Electronic Commerce Research and Applications 29 (2018) 1–11

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

journal homepage: www.elsevier.com/locate/ecra

Impact of product attributes on customer satisfaction: An analysis of online reviews for washing machines

Yuren Wang [a], Xin Lu [b,a,*], Yuejin Tan [a]

[a] College of Systems Engineering, National University of Defense Technology, Changsha, China
[b] School of Business, Central South University, Changsha, China

ARTICLE INFO

Article history:
Received 9 January 2018
Received in revised form 5 March 2018
Accepted 5 March 2018
Available online 10 March 2018

Keywords:
Customer behavior
Customer satisfaction
Online reviews
Product attributes
Product design

ABSTRACT

Online reviews are an important information source for companies analysing users' demands. We conducted a study of online reviews to measure how product attributes impact customer satisfaction. First, we attempted to infer through sentiment analysis whether a customer is satisfied with a purchase according to their review. Second, a logistic regression model was developed to estimate the impact of various product properties on customer satisfaction scores. Our estimates indicated that customer satisfaction is influenced by drainage mode, loading type, frequency conversion, type, display, colour, and capacity. We further investigate the impact of price and find that customers who buy cheap products should be treated differently from purchasers of expensive items because the relevance of design features on their satisfaction is different. Additionally, we observed that although customers are concerned about noise, perceived noise is not consistent with actual noise levels. We analysed specific reviews and then obtained more detailed information on customer attitudes.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Understanding customers' behaviours and reactions to product design and marketing is critical for manufacturers. Customers' satisfaction influences their loyalty to the company. Thus, satisfying both current and potential customers' needs has become a serious challenge for market-driven product design. In order to improve customer satisfaction, it is necessary to study how product attributes affect customer satisfaction.

Existing research on customer behaviour typically relies on interviewing respondents, either by means of traditional paper-and-pencil surveys or online questionnaires. These methods are usually expensive and time-consuming (Zhou et al., 2016). Furthermore, the quality of the data obtained from surveys depends on the willingness of respondents to participate in the study and might be biased by the length or complexity of the questionnaire (Groves, 2006). Internet retailing and e-commerce have provided new opportunities for improved customer behaviour analysis. As e-commerce becomes more popular, the number of customer reviews that a product receives grows rapidly. These reviews are an important information source for companies analysing users' demands (Guo et al., 2017) and can provide crucial timely feedback (Abrahams et al., 2012). In contrast to administered questionnaires, customer reviews are freely offered evaluations that can reflect customer concerns more accurately.

A large number of studies have been completed in recent years to determine how best to detect customers' priorities from online reviews. However, only a few studies discuss the value of online reviews for product design. Thus, this research studies how online reviews help product designers analyse customer requirements. A detailed analysis of washing machine reviews from Suning.com, one of the leading Chinese business-to-customer (B2C) online shopping platforms, was conducted. In particular, the level of satisfaction exhibited by each customer in 117,585 reviews was quantified through sentiment analysis. Each customer's satisfaction was then measured and related to a set of product attributes, such as color, drainage mode and loading type, etc. Next, this study investigated the impact of price on satisfaction and the differences between customers who bought products at different price points. Additionally, the relationship between a customer's perception of a washer's noise level and the unit's actual noise level was studied. Last, *latent Dirichlet allocation* (LDA) (Blei et al., 2003; Blei and Lafferty, 2007; Jo and Oh, 2011; Wei and Croft, 2006) was used to extract customer attitudes from reviews that discussed frequency conversion and drainage mode.

The research is designed to identify the following key factors. First, customer satisfaction was measured and the findings showed that type, color, drainage mode, capacity, frequency conversion,

* Corresponding author at: School of Business, Central South University, Changsha 410083, China.
E-mail address: xin_lyu@sina.com (X. Lu).

https://doi.org/10.1016/j.elerap.2018.03.003
1567-4223/© 2018 Elsevier B.V. All rights reserved.

display and loading type each have a statistically significant impact on satisfaction. Specifically, customers were more satisfied with washing machines which incorporate a pulsator, golden, large capacity, down drainage, frequency conversion, a liquid crystal display (LCD) screen, and front loading. However, other attributes – including automatic operation type, control mode, material, energy efficiency class, and auto power off – did not significantly affect customer satisfaction.

Second, our study investigated the differences between customers at different economic levels. Product price may reflect a customer's financial situation, so customers are divided into three groups according to the prices of washing machines they buy. In general, more expensive products are associated with higher levels of satisfaction. However, this is not always the case for all three customer income groups. Increased price is associated with a decrease in the satisfaction of customers who buy a low-priced product. Customers who buy high-priced products are not sensitive to model price. Moreover, results indicate that customers who buy expensive products and those who buy inexpensive products have different priorities for design features.

Third, through analysis of reviews discussing the noise of washing machines, it can be found that customers are concerned about washing machine noise, but their perceptual sensitivity to noise nuisance varies. Customers may complain that their washing machine is noisy even though its specified noise rating is not especially high. Finally, we assessed reviews related to these two specific factors. Our results show that both up-drainage and down-drainage designs have advantages, but almost all customers prefer frequency-conversion washing machines.

2. Literature review

The role of online communities, particularly in the context of new product development, has been discussed by many studies (e.g., Franke and Piller, 2003). Lee (2007) highlighted that online product reviews enable marketers and manufacturers to gain more complete understandings of customers. Zhu and Zhang (2010) proved that online customer reviews can be a good proxy for communicating customer experience by word-of-mouth. In Jin et al. (2014), an ordinal classification approach and an integer programming model were implemented to convert online reviews into the corresponding original customer satisfaction ratings.

First, the literature on sentiment analysis are evaluated. Sentiment analysis refers to the use of *natural language processing* (NLP), text analysis, computational linguistics, and biometrics to analyse opinions and discover how customers feel. Much research has addressed the problem of assessing emotional cues by analysing polarity in context (e.g., Wiebe and Riloff, 2005; Wilson et al., 2005). To improve the accuracy of sentiment classification, some special techniques (e.g., Mullen and Collier, 2004; Xia et al., 2013; Nakagawa et al., 2010; Hassan and Radev, 2010) have been developed. For example, Qiu et al. (2009) proposed a novel propagation approach that exploits the relations between sentiment words and features, to extract new sentiment words and features. Many studies have identified or extracted customer sentiments from online reviews. Hu and Liu (2004) mined the product features about which the customers expressed opinions. The researchers then used analytical methods to detect from textual criteria whether customer opinions are positive or negative. More recently, some authors have pointed out the power of research designs that are founded on Computational Social Science methods of inquiry, including the use of machine-based methods combined with explanatory econometrics for research involving NLP (Chang et al., 2014, Kauffman et al., 2017).

In Liu et al. (2005), a new technique based on language pattern mining is proposed as a way to extract details about product features from a particular type of reviews. Zagibalov and Carroll (2008) describe and evaluate a new method of automatic seed word selection for unsupervised sentiment classification of product reviews in Chinese. Lin and He (2009) proposed a joint sentiment/topic model, which detects sentiment and topic simultaneously in reviews. Hedegaard and Simonsen (2013) investigated the content of online reviews to plot the distribution of information in reviews according to different dimensions of usability and user experience.

Some studies have discussed the value of online reviews in product design processes. Decker and Trusov (2010) presented an econometric framework that allows inferences to be made about the relative effects of product attributes and brand names on the overall evaluation of products. Tucker and Kim (2011) proposed a robust framework for enriching product design processes by dynamically capturing customer preference trends from publicly available product review data. Goorha and Ungar (2010) described a system that monitors social and mainstream media to determine shifts in how people are thinking about a product. Wang et al. (2011) developed a systematic methodology for deriving product attributes from online reviews, constructing customer preference models by means of Bayesian linear regression, and using these models in design selection. Jin et al. (2012), proposed a supervised learning routine to identify product characteristics, which could then inform an ordinal classification algorithm to prioritize engineering characteristics for designers. Some work also has sought to capture signs of customers' changing requirements and to predict subsequent design trends (e.g., Lee, 2007; Tucker and Kim, 2011).

In particular, past research has shown how detecting product defects improves revised designs' performance. Social media surveillance, text classification, and sentiment analysis have been used successfully in previous work on defect detection. Abrahams et al. (2012, 2013, 2015) developed text analytic frameworks for defect detection and apply these methods to discovering flaws in the automotive and consumer electronics industries. Law et al. (2017) extended the text analytics framework for the detection of under-performance in large home appliances, the results of which will be beneficial for improving dishwasher appliance quality management methods. In Jin et al. (2016), aspects of product features and consumers' detailed reasons for choosing features were extracted from online reviews to inform designers about characteristics that lead to dissatisfied user experiences.

3. Data and method

3.1. Data description

The data used in this study involve washing machine review data extracted from the Suning.com website. Suning.com is a new B2C online shopping platform, which is filled with traditional home appliances, daily necessities, and other product categories. It has more than 500 million monthly active users and ranks among the top three Chinese B2C companies. Suning.com provided two datasets: a review dataset and a product dataset.

The review dataset recorded 117,585 reviews of customers, including product ID, product ratings (in the form of one-to-five stars) and full-text consumer comments. The reviews were published in the period from March 2012 to November 2014. After filtering out empty and duplicated records, 105,263 reviews remained. The product dataset consists of the information about each product's attributes and functions. For each washing machine, the data contains product ID, name, automatic type, control mode, etc. In order to obtain the product attributes of each review, the product dataset was joined with the review dataset. This yielded 71,909 review records for this research.

3.2. Sentiment analysis

It is common in China for some retailers to offer bonuses to customers if they give favourable comments online. Thus, customers might be inclined to give a product a higher numerical rating, even though they are not satisfied. Consumers are more likely to express their true feelings in online reviews than in one-to-five star summaries. As can be seen from the dataset, many numerical ratings that do not match the reviews, inasmuch as many customers have expressed dissatisfaction in their textual comments, but they have all given 5-stars in the rating to get the bonus. Because customer reviews often reflect more closely their real attitudes, we use the review comments instead of numbered ratings to evaluate customer satisfaction.

NLP technology was used to analyse the reviews. There are some key differences between Chinese- and English-language expressions (Zhou et al. 2016). SnowNLP, a Python library that specializes in analysing Chinese, is used to process the reviews. It has functions similar to TextBlob (a Python library for processing English textual data), such as part-of-speech tagging, noun phrase extraction, sentiment analysis, and text abstraction.

We first trained the dictionary in SnowNLP with the review commentary data. Some review samples were tagged manually and then stored in the existing dictionary to improve the accuracy of the sentiment analysis. SnowNLP can predict the probability that a sentence is positive or negative. Analysis obtained 55,291 positive reviews and 16,618 negative reviews. The sort of comments posted by each customer were then used to measure the level of satisfaction and entered in the model below.

3.3. Regression model

The model is based on the relationships between customer satisfaction (i.e., the sentiment analysis result) and product attributes. We define a binary variable $SATISFACTION_i$ with $SATISFACTION_i = 1$ if the sentiment analysis result of ith review record was positive, and $SATISFACTION_i = 0$ otherwise.

A series of dummy variables were created. For example, there are three types of washing machines, including pulsator, drum, and washer-dryer. The measure $TYPE_i$, consists of two dummy variables: *TYPE_PULSATOR* and *TYPE_DRUM*. They were estimated to capture the influence of type. Similarly, the impacts of color ($COLOUR_i$), material ($MATERIAL_i$), energy-efficiency class ($ENERGY_i$), and display screen ($DISPLAY_i$) were examined. They all contain a set of dummy variables. To account for the influence of automatic type, the dummy variable $AUTOMATION_i$ was constructed. Washing machines can be automatic and semi-automatic. The binary variable $AUTOMATION_i$ was defined with $AUTOMATION_i = 1$ if the ith washing machine was automatic and $AUTOMATION_i = 0$ if it was semi-automatic. Similarly, other attributes were considered by using variables denoting the control mode ($CONTROL_i$), drainage mode ($DRAINAGE_i$), loading type ($LOADING_i$), frequency conversion ($FREQUENCY_i$) and auto power off ($AUTOOFF_i$). In addition, the variable $CAPACITY_i$ was introduced as an indicator for the capacity of the washing machine.

We formulated a logistic regression to estimate the impact of the above variables on $SATISFACTION_i$.

$$\begin{aligned}\text{logit}[\Pr(SATISFACTION_i) = 1] = \beta_0 &+ \mathbf{X_i}\beta_1 + \beta_2 AUTOMATION_i \\ &+ \beta_3 TYPE_i + \beta_4 COLOR_i \\ &+ \beta_5 MATERICAL_i + \beta_6 ENERGY_i \\ &+ \beta_7 DISPLAY_i + \beta_8 CONTROL_i \\ &+ \beta_9 DRAINAGE_i + \beta_{10} LOADING_i \\ &+ \beta_{11} FREQUENCY_i + \beta_{12} AUTOOFF_i \\ &+ \beta_{13} CAPACITY_i \qquad (1)\end{aligned}$$

$\mathbf{X}_i$ is a set of variables that control for the underlying heterogeneity in product and customer characteristics, including price ($PRICE_i$), noise ($NOISE_i$), brand ($BRAND_i$), weight ($WEIGHT_i$), water consumption ($CONSUMPTION_i$), and family type ($FAMILY_i$). The integration of those attributes is not entirely dependent on production designers, so those attributes are included as control variables to correct for the heterogeneity. Table A1 in the Appendix provides a list of all the variables and controls.

4. Results

4.1. Effects of product attributes on customer satisfaction

Table 1 summarizes the results of estimating the model with customer satisfaction as a dependent variable based on the sample of 71,909 reviews. Only coefficients that are statistically significant are shown in the table. Table A1 in the Appendix provides results with all of the variables.

The regression results in Table 1 show that the coefficient for the dummy variable *TYPE_PULSATOR* ($\beta_{3\_pulsator}$) has a value of 1.012 and the coefficient for *TYPE_DRUM* (β_{3_drum}) has a value of 0.130. They indicate that compared to a combined washer-dryer machine, a pulsator or a drum washing machine led to a 175.1% or 13.9% increase in satisfaction, respectively. Also, satisfaction decreased if the color was changed from silvery to white or grey. However, the coefficient for *COLOUR_GOLDEN* (β_{4_golden}) was estimated to be 0.095 ($p = 0.008$), providing evidence that customer satisfaction improved by 10.0% if the washing machine was golden rather than silvery. The effect of display type was also significant. As Table 1 illustrates, an integral LCD screen (β_{7_LCD}) was associated with a 15.1% higher satisfaction. But the effect for a light emitting diode (LED) screens was not statistically significant at the 10% level, a result that does not support the common hypothesis that customers are more satisfied with LED screens than with no display screen.

The data also showed that up drainage had a negative effect ($\beta_{9\_up} = -0.106$, $p < 0.001$), indicating that up drainage was associated with 10.1% decreased satisfaction. Clearly, compared with up drainage designs, users were more satisfied with down drainage. When we examined the effect of loading type on customer satisfaction, we found that the coefficient is -0.702 ($p = 0.002$). Customers were more satisfied with front-loading than with top-loading washing machines. Next, we found that the coefficient for *FREQUENCY* ($\beta_{11\_conversion}$) was estimated to be 0.217 ($p < 0.001$), suggesting that customers were about 24.2% more satisfied if the washing machine had frequency conversion.

Furthermore, the results show that larger capacity ($\beta_{13} = 0.081$) was associated with higher satisfaction. Specifically, a unit's increase in *CAPACITY* led to an 8.4% increase in satisfaction. The coefficients for *MATERIAL*, *ENERGY*, *AUTOMATION*, *CONTROL* and *AUTOOFF* were not statically significant at the 10% level, suggesting that those attributes are not relevant to customer satisfaction. Thus, they should not be prioritized when manufactures design new washing machine models.

In summary, we found seven attributes related to customer satisfaction. From the standardized estimates of coefficients, it can be seen that the type of washing machine and loading type had relatively larger effects. Other attributes also had statistically significant impacts, but their influences were smaller. When designing a product, manufacturers should seriously consider the seven attributes that significantly influence customer satisfaction. Further, when resources are limited, it may be unnecessary for manufacturers to improve the performance attributes that have no significant impact.

Table 1
Effect of production attributes on customer satisfaction.

Variable	Coef (SE)	Odds Ratio	Standardized Estimate (SE)	p-value
$\beta_{3\_pulsator}$	1.012 (0.264)***	2.751	0.279 (0.264)***	<0.001
β_{3_drum}	0.130 (0.071)*	1.139	0.036 (0.071)*	0.066
β_{4_white}	−0.143 (0.033)***	0.867	−0.034 (0.033)***	<0.001
β_{4_grey}	−0.196 (0.034)***	0.822	−0.050 (0.034)***	<0.001
β_{4_golden}	0.095 (0.036)***	1.100	0.018 (0.036)***	0.008
β_{7_LCD}	0.141 (0.056)**	1.151	0.022 (0.056)**	0.013
β_{9_up}	−0.106 (0.035)***	0.899	−0.026 (0.035)***	0.002
β_{10_top}	−0.702 (0.223)***	0.496	−0.193 (0.223)***	0.002
$\beta_{11\_conversion}$	0.217 (0.065)***	1.242	0.060 (0.065)***	<0.001
β_{13}	0.081 (0.019)***	1.084	0.057 (0.019)***	<0.001
Likelihood ratio				<0.001

Note: Variables for $\mathbf{X}_i$ and not statistically significant at the 10% level are not displayed. ***, **, and * denote significance at the 1%, 5% and 10% levels. Standard errors are shown in parentheses.

Table 2
Descriptive statistics for price.

Mean	Std Dev	Min	Median	Max	Lower Quartile	Upper Quartile
2398.71	1635.14	339	2164.33	7199	953	3226

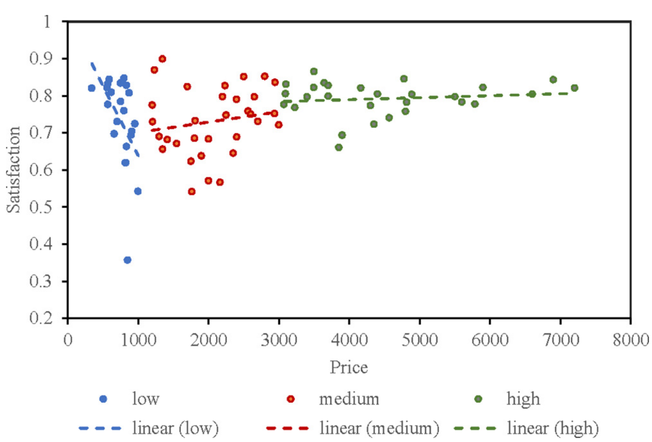

Fig. 1. Satisfaction as a function of price.

Table 3
Effect of attributes on customer satisfaction in different price groups.

Variable	Model 1 Coef (SE)	Model 2 Coef (SE)	Model 3 Coef (SE)
$\beta_{3\_pulsator}$		0.722 (0.302)**	
β_{4_white}	−1.051 (0.291)***	−0.238 (0.050)***	−0.323 (0.096)***
β_{4_grey}	−0.780 (0.351)**		−0.276 (0.149)*
β_{4_golden}	−0.745 (0.381)*		−0.149 (0.087)*
β_{7_LED}	0.205 (0.067)***		−0.374 (0.128)***
β_{7_LCD}			−0.319 (0.145)**
β_{9_up}		−0.273 (0.062)***	
β_{10_top}		−0.553 (0.276)**	
$\beta_{11\_conversion}$		0.314 (0.103)***	
β_{13}	0.006 (0.002)***	−0.0003 (<0.001)*	
β_{price}	−0.002 (0.001)***	0.0003 (<0.001)***	
β_{noise}	−0.006 (0.003)*	0.012 (0.003)***	0.011 (0.004)***
Likelihood ratio	$p < 0.001$	$p < 0.001$	$p < 0.001$

Notes: The table shows the estimates of coefficients. Variables for $\mathbf{X}_i$ and not statistically significant at the 10% level (provided in the appendices) are not displayed. Mode 1 is the model for low-priced product. Model 2 is for medium-priced products. Model 3 is for high-priced products. ***, **, and * denote significance at 1%, 5% and 10% levels. Standard errors are shown in parentheses.

4.2. Effect of price on customer satisfaction

We note that price also had a statistically significant effect on customer satisfaction. The coefficient *PRICE* was estimated to be 0.0002 ($p < 0.001$). We guess that the effect of price on satisfaction may be complicated. From Table 2, we see that the average price for the washing machine that customers reviewed was ¥2398.71, and the upper and lower quartiles were ¥3226 and ¥953, respectively. For simplicity's sake, we divided the price into three ranges: low (0–¥1000), medium (¥1000–¥3000) and high (¥3000 and above). The number of reviews in the three groups were 23,216, 30,114 and 18,579, respectively.

Consider the relationship between satisfaction (i.e., the proportion of satisfied customers) and price, as shown in Fig. 1. We observed a pattern: satisfaction decreased with price when the washing machine was cheap, but a higher price led to higher satisfaction when the washing machine was of moderate price or expensive. Thus, there may be some heterogeneities that distinguish customers who buy budget washers from those who buy expensive washing machines.

Thus, we have analysed the impact of attributes in different price ranges. The models for low-priced, medium-priced, and high-priced units are defined as Models 1, 2 and 3, and the results are shown in Table 3. Tables A2–A4 in the Appendix provide complete results with all variables for Model 1, 2 and 3.

In Model 1, the coefficient for *PRICE* had a value of −0.002 ($p < 0.001$). It suggests that a higher price was associated with lower satisfaction. In addition, some of the analysis results were consistent with previous findings. Specifically, customers were more satisfied with silver color than with white and grey finishes. Larger load capacity was associated with higher satisfaction.

Unlike the results of Model 1 though, higher price was associated with higher satisfaction in Model 2. The results show that customers were more satisfied with a silvery washing machines than white ones. Moreover, pulsator washing machines were more satisfying than drum washing machines. In addition, a down-draining washing machine was preferred to an up-draining one. And a washing machine with front loading and frequency conversion was associated with high satisfaction. Last, greater capacity was related to higher satisfaction too.

Next, consider the results for Model 3. The coefficient for *PRICE* was not statistically significant at the 10% level, which means price was not a key factor impacting the satisfaction of customers who bought expensive products. And some attributes which had

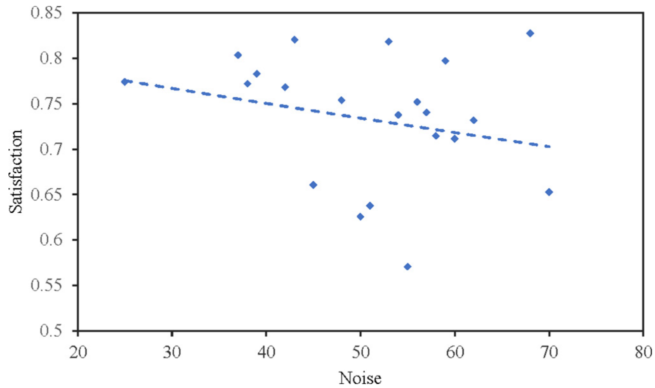

Fig. 2. Satisfaction as a function of noise. Note: Dashed line is linear-fitted curve.

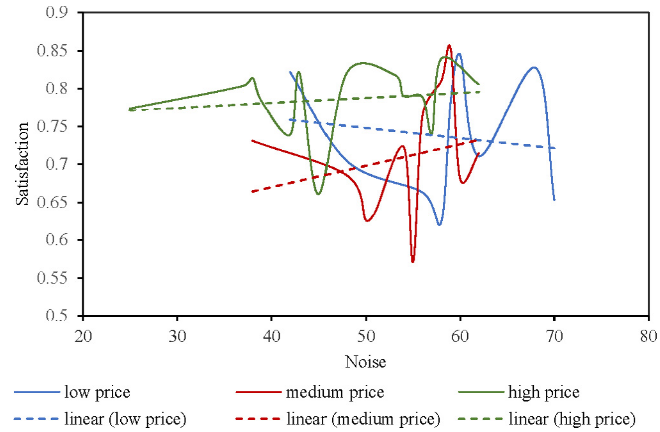

Fig. 3. Satisfaction as a function of noise in different price groups.

significant effects in Model 2 did not have statistically significant impacts in Model 3. Those attributes include drainage mode, loading type, frequency conversion, and capacity. The results suggest that affluent people were not concerned with these attributes.

The results for the three groups are different. For the products in the low-price and expensive groups, there were fewer attributes significantly related to user satisfaction. One explanation for this is that customers who buy low-priced products only expected basic functionality. As a result, the product's other attributes were not considered essential. By contrast, customers who bought high-priced products cared more about advanced functions and services, such as intelligent controls and automatic drying. Consequently, the other attributes likely went unnoticed. Their presence or absence showed no significant effects.

The impact of price on user satisfaction is also examined. Higher satisfaction was associated with lower price when the product was relatively cheap, but customers were more pleased with a higher-priced unit when they purchased a moderately-priced appliance. For expensive products, the influence of price on satisfaction was not significant. Such variances in responses are rational if customers' different financial conditions are taken into perspective. Customers with lower incomes might have wished the product could be cheaper, so their satisfaction decreased with the increase in price. Customers who bought the moderately-priced products might have had higher incomes than the former group of buyers. Purchasers of medium-priced units were primarily concerned with a washing machine's quality, which generally correlates with higher prices. Thus, the expensive product would likely enhance their satisfaction. Last, customers who bought the most expensive products would be relatively wealthier, so they were not sensitive to price.

4.3. Effect of noise on user satisfaction

Another result should be noted is that the coefficient for *NOISE* is 0.012 ($p < 0.001$) in Model 2 and 0.011 ($p = 0.002$) in Model 3. *NOISE* is the noise level generated by the unit's operation, which is officially rated by manufacturers as a specification parameter of the washing machine. Our estimates indicate that higher noise levels led to higher satisfaction when the washing machine was expensive – a discovery that does not seem to conform to common sense. Consequently, we have given careful consideration to possible explanations for this phenomenon.

First, we tries to find a relationship between customer satisfaction and *NOISE*. The reviews are divided into a number of groups according to the noise level, and the percentage of satisfied users in each group is calculated. Fig. 2 presents the relationship between satisfaction (i.e., the percentage of satisfied users) and noise level. We found a pattern that indicates satisfaction decreases with increasing noise level – a result that fits our theoretical understanding. However, this trend was not uniform for customers in all the three models.

Subsequently, we scrutinized the relationship between consumer satisfaction and noise level, classified by different price groups, the results of which are shown in Fig. 3.

The pattern is consistent with the previous regression results. When the washing machine was cheap, increased noise was associated with reduced satisfaction. Nevertheless, customers were more satisfied with a washing machine that emits high levels of noise when the machine's price was moderate or high.

This may be related to customers' different tolerances for noise or to their personal use experiences. To confirm this hypothesis, we conducted a text analysis on the reviews. First, phrases related to noise were extracted and a part of them marked as training samples. Phrases that describe low noise, such as "The washing machine is quiet", were marked positive, and phrases that describe high noise, such as "It is making too much noise", were marked negative. Both sets of phrases were stored in the dictionary of SnowNLP Python library. Next, we implemented a sentiment analysis based on the dictionary. Finally, we divided review phrases into groups based on the noise levels of washing machines and calculated the proportion of customers in each group who perceived high noise levels.

Fig. 4 presents the relationship between the perception of noise level and the empirically verified noise level of washing machine. Perception of noise level is defined as the proportion of customers whose perceived noise level is high. When the price of washing machine was less than ¥1000, customers' perception of noise level increases in accord with the increase in noise level specified by the washing machine's manufacturer. However, when the unit's price was higher than ¥1000, the owner's perception of the washer's operational noise did not necessarily increase to match the manufacturer's noise rating. These discrepancies might have been be due to differences of customers' noise tolerance. The operational environment of washing machine might also have affected customers' subjective experience. For example, the sounds of a washing machine might be normal during the day but seem much louder during the quiet of the night.

5. Discussion

From the results above, we clearly detected that manufacturer specified design attributes affect customers' satisfaction. To obtain a fully detailed, comprehensive picture of some particular customer's attitudes, we extracted the reviews related to a selection

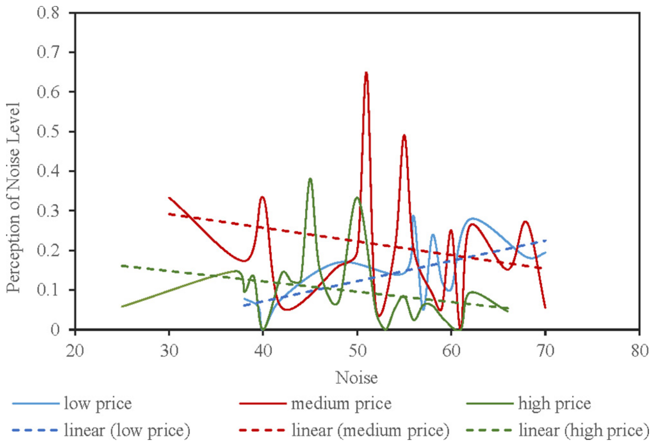

Fig. 4. Perception of noise level as a function of noise in different price groups.

of attributes and analysed them through an LDA model. Because drainage mode and frequency conversion have been shown to significantly impact customers' satisfaction, and the relevant reviews were sufficient, these comments were examined for indicators of owners' detailed attitudes.

The results above show that customers preferred draining water from the bottom rather than the top. Further analysis suggested the reasons for these preferences. Reviews related to drainage modes were isolated from the dataset and divided into up-drainage group and down-drainage groups. These were analysed through the LDA model too. Tables B1 and B2 in the Appendix provide the detailed results. From these, it can be seen that different customers have different attitudes toward the two modes of drainage. The operation of the up-drain washing machines is more complicated, but it allows customers to install the washing machine in more convenient areas. The operating noise of down-drainage washing machines is less than that of up-drainage washers. However, down-drainage washing machines sometimes fail to adequately clean the laundry. In general, most users are more satisfied with down drainage, but the up-drainage models also have advantages.

Similarly, reviews that comment on frequency conversion were researched. The results can be found in Tables B3 and B4 in the Appendix. Nearly all the phrases that describe washing machines with integrated frequency conversion were positive, while most phrases relate to non-frequency-conversion washing machines were negative. Our analysis demonstrates that customers were more satisfied with the frequency conversion designs, distinguished by low noise emission, simplified operation, consistently clean laundry, power efficiency, and other benefits. The non-frequency-conversion washing machines are noisier and do allow clothes to be added during wash cycles.

6. Conclusion and future work

When a product is in design, how planned attributes may affect customer satisfaction should be considered. This study presented a logistic regression model that offered strong indications of how satisfied customers will be with different product attributes. In analysing customer comments on washing machine purchases, we discovered that customer satisfaction is sensitive to a design's drainage mode, loading type, frequency conversion, type, display, color and capacity. Designers, thus, should prioritize these attributes when weighing the costs of manufacture against consumers' strong feelings towards key features. Other optional features, such as construction material and energy efficiency class, showed no statistically significant impacts on customer satisfaction. Such variables ought to give an appliance design team less concern.

The effects of price were also observed. The regression results for different price groups showed that consumers with strong budget concerns demonstratde different evaluative priorities than prospective purchasers of medium- and high-priced units. In addition, customer satisfaction relating to some attributes, such as noise, largely depended on the individual customer's perception. The results show that some owners perceived the noise level of low-noise operating washing machines as excessively loud. Manufacturers should reconsider whether strategies that prioritize performance improvements in attributes that elicit widely varying subjective reactions are sound.

A number of researchers (e.g., Song et al., 2016; Jiang and Rosenbloom, 2005; Qu et al., 2008) have reported that logistics services and retailer visibility play important roles in securing customers' satisfaction. Many reviews referred to the quality of logistics services and retailer support. Thus, future research needs to investigate how and to what extent these factors affect customer satisfaction. If optional features can have enormous impact on sales and brand loyalty, the psychological effects of many other product features must be controlled for.

A few attributes, such as drying function, are not included in our model. Nevertheless, manufacturers should not downplay the importance of the critical features that prominently figured in customers' assessments of their needs. Designers should also take note. New studies should be directed towards fine-tuning additional tests of how customers' personal attributes influence purchase satisfaction. However, the data gathering and estimation challenges associated with such detailed modelling are formidable.

Future research also should look at how more effective design strategies may enhance the ownership experiences of different customers. To this end, personal information about customers, such as occupation and income, should be collected. It will be particularly desirable to investigate how design options attract high-income customers, who, with little concern for price, may invest in more advanced features and better services. More detailed analyses of high-end, "intelligent" functions are required as a reference for improving many product lines.

The modelling approach and estimation technique utilized in our study can be generalized to improve the design and manufacturing of other products, especially complex ones. Complex products are often marketed as having diverse arrays of design features. Manufacturers will face challenges in deciding which attributes to improve and include. In addition, consumers may have much higher expectations of complex products.

Acknowledgments

The contact author acknowledges the Natural Science Foundation of China under Grant Nos. 71522014, 71771213 and 71690233. His co-authors were partially supported by the Natural Science Foundation of China under Grant Nos. 71731009, 71628103 and 71725001, and wish to acknowledge the support with thanks.

Appendix A. Logistic regression model

See Tables A1–A4.

Table A1
Effect of attributes on customer satisfaction.

Variable		Coef	SE	p-value	Standardized Estimate
Intercept		1.823	1.1405	0.110	0.162
Price		0.0002	0.00002	<0.001	−0.009
Brand	Bosch	−0.224	0.1708	0.190	0.055
Brand	Blomberg	0.638	0.1127	<0.001	0.012
Brand	Camel	0.209	0.1124	0.063	−0.029
Brand	Casarte	−0.677	0.1657	<0.001	−0.077
Brand	Daewoo	−1.716	1.1295	0.129	0.031
Brand	DeMuller	0.696	0.1580	<0.001	0.014
Brand	Hap	0.505	0.2201	0.022	−0.055
Brand	Haier	−0.261	0.0781	<0.001	0.030
Brand	Hisense	0.531	0.1339	<0.001	0.010
Brand	Keg	0.128	0.0999	0.120	−0.014
Brand	Konka	−0.178	0.0987	0.071	−0.028
Brand	LG	−0.316	0.1024	0.002	0.017
Brand	LittleDuck	−1.295	1.1737	0.270	−0.045
Brand	LittleSwan	−0.251	0.0830	0.003	0.004
Brand	Midea	0.025	0.0833	0.766	0.026
Brand	Oping	0.342	0.1484	0.021	−0.017
Brand	Panasonic	−0.134	0.0913	0.143	0.002
Brand	Patches	0.039	0.1806	0.828	−0.076
Brand	Ripu	−2.285	1.1510	0.047	0.012
Brand	Sakura	0.164	0.1212	0.177	−0.003
Brand	Samsung	−0.023	0.0994	0.814	−0.067
Brand	Sanyo	−0.354	0.0799	<0.001	−0.059
Brand	Siemens	−0.514	0.1011	<0.001	0.063
Brand	Snowflk	1.435	0.2526	<0.001	0.022
Brand	Skyworth	0.300	0.1121	0.008	0.051
Brand	Weili	0.504	0.0944	<0.001	−0.009
Brand	Whirlpool	−0.071	0.0850	0.406	0.057
Brand	Zanussi	1.305	0.2172	<0.001	0.030
Brand	Beko	0.674	0.1543	<0.001	0.018
Brand	Sevenstars	0.410	0.1576	0.009	
Automatic Type	Semi-auto	0			−0.018
Control Mode	Computer	−0.152	0.1307	0.244	0.279
Type	Pulsator	1.012	0.2637	<0.001	0.036
Type	Drum	0.130	0.0705	0.066	−0.034
Color	White	−0.143	0.0334	<0.001	−0.050
Color	Grey	−0.196	0.0340	<0.001	0.018
Color	Golden	0.095	0.0357	0.008	−0.017
Material	Stainless Steel	0			0.059
Drainage Mode	Up Drainage	−0.106	0.0350	0.002	−0.026
Loading Type	Top Loading	−0.702	0.2234	0.002	−0.193
Family Type	One	0.107	0.0742	0.149	0.010
Family Type	Two	0.107	0.0533	0.044	0.015
Family Type	Three	−0.062	0.0276	0.025	−0.017
Capacity		0.081	0.0190	<0.001	0.0571
Energy Effic Class	1	−1.235	1.1412	0.279	−0.340
Energy Effic Class	2	−1.0755	1.1418	0.346	−0.269
Energy Effic Class	3	−0.775	1.1412	0.497	−0.181
Energy Effic Class	4	0			
Noise		−0.002	0.0013	0.113	−0.011
Weight		−0.008	0.0036	0.035	−0.081
Freq Conversion	Yes	0.217	0.0651	<0.001	0.060
Auto Power Off	No	0.055	0.0404	0.172	0.010
Display Screen	LCD	0.141	0.0563	0.013	0.022
Display Screen	LED	0.050	0.0404	0.216	0.011
Water Consump		−0.0001	0.00012	0.224	−0.006

Table A2
Effect of attributes on customer satisfaction in low-priced group.

Variable		Coef	SE	p-value	Standardized Estimate
Intercept		0.812	0.7574	0.284	−0.186
Price		−0.002	0.0007	<0.001	−0.094
Brand	Camel	0.012	0.1499	0.937	−0.046
Brand	DeMuller	0.342	0.2225	0.124	−0.002
Brand	Hap	0.922	0.3029	0.002	−0.093
Brand	Haier	0.207	0.2279	0.365	−0.093
Brand	Hisense	−0.513	0.2455	0.037	−0.048
Brand	Keg	0.586	0.1989	0.003	−0.118
Brand	Konka	0.093	0.2032	0.647	0.140

(continued on next page)

Table A2 (*continued*)

Variable		Coef	SE	*p*-value	Standardized Estimate
Brand	LittleDuck	1.642	0.5306	0.002	−0.231
Brand	LittleSwan	−0.725	0.2229	0.001	−0.140
Brand	Midea	0.089	0.1507	0.554	0.051
Brand	Oping	1.342	0.3969	<0.001	−0.008
Brand	Patches	0.838	0.4574	0.067	−0.134
Brand	Ripu	−1.342	0.5364	0.012	−0.061
Brand	Sakura	0.450	0.2847	0.114	−0.218
Brand	Sanyo	−0.141	0.1585	0.374	0.095
Brand	Snowflk	2.189	0.4970	<0.001	−0.162
Brand	Skyworth	−0.632	0.4052	0.119	−0.129
Brand	Weili	0.656	0.1616	<0.001	−0.051
Brand	Whirlpool	−0.640	0.1997	<0.001	−0.151
Brand	Sevenstars	0.952	0.5794	0.100	
Automatic Type	Semi-auto	0			0.061
Control Mode	Computer	0.305	0.2496	0.222	−0.214
Color	White	−1.051	0.2910	<0.001	−0.208
Color	Grey	−0.780	0.3511	0.026	−0.057
Color	Golden	−0.745	0.3810	0.051	−0.198
Material	Stainless Steel	0			0.101
Drainage Mode	Up Drainage	0			
Loading Type	Top Loading	−0.158	0.1503	0.294	−0.020
Family Type	One	0.121	0.1274	0.341	0.027
Family Type	Two	−0.295	0.0819	<0.001	−0.081
Family Type	Three	0.138	0.0965	0.154	0.077
Capacity		−0.432	0.0975	<0.001	−0.119
Energy Effic Class	1	0			
Energy Effic Class	2	0			
Energy Effic Class		−0.006	0.0035	0.088	−0.039
Energy Effic Class		0.056	0.0194	0.004	0.144
Auto Power Off	No	0.056	0.4959	0.910	0.0080
Display Screen	LCD	0.138	0.1884	0.464	0.015
Display Screen	LED	0.204	0.0672	0.002	0.049
Water Consump		0.006	0.0019	<0.001	0.120

Table A3
Effect of attributes on customer satisfaction in medium-priced group.

Variable		Coef	SE	*p*-value	Standardized Estimate
Intercept		−0.462	0.5729	0.420	0.081
Price		0.00025	0.00005	<0.001	−0.050
Brand	Blomberg	1.515	0.4298	0.000	−0.049
Brand	Daewoo	−0.412	0.5069	0.416	−0.037
Brand	DeMuller	−0.099	0.5110	0.847	−0.359
Brand	Haier	0.552	0.4102	0.179	−0.050
Brand	Hisense	1.286	0.4397	0.004	−0.166
Brand	LG	0.505	0.4174	0.226	−0.264
Brand	LittleSwan	0.677	0.4173	0.105	−0.195
Brand	Midea	0.829	0.4212	0.049	−0.197
Brand	Panasonic	0.498	0.4229	0.239	−0.108
Brand	Samsung	1.022	0.4170	0.014	−0.355
Brand	Sanyo	0.236	0.4202	0.574	−0.144
Brand	Siemens	0.490	0.4261	0.251	−0.044
Brand	Skyworth	1.380	0.4317	<0.001	−0.149
Brand	Whirlpool	0.727	0.4201	0.084	−0.166
Brand	Zanussi	1.903	0.4563	<0.001	−0.009
Brand	Beko	2.031	0.4352	<0.001	0.183
Type	Pulsator	0.722	0.3023	0.017	−0.061
Color	White	−0.238	0.0498	<0.001	−0.020
Color	Grey	−0.087	0.0633	0.169	0.008
Color	Golden	0.047	0.0583	0.420	0.046
Drainage Mode	Up Drainage	−0.273	0.0620	<0.001	−0.070
Loading Type	Top Loading	−0.553	0.2758	0.045	−0.139
Family Type	One	0.903	0.2763	0.001	0.052
Family Type	Two	0.651	0.1362	<0.001	0.056
Family Type	Three	0.074	0.0507	0.1445	0.019
Capacity		0.080	0.0349	0.023	0.047
Energy Effic Class	1	−0.566	0.1387	<0.001	0.273
Energy Effic Class	2	−0.250	0.0908	0.007	0.337
Energy Effic Class	3	0			0.304
Noise		0.012	0.00317	<0.001	0.045
Weight		−0.011	0.00560	0.050	−0.095
Freq Conversion	Yes	0.314	0.1028	0.002	0.086
Auto Power Off	No	−0.079	0.0574	0.171	−0.015
Display Screen	LCD	0.095	0.1797	0.595	0.012
Display Screen	LED	0.060	0.1612	0.712	0.009
Water Consump		−0.0003	0.00013	0.052	−0.015

Table A4
Effect of Attributes on Customer Satisfaction in high-priced group.

Variable		Coef	SE	p-value	Standardized Estimate
Intercept		0.229	0.5389	0.671	0.030
Price		0.00005	0.00004	0.188	−0.012
Brand	Bosch	−0.149	0.2087	0.477	0.0028
Brand	Blomberg	0.024	0.1805	0.892	−0.064
Brand	Casarte	−0.760	0.2477	0.002	0.049
Brand	Daewoo	0.735	0.5240	0.161	−0.102
Brand	Haier	−0.480	0.1280	0.000	−0.034
Brand	LG	−0.315	0.1773	0.076	−0.016
Brand	LittleSwan	−0.081	0.1453	0.577	−0.112
Brand	Panasonic	−0.630	0.1628	0.000	−0.022
Brand	Samsung	−0.145	0.1523	0.340	−0.032
Brand	Sanyo	−0.200	0.1113	0.072	−0.112
Brand	Siemens	−0.609	0.1565	0.000	0.126
Type	Pulsator	0.640	0.4018	0.111	0.023
Type	Drum	0.094	0.0873	0.282	−0.071
Color	White	−0.323	0.0957	0.001	−0.029
Color	Grey	−0.276	0.1488	0.064	−0.039
Color	Golden	−0.149	0.0870	0.086	−0.018
Drainage Mode	Up Drainage	0.112	0.1170	0.341	0.031
Loading Type	Top Loading	0			
Family Type	One	1.195	0.3269	0.000	0.111
Family Type	Three	−0.003	0.0651	0.968	−0.001
Capacity		0.046	0.0608	0.447	0.029
Energy Effic Class	1	−0.084	0.1429	0.557	0.022
Energy Effic Class	2	0			0.028
Energy Effic Class	3	0			v
Noise		0.011	0.0035	0.002	0.056
Weight		0.011	0.0096	0.258	0.079
Freq Conversion	Yes	0			
Auto Power Off	No	−0.058	0.0763	0.446	−0.013
Display Screen	LCD	−0.319	0.1452	0.028	−0.069
Display Screen	LED	−0.374	0.1276	0.003	−0.090
Water Consump		0.0001	0.0045	0.981	0.001

Appendix B. LDA Model

See Tables B1–B4.

Table B1
LDA results for the reviews about up-drainage.

Class 1	Class 2	Class 3
Speechless	Not bad	Aha
It's not used	Drain off water	Good
Keep peddling shelves and drains	Draining of the water	Price and performance are OK
Complaint	I don't know how to use it	Recommend
I would like to contact the manufacturer for a replacement	Just bought it back like this	Praise the installation master
The machine drains off water while we add water.	Depressed	The master installed it patiently
What I want to buy is a washing machine, not a filter	Just so so	Teach me how to use
Washing powder box and emergency drainage do not match	The drain was completely put on the ground when it was first used	Drain pipe
I wonder if it's debugging or old	Wash clothes while draining	Inlet pipe
There is no present	I called customer service phone before I knew it needs to be hung up	I can pace the power line reasonably

Table B2
LDA results for the reviews about down-drainage.

Class 1	Class 2	Class 3
Bad	Simple operation	General appearance
Not bad	Low noise	Undrained
It is not bad	High performance price ratio	It has not been used
All right	Enough power	Depressed
It's very good	washing machine is good	while draining
Overall it is OK	Attractive appearance	Material is not very good
Basically no problem after using a period of time	Worth buying	Very good
It is quite enough for the couple	Large capacity	It's very good to use
The drain can be adjusted to the left or right	Water saving and electricity saving	The pipe is a little short
Praise	Practical function	The other is good

Table B3
LDA results for the reviews about frequency conversion.

Class 1	Class 2	Class 3
Variable-frequency motor	Large capacity	It can wash clothes clean
High performance price ratio	Simple operation	Major brand
Good	Low noise	Very satisfied
Low noise	The washing machine is good	Very good
It is quiet	Practical function	It is good to use
Frequency conversion	Enough power	Good
Large capacity	Attractive appearance	The service is good
It is variable-frequency	High performance price ratio	Ultra-large capacity
It can wash clothes clean	Variable frequency and mute	Variable frequency and power saving
Variable-frequency	Water saving and electricity saving	The sound of frequency conversion washing machine is very small

Table B4
LDA results for the reviews about non-frequency conversion.

Class 1	Class 2	Class 3
we should buy a frequency conversion washing machine	High performance price ratio	Go to the store to buy
The machine trembled when the master installed it	I am a little bit of regret	I want buy a frequency conversion washing machine
The tray cost 185	Awesome logistics	I was told it is frequency conversion
The machine doesn't tremble	Exhibits and photos	The goods arrived
Clothes can not be added when the fixed frequency laundry is working	It is ok	it's an ordinary motor
The functions are enough	Inadequate	New machine
Follow-up evaluation	No frequency conversion	The noise
The washing machine is received	Not 1 energy saving	The sound of dehydration is high
The noise is a bit high	we should buy frequency conversion	It is bad
There is no after-sales installation after receipt of the goods	I should choose the frequency conversion	A little regret

References

Abrahams, A., Fan, W., Wang, G., Zhang, Z., Jiao, J., 2015. An integrated text analytic framework for product defect discovery. Prod. Oper. Manag. 24, 975–990.

Abrahams, A., Jiao, J., Fan, W., Wang, G., Zhang, Z., 2013. What's buzzing in the blizzard of buzz? Automotive component isolation in social media postings. Decis. Support. Syst. 55, 871–882.

Abrahams, A., Jiao, J., Wang, G., Fan, W., 2012. Vehicle defect discovery from social media. Decis. Support. Syst. 54, 87–97.

Blei, D., Lafferty, J., 2007. A correlated topic model of science. Annals Appl. Stat. 1, 17–35.

Blei, D., Ng, A., Jordan, M., 2003. Latent Dirichlet allocation. J. Machine Learn. Res. 3, 993–1022.

Chang, M.R., Kauffman, R.J., Kwon, Y.O., 2014. Understanding the paradigm shift to computational social science in the presence of big data. Decision Support Systems 63, 67–80.

Decker, R., Trusov, M., 2010. Estimating aggregate consumer preferences from online product reviews. Int. J. Res. Mark. 27, 293–307.

Franke, N., Piller, F., 2003. Key research issues in user interaction with user toolkits in a mass customisation system. Int. J. Technol. Manag. 26, 578–599.

Goorha, S., Ungar, L., 2010. Discovery of significant emerging trends. In: Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM Press, New York, pp. 57–64.

Groves, R., 2006. Nonresponse rates and nonresponse bias in household surveys. Public Opin. Quart., 646–675

Guo, W., Liang, R., Wang, L., Peng, W., 2017. Exploring sustained participation in firm-hosted communities in China: the effects of social capital and active degree. Behav. & Inf. Technol. 36, 223–242.

Hassan, A., Radev, D., 2010. Identifying text polarity using random walks. In: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, Stroudsburg, PA, pp. 395–403.

Hedegaard, S., Simonsen, J., 2013. Extracting usability and user experience information from online user reviews. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM Press, New York, pp. 2089–2098.

Hu, M., Liu, B., 2004. Mining and summarizing customer reviews. In: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM Press, New York, pp. 168–177.

Jiang, P., Rosenbloom, B., 2005. Customer intention to return online: price perception, attribute-level performance, and satisfaction unfolding over time. Eur. J. Mark. 39, 150–174.

Jin, J., Ji, P., Kwong, C., 2016. What makes consumers unsatisfied with your products: review analysis at a fine-grained level. Eng. Appl. Artif. Intell. 47, 38–48.

Jin, J., Ji, P., Liu, Y., 2012. Product characteristic weighting for designer from online reviews: An ordinal classification approach. In: Proceedings of the 2012 Joint EDBT/ICDT Workshops. ACM Press, New York, pp. 33–40.

Jin, J., Ji, P., Liu, Y., 2014. Prioritising engineering characteristics based on customer online reviews for quality function deployment. J. Eng. Des. 25, 303–324.

Jo, Y., Oh, A., 2011. Aspect and sentiment unification model for online review analysis. In: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining. ACM Press, New York, pp. 815–824.

Kauffman, R.J., Kim, K., Lee, S.Y.T., Hoang, A.P., Ren, J., 2017. Combining machine-based on econometric methods for policy analytics insights. Electr. Commerce Res. Appl. 15, 115–140.

Law, D., Gruss, R., Abrahams, A., 2017. Automated defect discovery for dishwasher appliances from online consumer reviews. Expert Syst. Appl. 67, 84–94.

Lee, T., 2007. Needs-based analysis of online customer reviews. In: Proceedings of the Ninth International Conference on Electronic Commerce. ACM Press, New York, pp. 311–318.

Lin, C., He, Y., 2009. Joint sentiment/topic model for sentiment analysis. In: Proceedings of the 18th ACM Conference on Information and Knowledge Management. ACM Press, New York, pp. 375–384.

Liu, B., Hu, M., Cheng, J., 2005. Opinion observer: Analyzing and comparing opinions on the web. In: Proceedings of the 14th International Conference on World Wide Web. ACM Press, New York, pp. 342–351.

Mullen, T., Collier, N., 2004. Sentiment analysis using support vector machines with diverse information sources. Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, pp. 412–418.

Nakagawa, T., Inui, K., Kurohashi, S., 2010. Dependency tree-based sentiment classification using CRFs with hidden variables. In: Proceedings of Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, Stroudsburg, PA, pp. 786–794.

Qiu, G., Liu, B., Bu, J., Chen, C., 2009. Expanding domain sentiment lexicon through double propagation. In Proceedings of the Internatoin Joint Conference on Artificial Intelligence, pp. 1199–1204.

Qu, Z., Zhang, H., Li, H., 2008. Determinants of online merchant rating: content analysis of consumer comments about Yahoo merchants. Decision Support Systems 46, 440–449.

Song, G., Zhan, Y., Guo, Y., 2016. The effectiveness of online shopping characteristics and logistics service on satisfaction. In: Proceedings of the 2016 13th IEEE International Conference Service Systems and Service Management. IEEE Computer Society Press, Washington, DC, pp. 1–6.

Tucker, C., Kim, H., 2011. Predicting emerging product design trend by mining publicly available customer review data. In: Proceedings of the 18th International Conference on Engineering Design, Impacting Society through Engineering Design. Copenhagen, Denmark, pp. 15–19.

Wang, L., Youn, B., Azarm, S., Kannan, P., 2011. Customer-driven product design selection using web based user-generated content. In: ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers, New York, pp. 405–419.

Wei, X., Croft, W., 2006. LDA-based document models for ad-hoc retrieval. In: Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM Press, New York, pp. 178–185.

Wiebe, J., Riloff, E., 2005. Creating subjective and objective sentence classifiers from unannotated texts. In: Proceedings of International Conference on Intelligent Text Processing and Computational Linguistics. Springer, Berlin Heidelberg, pp. 486–497.

Wilson, T., Wiebe, J., Hoffmann, P., 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In: Proceedings of the conference on human language technology and empirical methods in natural language processing. Association for Computational Linguistics, Stroudsburg, PA, pp. 347–354.

Xia, R., Zong, C., Hu, X., Cambria, E., 2013. Feature ensemble plus sample selection: domain adaptation for sentiment classification. IEEE Intell. Syst. 28, 10–18.

Zagibalov, T., Carroll, J., 2008. Automatic seed word selection for unsupervised sentiment classification of Chinese text. In: Proceedings of the 22nd International Conference on Computational Linguistics. Association for Computational Linguistics, Stroudsburg, PA, pp. 1073–1080.

Zhou, Q., Xia, R., Zhang, C., 2016. Online shopping behavior study based on multi-granularity opinion mining: China versus America. Cogn. Comput. 8, 587–602.

Zhu, F., Zhang, X., 2010. Impact of online consumer reviews on sales: the moderating role of product and consumer characteristics. J. Market. 74, 133–148.