	513,338	55435	150923	15	8416	3,134,696

Topic Modeling

Results of Topics

We apply the GENSIM library in Python for LDA topic modeling and the TF-IDF algorithm for weighted optimization to extract the topics from all processed reviews of consumers of women's fashion on the AliExpress platform. Following the common practice in LDA topic model (Hao, Zhang, Wang, & Gao, 2017), the hyper parameters were set to $\alpha = 0.1$ and $\beta = 0.01$. Considering that the number of review topics for the same category of products is not excessive, we let K , the number of review topics, vary from 1 to 30, and calculate the corresponding coherence value to determine the optimal value of K . Results show that the coherence value is the largest when $K = 13$. Thus, we select 13 as the number of topics extracted from the consumer reviews.

After model fitting, we use words with the top 10 highest probabilities under each topic to characterize the topic (Buschken & Allenby, 2016). Table 3 lists the extracted topics from reviews and their detailed information. As shown in Table 3, the relevance of the words in Topic 2 and Topic 4 is low, and it is difficult to identify their specific topic. Thus, they are referred as shared topics. Except for the shared topics, the other topics are classified into three categories: *Product*, *Service* and *Logistics*.

Table 3: Results of LDA topic modeling and classification.

Topic category	Topic	Top 10 words with the highest probability under the topic
<i>Product</i>	Topic 1: length	height, cm, perfectly, size, took, sat, excellent, approached, length, parameters
	Topic 3: color	color, match, coat, white, transparent, photo, black, pink, different, la
	Topic 9: features	pleasant, suit, threads, stick, body, seams, fabric, soft, quality, good
	Topic 10: detail	critical, shines, beige, size, legs, translucent, little, narrow, thin, good
	Topic 11: accessories	elastic, leggings, tights, band, pantyhose, rubber, fleece, qualitatively, belt, bottom
	Topic 13: size	size, take, small, sizes, larger, good, xl, smaller, big, ordered

Service	Topic 5: seller's reply	seller, quickly, came, thank, corresponds, description, quality, delivery, recommend, order
	Topic 7: salesperson	soon, salesperson, arrive, good, thank, order, color, salesman, quick, third
	Topic 8: Personal feeling	fits, size, small, recommended, described, like, ordered, pictures, nice, good
	Topic 12: after-sale service	money, returned, never, refund, dispute, goods, seller, return, back, come
Logistics	Topic 6: logistics	fast, good, quality, nice, love, super, shipping, delivery, beautiful, great
Shared topics	Topic 2	fur, wife, thought, costume, middle, colors, yet, longer, gray, hair
	Topic 4	shirt, gift, best, item, put, top, likes, knitting, og, new

Variables Construction

To examine the weight differences given to review topics, the topic probability of consumer reviews was used as the core explanatory variable. To help interpret subsequent analyses, while examining content differences, we disregard shared topics.

Based on the topic modeling results, we further calculated the probability of each review belonging to each topic using Python, drawing on the method of Wang et al. (2019). Specifically, θ_{dk} denotes the probability of topic k discussed in review d posted on a given topic. Then, we added up the probabilities of all topics belonging to a category and calculate the total probability of each topic category.

$$p_{di} = \sum_{k \in category_i} \theta_{dk} \quad (2)$$

where i denotes the topic category, and $category_i$ denotes the set of topics belonging to topic category i .

We then conducted descriptive analysis on the review topic categories. As shown in Table 4, overall, consumers pay much attention to the products and services when shopping on CBEC platforms, while logistics as a third-party factor receives relatively little attention.

Table 4: Probability statistics of topics of poor reviews.

Topic categories	N	Mean	Std. Dev	Min	P50	Max
Product	513,338	0.489	0.244	0.080	0.493	2.128
Service	513,338	0.302	0.219	0.049	0.227	1.433
Logistics	513,338	0.096	0.157	0.011	0.028	0.766

RQ1: Topic Comparison between Cross-Cultural Countries

In this study, the U.S. and Russia were selected for further comparison of consumer differences for the following reasons. First, the U.S. and Russia are large and important export markets. Second, the U.S. and Russia are culturally different. According to the cultural cluster theory, the U.S. belongs to the Anglo-Saxon group and Russia belongs to the independent cluster. Also, the U.S. and Russia differs significantly in their scores of the cultural dimension of individualism (46 vs. 95) of the Hofstede's cultural dimension theory.

To examine the topic differences of the reviews between the U.S. and Russia, we conducted a series of two-tail t-tests on the probabilities of each category (Wang et al., 2019). The comparison and t-test results are shown in Table 5. Results of t-tests indicate that, reviews of the U.S. discuss more on topics in *Service* ($0.361 > 0.270$) and *Logistics* ($0.166 > 0.062$); while Russian consumers are significantly more likely to post reviews on the topic of *Product* ($0.559 > 0.359$). Results are all statistically significant ($p < 0.001$). These findings imply that, U.S. consumers are more concerned with *Service* and *Logistics* than Russian consumers, who are more concerned with *Product*. Meanwhile, the U.S. scores significantly lower on the cultural dimension of individualism than Russia. In summary, H1, H2 and H3 were all supported. These findings indicate that, U.S. consumers focus more on *Service* and *Logistics* in the process of CBEC shopping, but less on *Product*.

Table 5: Results of two-tail t-tests of category probability for reviews from the U.S. and Russia.

Topic category	Average topic probability of consumer reviews from the U.S. p_1	Average topic probability of consumer reviews from Russia p_2	Results of the comparison	p value
Product	0.359	0.559	$p_1 < p_2$	< 0.001
Service	0.361	0.270	$p_1 > p_2$	< 0.001

Logistics	0.166	0.062	$p_1 > p_2$	< 0.001
-----------	-------	-------	-------------	---------

RQ2: Across-Countries Impacts of Review Topics on Consumer Satisfaction

Regression Model

To explore the impact of extracted review topics on consumer rating in different cultural contexts, subsample linear regression was conducted on the dataset. Specifically, the subsample includes the focused target markets, the U.S. and Russia.

The dependent variable of this research is consumer satisfaction, and consumer rating is used as the proxy variable. The independent variables are the topic probabilities of each category, which are extracted from the textual information of consumer reviews using text mining techniques. Other factors that may affect consumer ratings are used as control variables. The regression model is shown in equation (3), where product, service, and logistics are the topic probabilities of the product, service, and logistics review categories respectively and i denotes the i th review.

$$\begin{aligned} \text{rating}_i = & \beta_0 + \beta_1 \text{product}_i + \beta_2 \text{service}_i + \beta_3 \text{logistics}_i + \beta_4 \text{shippingMethod}_i + \\ & \beta_5 \text{reviewTime}_i + \beta_6 \text{reviewNum}_i + \beta_7 \text{shelfTime}_i + \beta_8 \text{averageRating}_i + \\ & \beta_9 \text{originalPrice}_i + \beta_{10} \text{discount}_i + \beta_{11} \text{shopFollower}_i + \text{productTypeDummy}_p + \varepsilon \end{aligned} \quad (3)$$

Data Analysis

We explore the influence of the extracted review topics on consumer rating in different cultural contexts. First, a subsample regression was conducted on the review data of the U.S. (i.e., individualism) and Russia (i.e., collectivism), respectively. The results are shown in Table 6.

Table 6: Linear regression results for each subsample.

Variables	(1) rating of the U.S.	(2) rating of Russia
product_i	0.053 (0.0563)	0.319*** (0.0156)
service_i	0.142*** (0.0543)	0.171*** (0.0170)
logistics_i	1.319*** (0.0567)	0.834*** (0.0226)
shippingMethod_i	0.027 (0.0204)	-0.011 (0.00697)
reviewTime_i	0.001*** (0.000134)	0.000*** (4.56e-05)
reviewNum_i	-0.000*** (1.32e-05)	-0.000*** (2.41e-06)
shelfTime_i	-0.002*** (0.000299)	0.001*** (0.000115)
averageRating_i	1.431*** (0.0297)	1.496*** (0.0102)
originalPrice_i	0.001*** (0.000278)	-0.000 (9.51e-05)
discount_i	-0.008 (0.0381)	0.005 (0.0123)
shopFollower_i	0.000** (9.15e-08)	-0.000 (1.11e-08)
Constant	-2.425*** (0.170)	-2.929*** (0.0571)
Control productType	YES	YES
N	31,152	246,112
R-Squared	0.126	0.101

Note: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

From Table 6, all topic categories including *Product*, *Service* and *Logistics*, except for the *Product* variable of the U.S., have a significant positive effect on satisfaction ($p < 0.01$), indicating that the U.S. and Russia consumers are generally satisfied with the product, service and logistics of China's CBEC export women's fashion products.

Next, we compare the coefficients of *Product*, *Service*, and *Logistics* between the U.S. and Russia to explore their consumer differences. The coefficients of both *Product* ($0.319 > 0.053$) and *Service* ($0.171 > 0.142$) are greater in Russia than in the U.S., indicating that Russian consumers are more satisfied with products and services. A possible explanation for Russian consumers being more satisfied with products might be that, Russian consumers have greater demand for clothing. Russian textile production

has been in decay over the past two decades. Low production rates led to the trade deficit, resulting in dependence on apparel imports. Meanwhile, the COVID-19 pandemic boosted e-commerce sales in Russia, including the apparel segment. In the following years, the popularity of apparel shopping on the internet was projected to increase. All these have led to an increase in demand for online imported clothing from Russian consumers. Besides, U.S. consumers are dissatisfied with the service, probably because the transaction process of CBEC is relatively cumbersome, thus failing to meet U.S. consumers' higher expectations of services. According to Hofstede's Cultural Dimensions, the U.S. has lower score of power distance than Russia (40 vs. 93). Therefore, U.S. consumers tend to have higher expectation for service quality (Ladhari, Pons, Bressolles, & Zins, 2011). Moreover, the coefficients of *Logistics* ($1.319 > 0.834$) are greater in the U.S. than in Russia, which means that U.S. consumers are more satisfied with logistics.

Other results in Table 6 show that, first, U.S. consumers are less satisfied with products with longer shelf time than Russian consumers, indicating that they prefer latest products. Second, the price control variables *originalPrice* and *discount* both have no effect on Russian consumers' satisfaction, which means that Russians are much less sensitive to price factors.

DISCUSSION AND CONCLUSION

This study mainly focuses on three questions: First, the cultural differences in consumer reviews on cross-border e-commerce platforms. Second, the heterogeneous impacts of review topics on consumer satisfaction across countries. Third, the topics underlying negative online reviews.

To address these three questions, this study compares the preferences and satisfaction of the U.S. and Russian consumers through their posted reviews. We focus on reviews of women's fashion on the AliExpress platform and apply LDA topic modelling techniques to extract information from the review text. We obtained review topics from overseas consumers in three categories: *Product*, *Service* and *Logistics* and compared reviews from the United States and Russia. Among these topic categories, U.S. consumers focus more on *Service* and *Logistics*, while Russian consumers pay more attention to *Product*.

Then, the regression results are as follows: First, overall, consumers in the United States and Russia are satisfied with the *Product*, *Service* and *Logistics* of the AliExpress platform. Second, U.S. consumers are more satisfied with *Logistics*, while Russian consumers are more satisfied with *Product* and *Service*. Meanwhile, Russian consumers post significantly higher ratings than U.S. consumers. Third, U.S. consumers are more interested in new products, and they are more sensitive to the price of goods.

Theoretical and Managerial Implications

This study extends the research on CBEC export differentiation and provides the following theoretical contributions. First, the study uncovers the differences in consumer demand across countries by employing LDA on different sites of a CBEC export platform, enriching CBEC studies using textual analysis. Second, from the standpoint of helping CBEC export, this study bridges the gap that prior studies only focused on a specific country or region. Third, from a cross-cultural perspective, as we know, this is the first study that investigates and compares the underlying factors affecting consumer satisfaction in different overseas market.

This study also provides insights for the CBEC exports industry. First, the findings show that consumers in different countries have significantly different preferences, which provides enterprises with valuable information for using big data to determine consumer value. Second, the regression identifies factors that affect consumer satisfaction, which could help enterprises target their products and services in overseas markets. Third, understanding the similarities and differences of consumers in target markets could assist CBEC enterprises in adapting to local conditions and making timely strategic adjustments. This not only helps enterprises to innovate and transform, but also enhances their international competitiveness. This allows enterprises to make better use of the resources of the international market and cultivate new advantages for participating in international cooperation and competition.

Limitations and Future Directions

The study also has some limitations. First, the study used consumer feedback for a specific product type, women's fashion. The selection of a single category restricts the generalizability of the conclusions. Future research is suggested to extend the analytical framework to a wider range of product categories. Second, due to platform restrictions, we could only use six months of data. Future research may investigate cross-cultural differences in consumer satisfaction and its influencing factors using long-term data. Finally, the performance of LDA topic modeling is affected by the hyper-parameters. Therefore, future work could adjust the hyper-parameters to achieve improved modeling results.

ACKNOWLEDGMENT

This work is partially supported by National Natural Science Foundation of China (72102039, 72201269, 72172036).

REFERENCES

- Büschken, J., & Allenby, G. M. (2016). Sentence-based text analysis for customer reviews. *Marketing Science*, 35(6), 953-975. <https://doi.org/10.1287/mksc.2016.0993>
- Clemons, E. K., & Gao, G. G. (2008). Consumer informedness and diverse consumer purchasing behaviors: Traditional mass-market, trading down, and trading out into the long tail. *Electronic Commerce Research and Applications*, 7(1), 3-17. <https://doi.org/10.1016/j.elerap.2007.10.001>

- De Mooij, M., & Hofstede, G. (2002). Convergence and divergence in consumer behavior: implications for international retailing. *Journal of Retailing*, 78(1), 61-69. [https://doi.org/10.1016/S0022-4359\(01\)00067-7](https://doi.org/10.1016/S0022-4359(01)00067-7)
- Ding, Y., Gao, X., Huang, C., Shu, J., & Yang, D. (2018). Service competition in an online duopoly market. *Omega*, 77, 58-72. <https://doi.org/10.1016/j.omega.2017.05.007>
- Faqih, K. M., & Jaradat, M. I. R. M. (2015). Assessing the moderating effect of gender differences and individualism-collectivism at individual-level on the adoption of mobile commerce technology: TAM3 perspective. *Journal of Retailing and Consumer Services*, 22, 37-52. <https://doi.org/10.1016/j.jretconser.2014.09.006>
- Gao, W., & Li, W. (2015). Research on service innovation strategy of B2C e-commerce enterprise. *Management & Engineering*, (21), 13-16. <https://doi.org/10.5503/J.ME.2015.21.003>
- Pankaj, G. (2001). Distance still matters: the hard reality of global expansion. *Harvard business review*, 79(8), 137-147.
- Hao, H., Zhang, K., Wang, W., & Gao, G. (2017). A tale of two countries: International comparison of online doctor reviews between China and the United States. *International Journal of Medical Informatics*, 99, 37-44. <https://doi.org/10.1016/j.ijmedinf.2016.12.007>
- Jia, S. S. (2020). Motivation and satisfaction of Chinese and US tourists in restaurants: A cross-cultural text mining of online reviews. *Tourism Management*, 78, 104071. <https://doi.org/10.1016/j.tourman.2019.104071>
- Knight, G., Koed Madsen, T., & Servais, P. (2004). An inquiry into born-global firms in Europe and the USA. *International Marketing Review*, 21(6), 645-665. <https://doi.org/10.1108/02651330410568060>
- Ladhari, R., Pons, F., Bressolles, G., & Zins, M. (2011). Culture and personal values: How they influence perceived service quality. *Journal of Business Research*, 64(9), 951-957. <https://doi.org/10.1016/j.jbusres.2010.11.017>
- Lee, M., Kang, M., & Kang, J. (2019). Cultural influences on B2B service quality-satisfaction-loyalty. *The Service Industries Journal*, 39(3-4), 229-249. <https://doi.org/10.1080/02642069.2018.1495710>
- Leon, R. D. (2019). Hotel's online reviews and ratings: a cross-cultural approach. *International Journal of Contemporary Hospitality Management*, 31(5), 2054-2073. <https://doi.org/10.1108/IJCHM-05-2018-0413>
- Li, F. (2021, Jun 28-Jul 02). Research on China's Cross -Border E -Commerce Promotion Strategy of Foreign Trade Transformation and Upgrading Based on the Perspective of Industrial Integration in the Internet Era. *In the 17th IEEE International Wireless Communications and Mobile Computing Conference (IEEE IWCMC)*, ELECTR NETWORK.
- Li, Y. Q., & Shen, H., & (2018). Cross Cultural Study of Customers Online Reviews Intention of the E-commerce Platforms. *In 2018 International Conference on E-Commerce and Contemporary Economic Development (Eced 2018)*, 18-22.
- Ludwig, S., De Ruyter, K., Friedman, M., Brüggem, E. C., Wetzels, M., & Pfann, G. (2013). More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. *Journal of Marketing*, 77(1), 87-103. <https://doi.org/10.1509/jm.11.0560>
- Maurer, C., & Schaich, S. (2011). Online customer reviews used as complaint management tool. *In Information and Communication Technologies in Tourism 2011* (pp. 499-511): Springer , Vienna. https://doi.org/10.1007/978-3-7091-0503-0_40
- Mou, J., Ren, G., Qin, C., & Kurcz, K. (2019). Understanding the topics of export cross-border e-commerce consumers feedback: an LDA approach. *Electronic Commerce Research*, 19, 749-777. <https://doi.org/10.1007/s10660-019-09338-7>
- Punel, A., Hassan, L. A. H., & Ermagun, A. (2019). Variations in airline passenger expectation of service quality across the globe. *Tourism management*, 75, 491-508. <https://doi.org/10.1016/j.tourman.2019.06.004>
- Qin, X., Liu, Z., & Tian, L. (2020). The strategic analysis of logistics service sharing in an e-commerce platform. *Omega*, 92, 102153. <https://doi.org/10.1016/j.omega.2019.102153>
- Rose, S., Clark, M., Samouel, P., & Hair, N. (2012). Online customer experience in e-retailing: an empirical model of antecedents and outcomes. *Journal of Retailing*, 88(2), 308-322. <https://doi.org/10.1016/j.jretai.2012.03.001>
- Tian, R. G., & Lan, X. (2009). E-commerce concerns: Cross-cultural factors in internet marketing. *In 2009 International Conference on Electronic Commerce and Business Intelligence* (pp. 83-86). IEEE. <https://doi.org/10.1109/ECBI.2009.55>
- Tsiotsou, R. H. (2022). Identifying value-creating aspects in luxury hotel services via third-party online reviews: a cross-cultural study. *International Journal of Retail & Distribution Management*, 50(2), 183-205. <https://doi.org/10.1108/IJRDM-04-2021-0207>
- Wang, F., Yang, Y., Tso, G. K., & Li, Y. (2019). Analysis of launch strategy in cross-border e-Commerce market via topic modeling of consumer reviews. *Electronic Commerce Research*, 19, 863-884. <https://doi.org/10.1007/s10660-019-09368-1>
- Wu, G. (2016). Build cross-border E-commerce" Overseas warehouse" platform effectively. *In 2016 2nd International Conference on Social Science and Technology Education (ICSSTE 2016)* (pp. 402-406). Atlantis Press. <https://doi.org/10.2991/icsste-16.2016.74>
- Xiao, L. (2018). Analyzing consumer online group buying motivations: An interpretive structural modeling approach. *Telematics and Informatics*, 35(4), 629-642. <https://doi.org/10.1016/j.tele.2018.01.010>
- Zhu, D. H., Ye, Z. Q., & Chang, Y. P. (2017). Understanding the textual content of online customer reviews in B2C websites: A cross-cultural comparison between the US and China. *Computers in Human Behavior*, 76, 483-493. <https://doi.org/10.1016/j.chb.2017.07.045>