ABSTRACT

Title of Document: PRODUCT DESIGN SELECTION USING ONLINE CUSTOMER REVIEWS

Lulu Wang, Doctor of Philosophy, 2011

Directed By: Dr. Byeng D. Youn
School of Mechanical and Aerospace Engineering

Dr. Shapour Azarm
Department of Mechanical Engineering

Dr. P. K. Kannan
Department of Marketing
R. H. Smith School of Business

Product design selection is heavily constrained by its customer preference data acquisition process. Traditionally, the customer preference data is collected through survey-based methods such as conjoint; sometimes product prototypes are generated and evaluated by focused groups of customers. In this way, the data acquisition process can become costly and require a significant amount of time.

The goal of this dissertation is to overcome the limitation of the traditional customer preference data acquisition process by making use of a new type of customer data – online customer reviews. Because online customer reviews are, to a large extent, freely available on the Internet copiously, using them for product design can significantly reduce the cost as well as the time. Of course, the data obtained from online reviews have some disadvantages too. For example, online reviews are freely expressed and can contain a lot of noise.

In this dissertation, a new methodology is developed to extract useful data from online customer reviews from a single website, construct customer preference models and select a product design that provides a maximum expected profit. However, online customer reviews from a single website may not represent the market well. Furthermore, different websites may have their own procedures and formats to acquire customer reviews. A new approach is developed to systematically elicit customer data from multiple websites, construct customer preference models by considering website heterogeneity, and select a product design. The model from multiple websites is also extended to account for customer preference heterogeneity. The models obtained from the online customer reviews for single and multiple websites are compared and validated using a set of out-of-sample data. To demonstrate the applicability of the proposed models, a smartphone case study is used throughout the dissertation.

PRODUCT DESIGN SELECTION
USING ONLINE CUSTOMER REVIEWS

By

Lulu Wang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:

Assistant Professor	Byeng D. Youn, Chair
Professor	Shapour Azarm, Co-Advisor
Professor	P.K. Kannan, Co-Advisor
Associate Professor	Linda Schmidt
Associate Professor	Janet Wagner (Dean's Representative)

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere appreciation to my advisors, Dr. Byeng D. Youn, Dr. Shapour Azarm and Dr. P.K. Kannan, whose insightful guidance, and tremendous patience and dedication to my research made the pursuit of my degree successful. It is my great honor and fortune to have an opportunity to learn from them – the rigorous scientific attitude, the hard-working spirit as well as the admirable personality. What I learned from them will be extremely beneficial for my future career development.

I would also like to thank Dr. Linda Schmidt, Dr. Janet Wagner for serving on my dissertation committee. Their invaluable comments and suggestions have significantly improved my dissertation.

Special thanks go to my colleagues at General Motors – Ms. Artemis Kloess, Dr. Peter Fenyes, Dr. Mark Beltramo, Dr. John Cafeo, Dr. Stacey Gu and Mr. Joseph Donndelinger for their kind encouragements, help and guidance during my internship there. The cooperation with them enlarged my research knowledge and improved my practical skills. I am also thankful to my colleagues and friends at UMD – Dr. Pingfeng Wang, Dr. Zhimin Xi, Mr. Zhichao Wang, Dr. Chao Hu, Dr. Byungchang Jung, and Dr. Soobum Lee, for their friendly support over the past years.

Finally, I would like to thank my parents for their continuous encouragements and support. Their love helps me pass through the most difficult time during my study. More importantly, I am grateful to my husband, Hua, whose love and understanding always give me the power to overcome every single challenge in my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AIC	=	Akaike information criterion
CAIC	=	Consistent AIC
CI	=	Confidence interval
DIC	=	Deviance information criterion
MAE	=	Mean absolute error
MAIC	=	Marginal AIC
MAPE	=	Mean absolute percentage error
MCMC	=	Markov Chain Monte Carlo
NRMSE	=	Normalized root mean squared error
OS	=	Operating system
QFD	=	Quality function deployment
RMSE	=	Root mean squared error
SVM	=	Support vector machine
UGC	=	User-generated content

NOMENCLATURE

a,c	=	Constants in the relationship between P and R
b	=	Offset parameter of the desired surface in SVM
C_i	=	Cost of the i^{th} design alternative
D_i	=	Demand for alternative i

e	=	Customer random error term
$\mathbf{E}$	=	Product specification vector
f	=	Probability density function
G	=	Number of parameters in a model
h_r	=	Height resolution in pixel
i	=	Design alternative index (i=1,…,I)
J	=	Number of observations
j	=	Observation index (j=1,…,J)
k	=	Training dataset index (k=1,…,K)
K	=	Number of training datasets
L	=	Likelihood function
l	=	Website index (l=1,…,Q)
M	=	Number of segments
MS	=	Market size
N	=	Total number of customers
n	=	Customer index (n=1,…, N)
P	=	Probability of purchase
PC	=	Product price
PF	=	Expected profit
Q	=	Total number of websites
R	=	Product rating
RK	=	Rank of expected profit
s	=	Segment number (s = 1,…, M)

$\mathbf{S}_k$	=	The k^{th} training dataset
t	=	Observation index ($t = 1,\ldots, T_n$)
T_n	=	Total number of observations from Customer n
u	=	Class output
$\mathbf{v}$	=	Input vector
w_r	=	Width resolution in pixel
w_s	=	Weight for Segment s
$\mathbf{W}$	=	SVM weight vector
$\mathbf{X}$	=	Attribute aggregate rating vector
$\mathbf{Y}$	=	Attribute rating vector
y	=	Observed ratings
z	=	Dependent variable of a regression model
α	=	Random effects of attribute aggregate ratings on product ratings between the websites
β	=	Customer preference model parameter
γ	=	Random effects of attribute ratings on attribute aggregate ratings between the websites
θ	=	Random effects of product specifications on attribute ratings between the websites
σ^2	=	The variation of customer preference model error
τ^2	=	Between-website variance
ε	=	Customer preference model error
ς	=	Customer-specific error term

η	=	Location shift of scale usage
λ	=	Scale shift of scale usage
$\boldsymbol{\xi}$	=	Bayesian model parameters
$\mathbf{\Lambda}$	=	Independent variables of a regression model
ϕ	=	Probability of a binary variable equal to 1
$\sum$	=	Covariance matrix of multivariate normal distribution
ρ	=	Correlation between two variables

Chapter 1: Introduction

The goal in this dissertation is to develop a new methodology that will make use of online customer reviews for product design selection. Product design selection aims at selecting a best design which satisfies customer needs and preferences; thus increasing profit for a firm. However, customer data required for design selection are traditionally collected by way of a time- and cost-consuming customer survey process. This has motivated the research in this dissertation to explore for an alternative source of customer data, that is, online customer reviews. Online customer reviews are considered to be "the voice of customer in the 21st Century" [1]. By making use of online customer reviews for product design, it is possible to overcome the above mentioned limitations of the traditional survey-based customer-driven product design. This proposed research (using online customer reviews for product design selection) is applicable to all customer durable products, such as a mobile phone, an appliance, a power tool, or an automobile. However, it is not an easy task to utilize online reviews since these reviews are not customer data targeted and thus can include a certain amount of noisy content that is irrelevant to customer preferences.

The goal for the dissertation is accomplished by way of three research thrusts. In Research Thrust 1, online reviews from a single website is processed and modeled for design selection. However, selecting reviews from a single website may not represent the entire market. Therefore, reviews from multiple websites are considered as well. However, integrating customer reviews from multiple websites exacerbates the complexity of the problem. This is because different websites have their own procedures and formats for customers to input the reviews. Such differences are called website heterogeneity. Online customer reviews from several websites are processed and integrated for selection in Research Thrust 2. In Research Thrust 3, an extension of the model in Research Thrust 2 is made by taking into account customer preference heterogeneity as well. These models are compared and validated through a set of out-of-sample data points.

1.1 Motivation and Objective

Two or three decades ago, the product design process was purely an engineering design process. Engineers sought product design which had the best performance – reliability, durability and so on. However, this engineering-centric design process can be highly risky as it ignored taking into account customer needs and preferences. Researchers gradually realized that involving the voice of the customer at the early stage of product design could mitigate the risk of an unsuccessful design – that is, failing in the market. This type of design process is called a customer-driven product design. Customer-driven product design has become an important research topic as evidenced by extant research on this topic [2-7]. Customer-driven product design enables a company to better understand customer needs, select a successful product design according to the market needs, and ultimately improve the financial well-being of the company. However, selecting an appropriate design relies heavily on acquiring appropriate customer data and building a related knowledge set. The acquisition of customer data is a difficult process and often constrained by limited resources, such as budgetary and time constraints. This difficulty has prevented the development of a reliable customer-driven product design.

The overall objective of this dissertation is to propose and develop a new methodology to overcome the limitations of traditional customer-generated data acquisition and processing for product design selection by making use of online customer reviews.

With the development of Web 2.0 technology, web users have been able interact with each other and share their opinions about a product on the web. The term User Generated Content (UGC) has emerged into the mainstream since then. UGC refers to a range of media content such as discussion boards, blogs, wikis, online customer reviews, social networking content, and so on. The focus in this dissertation is one type of UGC – that is, online customer reviews. However, utilizing online customer reviews for product design selection is not a trivial task because the reviews are freely expressed and written by customers without any constraints, structure and bounds. While this freedom can eliminate the response biases imposed on customers by conventional survey techniques, it also brings in the difficulties to process customer reviews which can contain a lot of noise, variability and even bias induced by their writers. In order to make use of these reviews for product design selection, the following research questions are explored in this dissertation: *How should online customer reviews be processed properly? How should these reviews be used in constructing preference models? How should the reviews be used for design selection?*

Although UGC in general and online customer reviews in particular have raised a wide range of interests, there are concerns in using them for design selection. One concern is the sample size. A single website might be characterized by a certain group of customers who visit the site. The information from a single website may be biased by that small group of customers who may not represent the customer preferences as a whole. One possible solution is to combine customer preference information from multiple websites. In this way, online reviews from multiple websites can provide heterogeneous customer data and, therefore, have to be integrated carefully since the data from different websites do not have the same format and thus cannot be simply aggregated. Simple combination of data from multiple websites cannot guarantee the accuracy of customer preference models. This leads to another research question that is investigated in this dissertation: *How should customer data from multiple websites be combined by considering the differences among the websites?*

This dissertation seeks to shed some light on all of the above mentioned questions in the context of product design selection.

1.2 Research Thrusts

To achieve the overall objective, three research thrusts are included as presented in Figure 1.1. Research Thrust 1 focuses on a single website. In Research Thrust 2, reviews from multiple websites are integrated and processed for design selection. In Research Thrust 3, two of the assumptions from Research Thrust 2 are relaxed and the models in Thrusts 2 and 3 are compared and validated.

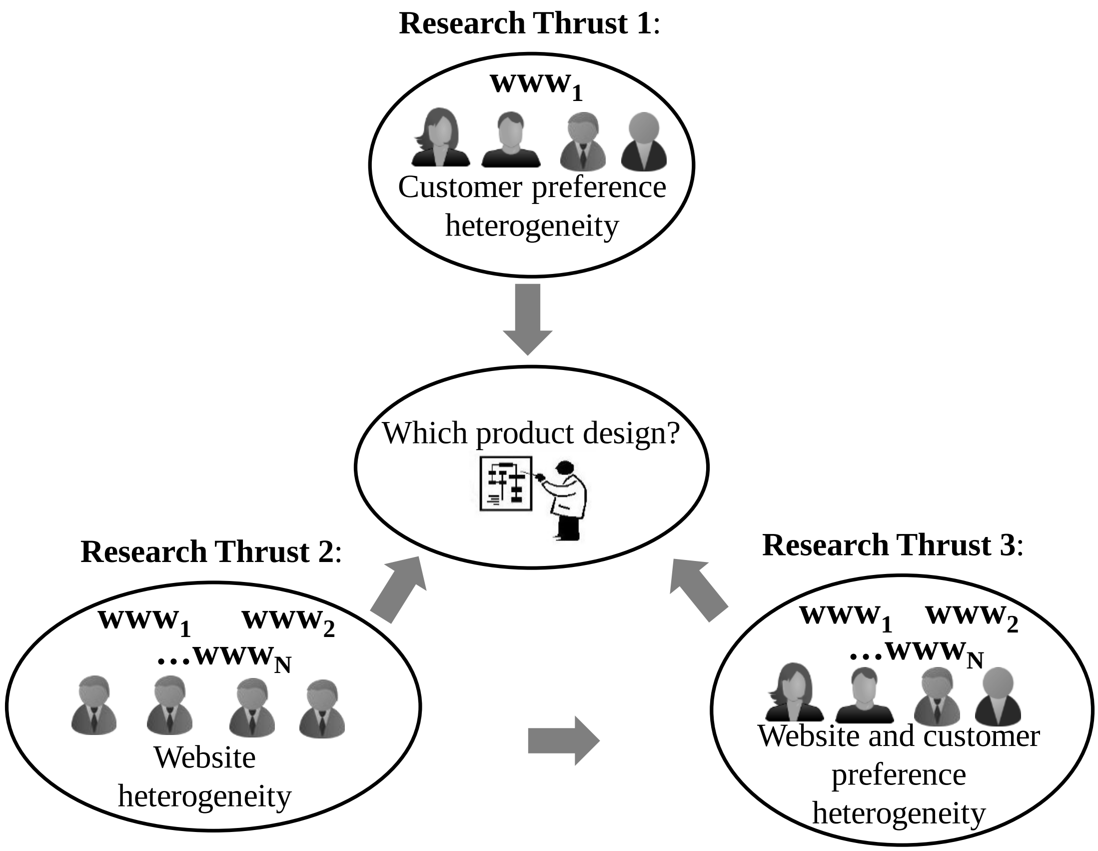


Figure 1.1 Thrusts Structure

1.2.1 Research Thrust 1: Design Selection with Online Customer Reviews from a Single Website

The main assumption here is that online customer reviews from a single website are good representation of the market. However, this assumption is relaxed in subsequent chapters. The purpose of this approach is to elicit meaningful customer data from online reviews, construct customer preference models and use the models for product design selection.

1.2.2 Research Thrust 2: Design Selection with Online Customer Reviews from Multiple Websites with Website Heterogeneity

It may not be possible that a single website can represent customer opinions for the entire market. Thus in this research thrust, online customer reviews from multiple websites are integrated and modeled for product design. The goal of this research thrust is to present a systematic methodology to elicit customer data from multiple websites, integrate the data from multiple websites and use the integrated customer preference model for product design.

1.2.3 Research Thrust 3: Design Selection with Online Customer Reviews from Multiple Websites with Website and Customer Preference Heterogeneity

In Research Thrust 2 the online customer reviews from multiple websites have been processed, integrated and used for design selection under several strong assumptions. Research Thrust 3 relaxes two such assumptions (imposed as part of Research Thrust 2): The first assumption is that there is no heterogeneity in customer preferences and the second assumption is that the multiple responses from a single customer are independent. These two assumptions, as considered in Research Thrust 2, may not be realistic because customers may have different tastes and preferences for a product. For a single customer, his or her different responses may be correlated because the statistical errors of the responses are caused by his or her biases due to inherent habits and backgrounds. Therefore, in Research Thrust 3, several models are proposed – one model considering both website and customer preference heterogeneity and the others considering the two types of heterogeneity plus the error correlations. The proposed models along with the model in research Thrust 2 are compared and validated using a set of out-of-sample data. The out-of-sample data are extracted from three websites which were not used for the model development.

1.3 Assumptions

In utilizing online customer reviews for product design selection, several assumptions are made:

- A mature product category is considered, one that already exists in the market. The overall objective is to improve the design of an existing product rather than to develop a new product;
- Product design alternatives have already been generated and the goal here is to make a selection out of these alternatives;
- Customer’s purchase decision (the probability of purchase) is assumed to be a function of product ratings. The function is assumed to be of linear, binary logit and exponential forms;
- Demand is linearly dependent on the probability of customer purchase decision;
- The online customer reviews are assumed to be written only by regular customers, the end users of the product. The effects from professional technicians/reviewers and active managements on online customer reviews are ignored. In practice, it is possible that the reviews are provided by professional technicians or reviewers who work for the firm. One of the main service providers for the management of online customer reviews are Bazaarvoice [10].
- In online customer reviews, freely written pros/cons are basically a summary

of the general comments.

- Customer preferences for a product are assumed to be unchanged in a certain time period, for example, one year or so for the smartphone. Under this assumption, online customer reviews can be downloaded during that time period and used for constructing the customer preference models.

There are several assumptions made specific for each research thrust.

- Assumption for Research Thrust 1:
    - Online customer reviews from a single website can represent the whole market (relaxed in Thrust 2);
- Assumptions for Research Thrust 2:
    - Customer reviews from multiple websites are a good representative of customer voices from the whole market;
    - No customer preference heterogeneity considered (relaxed in Research Thrust 3);
    - Statistical independence of multiple responses from a single customer is considered (relaxed in Research Thrust 3);
- Assumption for Research Thrust 3:
    - Customer reviews from multiple websites are a good representative of customer voices for the whole market;

1.4 Organization of Dissertation

The rest of the dissertation is organized as follows. As shown in Figure 1.2, Chapter 2 introduces background and terminology, including online customer reviews and comparing them from multiple websites. Some key techniques which the engineering design community may not be familiar with are introduced (e.g., text mining techniques, meta-analysis techniques). Chapter 3 is about Research Thrust 1 wherein data obtained from online customer reviews from a single website are processed and elicited using a modified text mining technique. The elicited customer data is used to construct customer preference models and select a product design. In order to consider more online customer reviews, the data from several websites with different formats are processed and integrated in Chapter 4 wherein Research Thrust 2 is considered. Subsequently, an extension of the model in Chapter 4 is made for integrating online customer reviews from multiple websites in Chapter 5 wherein website and customer preference heterogeneity are considered. Additionally, a set of out-of-sample data is used to validate the models. Finally, Chapter 6 provides concluding remarks of the dissertation as well as the main contributions and possible future research directions.

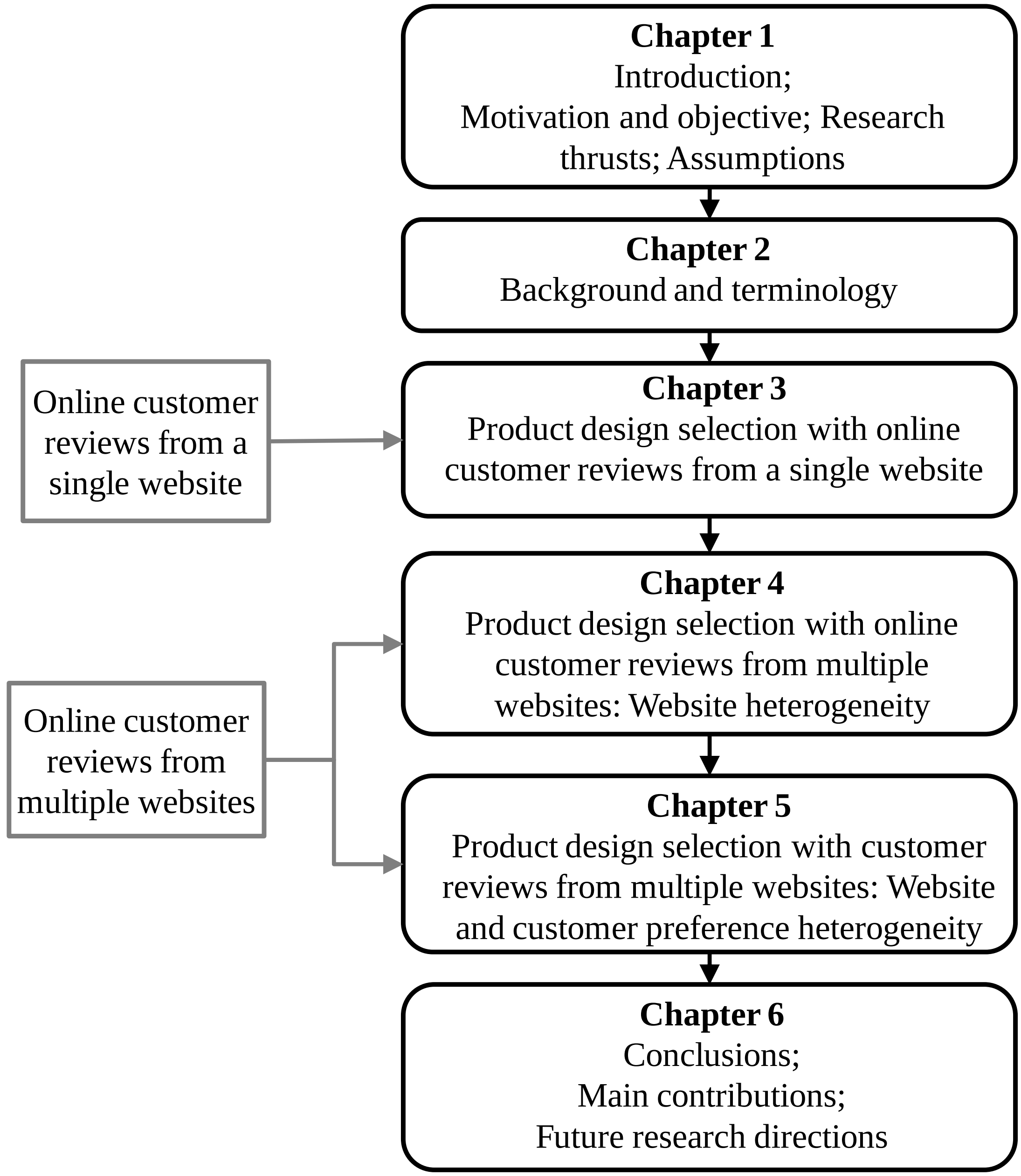


Figure 1.2 Organization of Dissertation

Chapter 2: Background and Terminology

2.1 Introduction

Several terms and background information used throughout this dissertation are defined and provided in this chapter. First, in Section 2.2, an overview of customer-driven product design selection is presented. In Section 2.3, different customer preference models are introduced. Then the terms UGC (in the context of online customer reviews) and text mining techniques required are described in Sections 2.4 and 2.5, respectively. The meta-analysis techniques, which are used for integrating customer reviews from multiple websites, are discussed in Section 2.6. Finally in Section 2.7, model performance evaluation and validation are described.

2.2 Customer-Driven Product Design Selection

Customer-driven product design is a process that integrates customer requirements into product design at early stages of the engineering design process. Customer-driven product design is stated as “just too expensive” [11]. One reason for this high cost is that an incorrect decision at the design selection stage can be very costly to fix later on at the production stage, marketing stage, service stage and so on. The other reason for the cost is that the process of eliciting customer data in customer clinics (e.g., focus groups). Indeed, the quantity and quality of customer data are restricted by budgetary or other resource limitations.

Customer-driven product design selection usually includes several main steps: customer data acquisition, customer preference modeling, design generation, demand and profit estimation, and selection. The related research for each step is reviewed in detail in the following sub-sections.

2.2.1 Conventional Customer Data Acquisition

In recent years, great emphasis has been put on the understanding and modeling of customer preferences in marketing research [12-22] and product design [2-7,23,24]. From the product design aspect, customer data are conventionally collected via interviews, surveys or focus groups. For example, the elicited customer data are used in Quality Function Deployment (QFD) [23,24] to link customer needs and design attributes. However, QFD is essentially a qualitative process. From the marketing research aspect, conjoint analysis [20] is a popular technique used to quantitatively determine how customers evaluate different attributes of individual products. In conjoint, respondents (as representative customers) choose from or rate a set of products or product attributes. Sometimes, product prototypes are used in the conjoint. The importance of each product attribute can be estimated from customer responses. Moreover, the estimated results of conjoint analysis can be used for new product development [2-7,22]. Using conventional methods like conjoint, the acquisition of customer data can be a time-consuming and costly process.

2.2.2 Product Design Selection

Plenty of research efforts have been made to solve the product design problem as a trade-off between performance and the financial objectives (e.g., profit) [2-7,15, 17]. From the financial point of view, the profit can be formulated using a demand model. Extant research has focused on estimating demand and market share of a product taking into account competitive offerings as well as uncertainty in estimating customer preferences [25,26]. Most research has focused on optimizing the multinomial logit function [27, 28] assuming that the demand is linearly dependent on the probability of customer purchase. In the design selection context, Michalek et al. [14] use Analytical Target Cascading to link marketing and engineering product design. This linking assures the feasibility and optimization of product design in the marketing and engineering domains. Kumar et al., [22] incorporate customer preferences in vehicle package design. They treat product qualitative attributes (e.g., vehicle roominess) as a function of engineering design variables, customer demographic and anthropometric characteristics and use multinomial logit models to estimate demand and market share. Similarly Luo et al., [29] use multinomial logit models to estimate demand while taking into account the variations in both product performance and customer preferences.

2.3 Customer Preference Models

Understanding customer preferences has been a popular research area in marketing research for four decades [12-22]. Lancaster first modeled customer preferences by their utilities over each characteristic of the product in 1966 [30,31]. Numerous types of regression models can be used to model customer preferences depending on the types of customer data. Two regression models used in this dissertation are introduced here – linear regression and binary logistic regression.

Customer ratings of products and product attributes are measured using rating scales. The most commonly used rating scale is the Likert scale [30] – which elicits responses in multiple scales (e.g., five "agreement" scales: strongly disagree, disagree, neutral, agree and strongly agree). In the linear regression model, the discrete rating data is assumed as a continuous variable. Let n=1,…, N represent individual customers, the rating of a product or product attribute for customer n is specified as

$$z_n = \boldsymbol{\beta} \cdot \boldsymbol{\Lambda} + \varepsilon_n \tag{2.1}$$

where z_n denotes a dependent variable, such as the rating from customer n; $\boldsymbol{\Lambda}$ denotes a vector of the independent variables, such as product attribute ratings; $\boldsymbol{\beta}$ represents a vector of model parameters; ε_n is an error term, usually assumed to have a normal distribution.

Customer data is not always of the form of ratings data. Sometimes, customer data can be purchase or not purchase decision or other decision variables such as specification or no specification of an attribute as a pro. In that case, the dependent variable z_n is a binary variable, equal to 1 or 0; 1 indicates the customer decides to do something and 0 indicates not to do something. For example, the dependent variable can be the customer decision variable of specifying an attribute as a pro – 1 indicates a customer specifies an attribute as a pro and 0 indicates the customer does not specify the attribute as a pro. Binary logistic regression can be used to model such decision variables [32]:

$$z_n = \boldsymbol{\beta} \cdot \boldsymbol{\Lambda} + \varepsilon_n \tag{2.2}$$

$$\operatorname{logit}(\phi) = \ln \frac{\phi}{1-\phi} = \boldsymbol{\beta} \cdot \boldsymbol{\Lambda} \tag{2.3}$$

where ϕ denotes the probability of z_n equal to 1; ε_n is a statistical error term – independent and identically distributed extreme values.

Customers often have different needs and tastes for the products. Thus customer preferences can be heterogeneous. An important approach to model customer preference heterogeneity is described here – the mixture model. In general, the mixture model is a statistical model for representing the presence of sub-populations within an overall population. Here, the overall population can be regarded as all the customers and the sub-populations can be the individual customers or customer groups. When the sub-populations are individual customers, the mixture model [33] is formulated by Equations (2.4) and (2.5). In this model, customer preference parameters $\boldsymbol{\beta}_n$ are not fixed but random variables. The parameters $\boldsymbol{\beta}_n$ are randomly distributed across customers.

$$z_n = \boldsymbol{\beta}_n \cdot \boldsymbol{\Lambda} + \varepsilon_n \tag{2.4}$$

$$\boldsymbol{\beta}_n \sim \text{Multivariate normal } (\boldsymbol{\mu}, \textstyle\sum) \tag{2.5}$$

where $\boldsymbol{\beta}_n$ are model parameters for n^{th} customer. $\boldsymbol{\mu}$ and $\sum$ denote the mean value and the covariance matrix of the population distribution for $\boldsymbol{\beta}_n$.

Customer preference heterogeneity can be obtained by grouping customers into different latent segments based on their response parameters [34]. Within one segment, customers are considered to be homogeneous – having similar preferences for the products, responding similarly to a market stimulus and so on. Yet, the customer needs and customer responses can be very different across different groups. One of the most powerful segmentation methods is mixture regression models [35]. These models identify customer segments on the basis of the estimated relationship between a dependent variable and a set of independent variables. One of the mixture regression models – the finite mixture regression – estimates a number of unobserved classes in the data and simultaneously relates a dependent variable (e.g., customer rating of a product) with a set of independent variables (e.g., product attributes) through a generalized linear regression model.

Assuming there are M discrete latent classes (segments), the conditional probability density function of a dependent variable z is specified as the weighted sum of the conditional probability density function f_s for Segment s (s=1,…, M), as formulated in Equation (2.6).

$$f\left(z\middle|\mathbf{\Lambda}\right)=\sum_{s=1}^{M} w_s \times f_s\left(z\middle|\mathbf{\Lambda},\boldsymbol{\beta}_s\right) \tag{2.6}$$

where w_s is the weight for Segment s satisfying $0 \leq w_s \leq 1$ and $\sum w_s = 1$. Suppose the dependent variable z is a linear function of an independent variable vector $\mathbf{X}$. For each segment, it can be stated as

$$z_{ns} = \boldsymbol{\beta}_s \cdot \mathbf{\Lambda}_{ns} + \varepsilon_{ns} \tag{2.7}$$

where z_{ns} denotes a dependent variable (e.g., customer rating of a product) for Customer n in Segment s, following a condition probability density function $f_s(z_{ns}|\mathbf{\Lambda}_{ns})$; $\mathbf{\Lambda}_{ns}$ denotes independent variables for Customer n in Segment s; ε_{ns} is a linear regression error; and $\boldsymbol{\beta}_s$ is the model parameter vector in Segment s. The log-likelihood of $f_s(z_n|\mathbf{\Lambda})$ can then be written as

$$\log L = \sum_{n=1}^{N} \log\left(\sum_{s=1}^{M} w_s \times f_s\left(z_n\middle|\mathbf{\Lambda},\boldsymbol{\beta}_s\right)\right) \tag{2.8}$$

where N represents the total number of customers. By maximizing the log-likelihood function, the optimal model parameter vector $\boldsymbol{\beta}_s^*$ can be estimated. The parameters $\boldsymbol{\beta}_s^*$ represent customer preferences, identical within each segment and different across the segments. The differences in $\boldsymbol{\beta}_s^*$ across the M segments represent the degree of customer preference heterogeneity across the segments.

In the mixture regression models, the number of segments is specified before the model estimation. It is recommended to estimate the models for a different number of segments and select the model with the best performance. The performance of a model is judged by a set of information criterion values. The smaller the information criterion value is; the more fit the model is. Three types of information criteria (AIC (Akaike information criterion), CAIC (Consistent AIC) and MAIC (Marginal AIC)) are commonly used [35-38]:

$$\text{AIC: } -2(\log L)_{\max}+2G_M \tag{2.9}$$

$$\text{CAIC: } -2(\log L)_{\max} + G_M[\log(J)+1] \tag{2.10}$$

$$\text{MAIC: } -2(\log L)_{\max} + 3G_M \tag{2.11}$$

where $(\log L)_{\max}$ denotes the maximum likelihood; G_{M} is the number of parameters in a model with M segments; J represents the number of observations.

Sometimes, to consider more evidence from individual customers, multiple observations from a single customer are collected. The multiple observations can be multiple ratings for different products or multiple decisions (such as specifying or not specifying an attribute as a pro) made by a single customer. Recall Equations (2.1) and (2.2), customer data z is related to a summation of $\boldsymbol{\beta} \cdot \boldsymbol{\Lambda}$ and the error term ε. The error term ε is induced by customers' biases due to their inherent habits or backgrounds. For example, if a customer is more conservative in rating a product than other customers, his or her ratings may be consistently lower than an average, thus his or her error term is larger than others. Therefore, the error terms ε's for different observations z's from the same customer should be correlated. Two models for the correlation of multiple responses are introduced here.

The first one is to model multiple decisions from a single customer [39]. In this first model, the error term for decision data is considered to be independent and with identically distributed extreme values as specified in Equation (2.2), which can be represented by two components as

$$\varepsilon_{n,t} = \varsigma_n + e_{n,t} \tag{2.12}$$

where $\varepsilon_{n,t}$ is the error term for the t^{th} observation from the n^{th} customer; ς_n is a customer-specific error term; $e_{n,t}$ is purely random across customers and observations, and it is also independent and identically distributed extreme values. By this formulation, the error term $\varepsilon_{n,t}$ is correlated across observations within the n^{th} customer.

The second one is to model the errors caused by customer habits of using rating scales [40], such as the five "agreement" scale: 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree and 5 = strongly agree. Customers often have different habits to use the rating scales. For example, some customers like to use the high-end of the scale (e.g., use only 4 and 5 of the five "agreement" scale); some just use a narrow interval of the scale (e.g., use only 2 to 4 of the five "agreement" scale). Two broad scale usage patterns that have been identified are called location shift and scale shift. Location shift is the shift of the individual customer's mean response, which is the tendency of some customers to use either the low or high end of the scale. Scale shift is caused by the tendency to use a wide or narrow interval of the scale.

The t^{th} observation from n^{th} customer is specified as

$$z_{n,t} = \boldsymbol{\beta} \cdot \boldsymbol{\Lambda} + \eta_n + \lambda_n \cdot \varepsilon_{n,t} \tag{2.13}$$

$$\varepsilon_{n,t} \sim N(0, \sum) \tag{2.14}$$

where η_n and λ_n represent the location and scale shift of n^{th} customer, respectively; they are modeled via a bi-variate normal prior:

$$\begin{bmatrix} \eta_n \\ \ln \lambda_n \end{bmatrix} \sim \mathrm{N}(\boldsymbol{\varphi}, \boldsymbol{\Sigma}) \tag{2.15}$$

2.4 Online Customer Reviews

2.4.1Customer Reviews in a Single Website

Online customer reviews are written by customer product users. With more people becoming familiar with the worldwide web, the amount of such reviews is increasing over time. Online customer reviews have the advantage of low cost and high volume. Figure 2.1 presents an example of an online customer reviews from www.bestbuy.com.

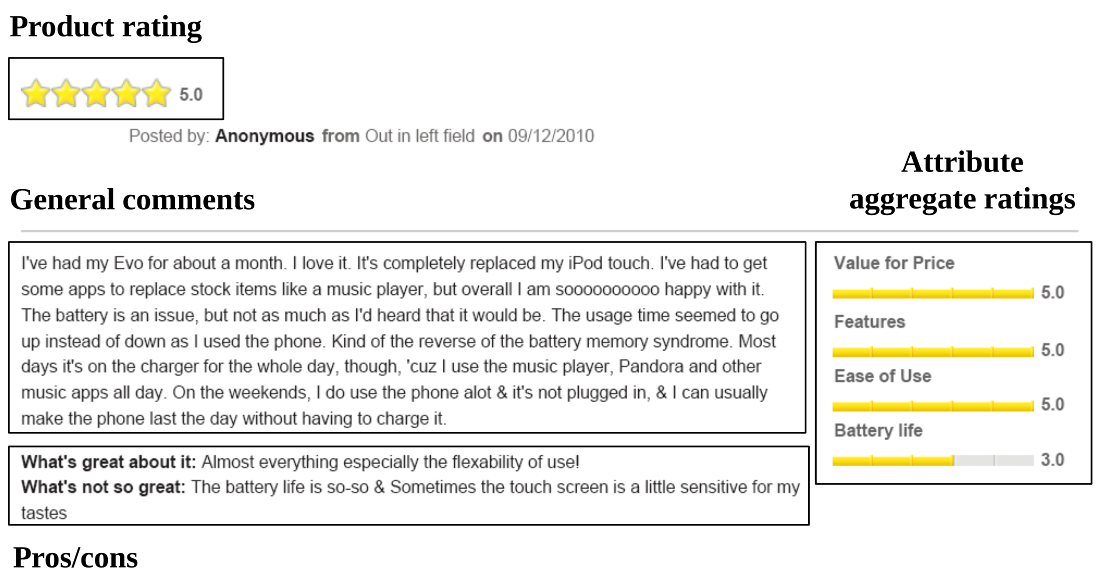


Figure 2.1 Customer reviews – www.bestbuy.com

The customer reviews may include numerical ratings and textual reviews. As highlighted in Figure 2.1, a customer provides two types of product ratings: an overall rating of a product and ratings of product attribute aggregates (e.g., four aggregates in Figure 2.1: “value for price”, “features”, “ease of use”, and “battery life”). As shown in Figure 2.1, customer textual reviews are composed of three parts: open-ended comments (called general comments), “what’s great about it” (the pros of a product), and “what’s not so great” (the cons of a product). General customer review comments can contain a lot of noise and as such a significant portion can become irrelevant to product attributes. Too much noise can exacerbate the complexity of a text mining process and thus impair the accuracy of the results. The pros/cons are summarized comments and simply listed phrases or sentences. References [16,41] state that text mining results using general comments are significantly worse than those from review summaries (pros/cons).

2.4.2 Customer Reviews in Multiple Websites

Different public websites usually contain their own formats of customer reviews and require different procedures to acquire customer reviews. To better understand such heterogeneity, the popular websites containing reviews can be categorized into four groups depending on the types of customer data included. As shown in Table 2.1, customer reviews can include four types of customer data: overall product ratings, attribute aggregate ratings, general comments and pros/cons. Group I of the websites includes all the four customer data types (Figure 2.2); Group II contains three data types except pros/cons data (Figure 2.3); Group III comprises three data types except attribute aggregate ratings (Figure 2.4); and Group IV includes two data types – product ratings and general comments (Figure 2.5). It is found that two common types of customer data are product ratings and general comments. The four groups are distinguished from each other by whether they contain attribute aggregate ratings and/or pros/cons.

Additionally, there exist two extra format differences among the websites: (i) the attribute aggregates are not identical; and (ii) the pros/cons data is not collected following the same procedures. For the websites with attribute aggregate ratings (Groups I and II), the attribute aggregates are determined by each website. Thus attribute aggregates are distinguished from different websites. For instance, www.bestbuy.com has four attribute aggregates (shown in Figure 2.2) – "Value for price", "Features", "Ease of use" and "Battery life". The website www.letstalk.com has four aggregates (shown in Figure 2.3) – "Call quality", "Ease of use", "Design" and "Battery life". The two websites only have two attribute aggregates in common. For the websites with pros/cons data, pros/cons are collected in two different ways. Some websites allow customers to summarize pros/cons freely (as www.bestbuy.com shown in Figure 2.6); other websites provide a checklist of pros/cons for customers (as www.tmobile.com shown in Figure 2.6). All of the distinctions among the websites will directly require different procedures of customer data elicitation and processing.

Table 2.1 Comparison of different websites

Groups	Product rating	Attribute aggregate rating	General comments	Pros/cons
I	✓	✓	✓	✓
II	✓	✓	✓	
III	✓		✓	✓
IV	✓		✓	

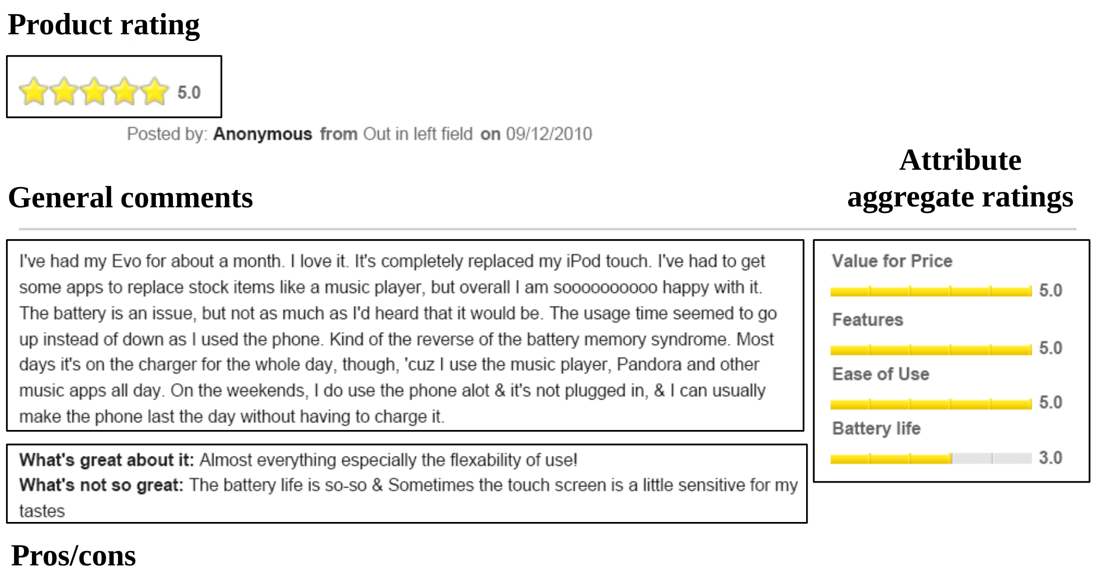


Figure 2.2 Group I sample – www.bestbuy.com

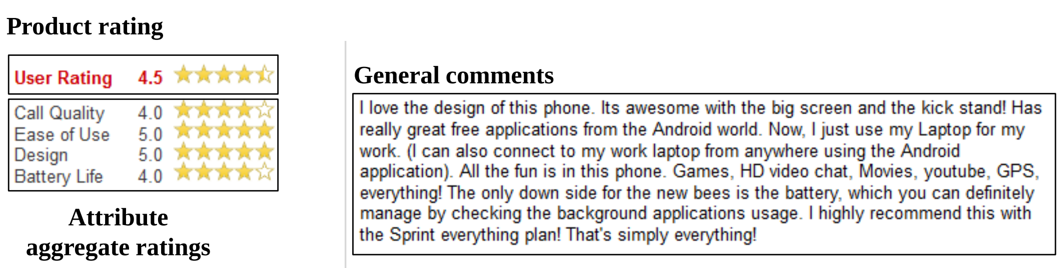


Figure 2.3 Group II sample – www.letstalk.com

Product rating

"Excellent phone overall"

Pros/cons

Pros: Size, useful apps, fairly easy to use (Android 2.2).

Cons: Terrible notifications bar (tells very little unless you drop down a special window. Lame).

Summary: Generally very happy with this phone. Notifications are poorly implemented. I was concerned about battery life, but now with Android 2.2, battery life has improved quit a bit. I'm happy with the switch from Pre to EVO. Great phone overall.

General comments

Figure 2.4 Group III sample – www.cnet.com

Product rating

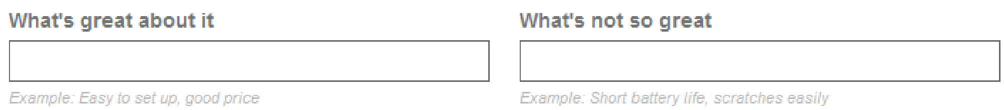
Lives up to its name once again!, May 2, 2011

General comments

I love this phone to death. I constantly use it all day long and get about a day's charge out of the battery which is fine and above average for most Android phones. It's ease of use, responsive touch and input make this phone fantastic. Everything is fast and smooth, the CPU installed is a powerhouse and fast enough for most people. Don't care about the lack of 4G since my area won't have it for a year or two at least. I would recommend this phone to anyone who likes Android and HTC Sense.

Figure 2.5 Group IV sample – www.amazon.com

www.bestbuy.com



www.tmobile.com

Figure 2.6 Different pros/cons formats

2.5 Text Mining Techniques

Textual customer reviews are very informal, even the pros/cons types of reviews. Without text mining, it is very difficult to extract any useful information for product design selection from textual comments. The general steps of text mining include pre-processing, text representation and content analysis [16,42].

The purpose of the pre-processing is to clean and normalize the textual data. It includes two steps: removing stop words and stemming. The stop words from a standard list [43,44] including articles, conjunctions, prepositions and some commonly used but meaningless words, like “is”, “been”, are removed from each phrase. For instance, after removing the stop words, the phrase “the battery life is so-so” becomes “battery life so-so”. The stemming process attempts to normalize the words by reducing the words into their stem or root form [44]. The Porter’s algorithm

[45] is widely used for stemming, which provides explicit steps how to normalize the words. For instance, at Step 1, the suffix “sses” is replaced by “ss”. At Step 2, the suffix “tional” is replaced by “tion”. The details of the algorithm are omitted in this dissertation. As an example of the stemming, “happy” and “happiness” are both stemmed as “happi”. In most cases, different variations of a stem are similar in semantic interpretations and that is why stemming is used. After this step, different variations of a stem are made equivalent and normalized into a single stem. Text representation is a step to transform raw text into numeric vectors. The most widely used representation is the vector space model [46]. According to this model, the text (e.g., a review sentence) is represented by a vector whose dimension is the number of features (i.e., the words of interest) and the components represent the appearance of the features (e.g., the frequency of the features). This step can prepare text for a possible follow-up content analysis.

The content analysis includes classifying reviews, determining the orientations of customer reviews and so on. A great deal of effort has been made to identify product attributes, determining the orientations (positive or negative) of customer reviews from online customer reviews (one type of reviews). Some literatures [47-50] used a supervised classification method to identify attributes from customer reviews. Supervised classification methods [47-50] take both training and testing steps. The training step develops a classification rule by analyzing training datasets. Each dataset includes a pair of an input vector (e.g., a review sentence) and a class output (e.g., a corresponding product attribute). In the literatures, the class outputs for the dataset are manually assigned. The testing step predicts the class outputs of new review sentences (testing datasets). Some references [41,42,51-54] identify product attributes by identifying the frequently-used noun and noun phrases. Refs [16,55] present an automated procedure to obtain conjoint attributes and levels from the pros/cons list of specific public websites. Some research efforts [42,56-58] in using online customer reviews have been made to determine the orientations (positive or negative) of customer reviews based on the occurrences of particular sentiment phrases. Some literature from marketing research [59-63] focuses on the impact of product reviews on product sales. One paper [64] uses a sophisticated text mining technique to incorporate customer review content into sale prediction models. Some recent work [65] makes efforts to extract market structure information from online customer reviews through text mining techniques. From an engineering perspective, a recent paper [66] presents a web-based framework to enable collective innovation (innovation through collaboration) in the early stage of product development.

A supervised classification method is introduced in the following. Support Vector Machine (SVM) [67] is a popular supervised classification method with excellent precision. Let $\mathbf{S}_k = \{(u_k, \mathbf{v}_k)\}$ denote the k^{th} example in the training set, $u_k \in \{\pm 1\}$ be the class for the input vector $\mathbf{v}_k$. For example, for the input vector ($\mathbf{v}$) for the review sentence "Great call quality", the class for this input is assigned as a pro (u = 1). For the input for the review sentence "Terrible signal", the corresponding class is a con (u = –1). Figure 2.7 illustrates an SVM problem when the input vector $\mathbf{v}$ is in two dimensions. The solid dots in Figure 2.7 represent the examples in the class u = 1 and the circles represent the examples in the class u = –1. SVM is a binary classification method aimed at finding a hyper-plane that best separates the two classes of examples in the training set. The desired hyperplane can be defined as $\mathbf{W} \cdot \mathbf{v} - b = 0$, where " $\cdot$ " denotes the dot product and **W** is a weight vector. **W** and b are obtained to maximize the minimum distance between the two classes. As shown in Figure 2.7, the hyperplanes parallel to the desired hyperplane and in the margin for the two classes are $\mathbf{W} \cdot \mathbf{v} - b = 1$ and $\mathbf{W} \cdot \mathbf{v} - b = -1$, respectively. The two hyperplanes should be as far apart as possible but still separating the two classes of examples. For the examples belonging to the class ($u = 1$), there is $(\mathbf{W} \cdot \mathbf{v} - b) \geq 1$. Similarly, for the examples in the class ($u = -1$), there is $(\mathbf{W} \cdot \mathbf{v} - b) \leq -1$. The two inequalities can be merged and rewritten as $u \times (\mathbf{W} \cdot \mathbf{v} - b) \geq 1$. Using the geometry, one can find the distance between the two hyperplanes to be equal to: $2/\|\mathbf{W}\|$. The objective of maximizing the distance is equivalent to the objective of the optimization problem in Equation (2.16) – to minimize $\|\mathbf{W}\|$. The objective is subject to that all the training examples are separated by the two hyperplanes (which is the constraint in Equation (2.16): $u_k \times (\mathbf{W} \cdot \mathbf{v}_k - b) \geq 1$ for all k).

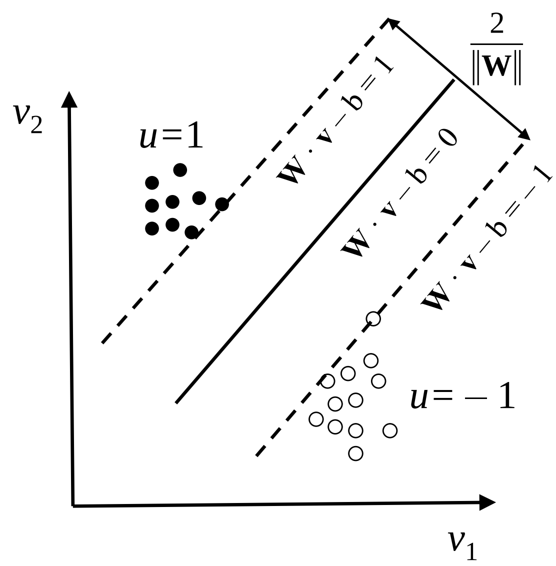


Figure 2.7 SVM problems for two-dimensional input vectors

$$\begin{cases} \underset{\mathbf{w},b}{\text{minimize}} \ \|\mathbf{W}\| \\ \text{subject to: } u_k \times \left(\mathbf{W}\cdot\mathbf{v}_k - b\right) \geq 1, \text{for all } k \end{cases} \tag{2.16}$$

Once the optimization problem is solved, new vectors $\mathbf{v}_{new}$ can be classified by comparing with the hyperplane $\mathbf{W} \cdot \mathbf{v} - b = 0$. If $(\mathbf{W} \cdot \mathbf{v}_{new} - b) \geq 0$, this vector will be classified into the class (u = 1); otherwise, it will be classified into the class (u = –1).

2.6 Meta-Analysis Techniques

Meta-analysis has been commonly executed in the area of health sciences. It is a statistical method to combine results from multiple medical clinics. If several clinics address the same research question, then it is useful to combine information from all of them. The meta-analysis of multiple clinics provides a more precise estimation of medical treatment effects and may provide valuable information regarding the differences between the clinics. The earliest example of meta-analysis proposed by Karl Pearson in 1904 [70] was a combination study of typhoid vaccine effectiveness.

He tried to overcome the problem caused by small sample sizes and analyzed the results from multiple clinics. In the past 30 years, meta-analysis has gained acceptance as a method for integrating relevant and comprehensive evidence in medical and clinical research [71]. Today, it is also used for psychological, educational, social science, and market research as well as other fields [71-73].

In a medical clinic, different treatments are usually used. For example, some patients are treated with a new medicine but others are not. A treatment effect is the effect of the treatment differences on how patients respond. Two frequently-used models in meta-analysis are the fixed effects model and the random effects model [71,80]. The two models are distinguished by the way one treats the variation in estimated treatment effects between the clinics. In the fixed effects model, the assumption is made that there is a global average effect and the clinic differences from the average effects are caused solely by sampling variation. Individual clinic results are combined to estimate this global average effect. In contrast, the random effects model views the between-clinic variation of the estimated treatment effects due to random variations in the way individual clinics are designed, conducted and measured. The "true" treatment effects estimated for individual clinics are allowed to be different (not assumed common as in the fixed effects model). The effects are regarded as drawn from a population of clinic effects. The mean of this population distribution is the overall treatment effect and its variance represents the uncertainty added by the differences between clinics. A fundamental assumption of the random effects model is that the individual clinics are exchangeable, which means that they are not identical replications but similar enough to be useful for estimating the parameters of the population distribution of clinic effects [75]. Sometimes, the independent variables are not completely identical across all the clinics. Reference [72] applied the random-effects meta-analysis for combining information from several medical clinics with incomplete information.

Because of its appeal, Bayesian methods have also been broadly used for meta-analysis to deal with the fixed effects and random effects model. The advantages of Bayesian-based methods are their great flexibility, the ability to quantify uncertainties of parameters and the ability to handle models in a complex fashion. Through Bayesian methods, researchers can express their prior belief of the clinic effects and then update the belief by taking into account the emerging clinic data. The computational complexity of Bayesian methods can be improved by employing Markov Chain Monte Carlo (MCMC) simulation methods [81].

2.7 Model Performance Evaluation and Validation

The performance of a model can be evaluated by a pseudo-r^2 value [82]. The pseudo-r^2 value, as defined in Equation (2.17), is defined as the degree of the predictive capability of a model. The pseudo-r^2 value can be anywhere from 0 to 1. The higher the value of the pseudo-r^2 is, the better the predictive capability of the model. The pseudo-r^2 is defined as:

$$r^2 = \sum_{j=1}^{J}\left(\hat{y}_j - \bar{y}\right)^2 \Big/ \sum_{j=1}^{J}\left(y_j - \bar{y}\right)^2 \qquad (2.17)$$

where $\hat{y}_j$ and y_j indicate predicted and observed ratings of the j^{th} observation, respectively; $\bar{y}$ indicates the average value of actual ratings; J indicates the number of observations.

The errors between the predicted values and observed values can be quantified by the root mean squared error *RMSE*, normalized root mean squared error *NRMSE*, mean absolute error *MAE* and mean absolute percentage error *MAPE* values [83-86]. All the error quantifications can be used for evaluating model performance as well as model validation. The *RMSE* defined in Equation (2.18) is the difference between the predicted and observed values. The *RMSE* value can range from 0 to ∞. The smaller the value of *RMSE*, the better the predictive power of the model. *RMSE* for different data sets cannot be directly compared unless it is appropriately normalized. The *NRMSE* in Equation (2.19) is a normalized measure of *RMSE* by dividing the range of observed values y [84]. The *NRMSE* value ranges from 0 to 1.

$$RMSE = \left[\sum_{j=1}^{J} \left(\hat{y}_j - y_j \right)^2 \Big/ J \right]^{1/2} \tag{2.18}$$

$$NRMSE = RMSE \big/ \left(y_{\max} - y_{\min} \right) \tag{2.19}$$

where $y_{\max}$ and $y_{\min}$ represent the maximum and minimum observed ratings, respectively.

The *MAE* and *MAPE* as defined in Equations (2.20) and (2.21) are two commonly used error quantification values.

$$MAE = \sum_{j=1}^{J} \left| \hat{y}_j - y_j \right| \Big/ J \tag{2.20}$$

$$MAPE = \sum_{j=1}^{J} \left| \left(\hat{y}_j - y_j \right) \big/ y_j \right| \Big/ J \tag{2.21}$$

The deviance information criterion *DIC* [87,88] values defined as in Equation (2.22) are more useful for Bayesian model selection problem.

$$DIC = \mathrm{D}\left(\bar{\xi} \right) + 2 p_D \tag{2.22}$$

where $\mathrm{D}(\bar{\xi}) = -2\log\left[f\left(y|\bar{\xi}\right)\right]$ is the deviance obtained by substituting the posterior mean values $\bar{\xi}$ of the model parameters ξ into the log-likelihood function, y is the observed data, and p_D denotes the effective number of model parameters, defined as $p_D = \mathrm{E}\left\{-2\log\left[f\left(y|\xi\right)\right]\right\} + 2\log\left[f\left(y|\bar{\xi}\right)\right]$. Generally, a smaller *DIC* value is associated with a better model–data fit.

2.8 Summary

In this chapter, the background knowledge required for this dissertation and some terminology which may not be familiar were introduced. Online customer reviews are the focus of this dissertation, therefore, the relevant literature and techniques are carefully reviewed.

In the next chapter, customer reviews from a single website are processed and used for customer-driven product design selection in order to overcome the limitation of conventional customer-driven product design.

Chapter 3: Product Design Selection using Online Reviews from a Single Website

As stated before, acquisition of the customer data for product design selection using traditional customer survey techniques can be a time-consuming and costly undertaking. The aim of this chapter is to overcome this limitation by using online customer reviews as an alternative to the traditional customer survey techniques such as conjoint. So far, there has not been any systematic effort in using online customer reviews in design selection for a durable product. Using online reviews in product design selection is not an easy task because the reviews are not specifically survey-designed and collected for product design.

This chapter develops a systematic methodology for eliciting product attributes from online reviews, constructing customer preference models and using these models in design selection. To demonstrate the proposed method, design selection of a smartphone is considered.

Section 3.1 gives a brief introduction and review of previous work. Section 3.2 defines the problem. Section 3.3 describes the proposed methodology. Section 3.4 presents a smartphone case study that demonstrates an application of the proposed method. Finally, Section 3.5 gives some concluding remarks.

3.1 Introduction

Web based online reviews for products have two main advantages: They are available at very low cost and copious quantity of data. But online reviews have an obvious disadvantage too: They include free expressions with a lot of noise. In order to overcome this particular disadvantage in online reviews and using it for product design selection, this chapter makes the following contributions, as discussed in the next two paragraphs.

The first contribution is the use of a new data source, online reviews, to overcome the current limitations in customer-generated data acquisition and processing. To the best of our knowledge, no systematical effort has been made so far to use such reviews in the context of customer-driven product design selection. Some research of studying online reviews [56-58] was on determining the orientations (positive or negative nature) of customer reviews by the occurrences of particular sentiment phrases. Part of the previous research focused on identifying product attributes discussed in customer reviews [16,42,47-55]. Some marketing researchers [59-63] make efforts in studying the impact of product reviews on product sales. From the engineering perspective, a recent paper [66] presents a web-based framework to enable collective innovation by learning online reviews. It is expected that the introduction of online reviews, as presented in this chapter, will reduce the cost of design selection schemes and more importantly dramatically reduce the time in obtaining customer data.

The second contribution of this chapter is that it proposes a new methodology for using online reviews for product design selection by extending and integrating several existing methods of customer preference modeling and customer-driven product design. As mentioned before, online reviews contain a lot of noise, variability and even bias induced by the customers themselves. Meanwhile, online reviews rarely contain detailed information on customers, like gender, income and other such customer-specific data. Lack of detailed information can make it harder to model customer preferences because customer preferences are often subjective and can be explained by customer-specific data. This second contribution makes it feasible to use customer data elicited from online reviews, construct customer preference models, and help in durable product design selection. More specifically, the existing text mining techniques are extended in this chapter to make the data elicitation from online reviews applicable for customer preferences modeling. The existing customer preference modeling methods and customer-driven product design methods are tailored and integrated to make use of the customer preference information obtained from online reviews for product design selection.

3.2 Problem Definition

As shown in Figure 3.1, a design engineer desires to make a design selection for a consumer durable product such as a mobile phone, an appliance, or an automobile. Given a series of design alternatives, the design engineer wants to select a design alternative which maximizes the expected profit for a manufacturer. The customer data is collected from online customer reviews. Both quantitative (numerical) and qualitative (textual) data exist for developing customer preference models and ultimately for selection. The objective of the proposed approach is to elicit customer data from online reviews, construct customer preference models and select a product design alternative.

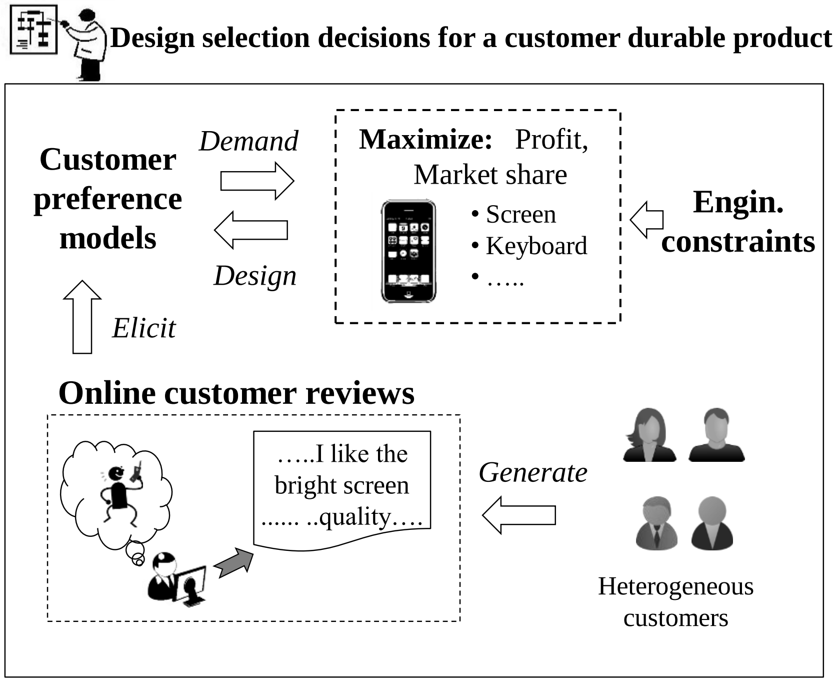


Figure 3.1 Problem definition

3.3 Approach

The proposed approach, as shown in Figure 3.2, involves three tasks: (Task 1) elicit product attributes and customer data; (Task 2) construct hierarchical customer preference models; and (Task 3) select a product design. In Task 1 (Section 3.3.1), online customer reviews are processed by a text mining technique to identify important product attributes. Customer ratings data is then collected from online reviews for the analysis in the subsequent tasks. To resolve the difficulty of modeling customer preference heterogeneity, Task 2 (Section 3.3.2) will focus on modeling the unobserved customer preference heterogeneity by segmenting customer data using mixture regression models. Simultaneously, the customer preference models are constructed for each segment using both customer ratings data and publicly known product specification data, which are available on the Internet. Specifically, the customer preference models considered in this chapter are multi-level, linking the customer ratings on products, product attributes and product specifications. In Task 3 (Section 3.3.3), a profit function is formulated using the customer preference models given the relationships between the probability of product purchase and product ratings. The profit is incorporated as a design objective which is maximized over all design alternatives for product design selection. The three tasks are detailed in the next three subsections.

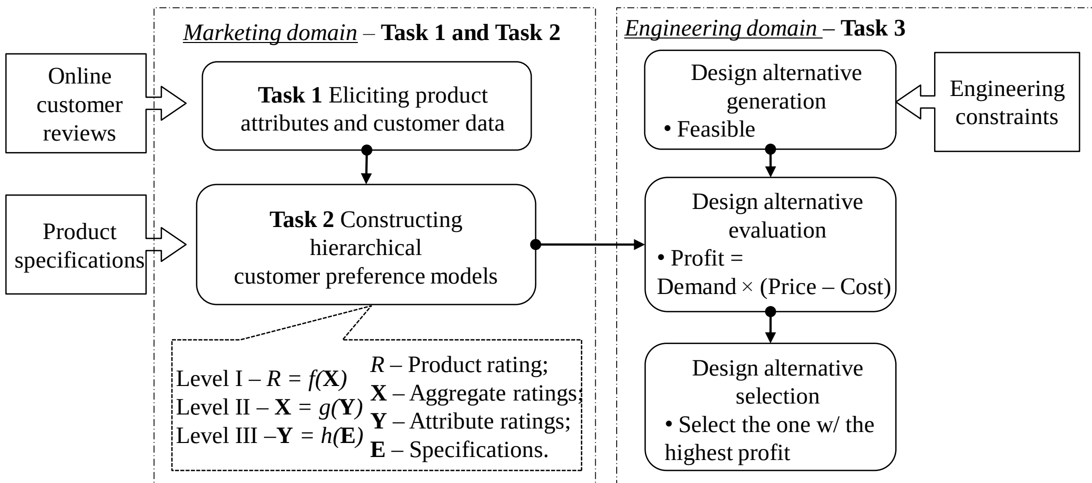


Figure 3.2 The methodology framework

3.3.1 Eliciting Product Attributes and Customer Data

The goal of this task is to identify important product attributes and elicit customer data from online customer reviews. As shown in Figure 3.3, Task 1 includes two main steps: identifying product attributes and eliciting customer data. In the first step, in order to identify product attributes, frequent words are identified as candidate attributes and meaningless frequent words are pruned. Meanwhile, a dictionary defining the words related to each attribute is constructed. In the second step, using the constructed dictionary, customer ratings for each attribute can be identified and elicited. These are described in detail next.

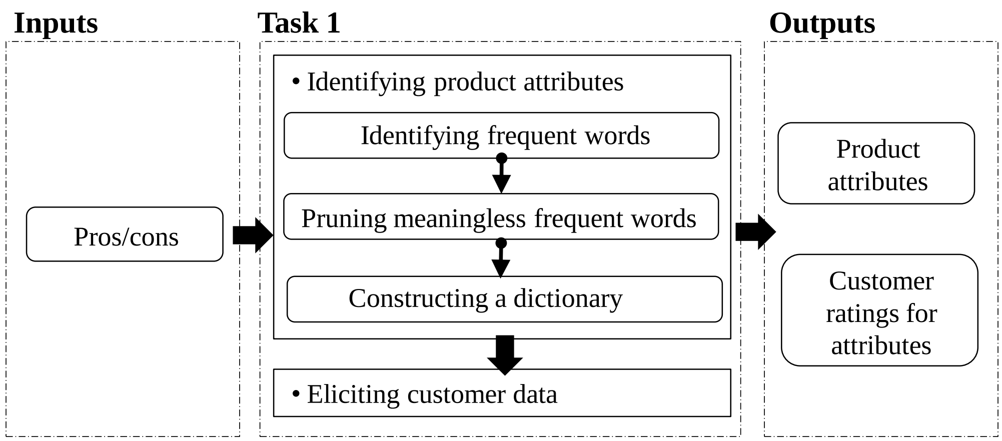


Figure 3.3 Task 1 framework

Identifying Product Attributes: The first crucial step in Task 1 is to identify important product attributes from customer reviews. To avoid excessive noise in customer reviews, this step only considers the lists of pros/cons to identify product attributes in this study. Each pro or con list cannot be directly used as inputs for identifying product attributes. They have to be partitioned into separate phrases or sentences. The phrases or sentences in each pro or con list are divided by standard separators including commas, slashes, and semicolons. Each phrase or sentence is a single input for identifying product attributes and assumed to correspond to only one single product attribute [16]. For example, the con list in Figure 3.4 contains two sentences and corresponds to two product attributes “battery life” and “touch screen”.

Figure 3.4 An example of a customer review from www.bestbuy.com

It is proposed to improve a previously developed part-of-speech tagger based technique, e.g., [42,54], in order to identify product attributes better. The text mining technique can identify frequent words from noun and noun phrases as candidate product attributes and then figure out product attributes from the candidates. Two improvements are made. First, the way to identify candidate product attributes is improved by accounting for noun and verb and their phrases, while the literature focuses on noun (e.g., "screen") and noun phrases (e.g., "battery life") only [42,54]. This improvement is based on the observations made from customer reviews. Some product attributes are described using verb phrases like "easy to use", "easy to text". If only noun and noun phrases are considered, the product attributes described using verb and verb phrases may be missed.

Secondly, the approach also improves the way to identify product attributes from candidate attributes. Two pruning rules (the compactness pruning rule and redundancy pruning rule) [42,54] are initially considered to check the candidate attributes. Compactness pruning is to check candidate attributes with at least two words and remove those that are likely to be meaningless. Redundancy pruning focuses on removing redundant candidate attributes with a single word. The candidates that do not satisfy these rules are removed. The retained candidates are identified as product attributes. However, it was found that the compactness pruning rule is not appropriate for the study here because this rule is aimed at removing the meaningless candidates containing at least two words. It is suggested [42,54] that in a natural language sentence, when the words in the candidate attributes appear together and in a specific order, then they are most likely to be meaningful. However, the candidates from pros/cons are very short; most contain only one word and a few contain two words. Thus, the compactness pruning rule is not applicable in the proposed method because it is not able to check the candidates according to the order of words.

An independency pruning rule is proposed to replace the above mentioned compactness pruning rule – the case with short candidate attributes. The consideration of the independency pruning rule is based on the assumption that each phrase from pros/cons corresponds to one product attribute. Each phrase may contain candidate attributes and other words. If the candidate attributes are meaningful product attributes, then in the phrases, the candidates can be used independently and do not need to be used with other noun or verb words to describe a product attribute. For example, consider the phrase in Figure 3.4: “The battery life is so-so”. Assume that “battery life” is a candidate attribute. In the phrase “The battery life is so-so”, the candidate attribute “battery life” is not used with other nouns or verbs. It is only used with an adjective word “so-so”. (“The” and “is” are presumed to be already removed as stop words during a pre-processing step.) Thus, “battery life” is used independently and can be identified as a meaningful product attribute. A candidate attribute is defined as independent in a phrase if it is not used together with other nouns or verbs. The independence pruning rule is: The candidates that cannot be determined as independent in any phrases from pros/cons will be removed.

Two examples are given to demonstrate how the independency pruning rule proposed here works. Consider a candidate attribute “texting” and three phrases from the pros/cons containing this candidate: *“texting is difficult”*, *“texting”*, *“texting capabilities”*. The candidate “texting” is independent because two sentences use the candidate without the support of other nouns/verbs. As another example, there is a candidate attribute “time” and three phrases from the pros/cons containing this candidate: *“slow response at times”*, *“long time for shipping”*, *“there is a lag all the time”*. The candidate attribute “time” is not independent and should be removed because all the phrases contain other noun/verbs. This indicates that the candidate “time” is not a meaningful product attribute and cannot be used independently to describe a product attribute.

Using the two pruning rules (the redundancy pruning rule and the independency rule), the important product attributes can be identified. The attributes are identified individually from the pros and the cons and then merged to be the attributes of interest. The reason to identify attributes individually from pros and cons is that some attributes may only appear in pros or cons. For example, the attribute “ease of use” is widely mentioned as pros but barely specified as cons.

Due to the variations in the use of language, customers tend to use different words to describe even the same thing. For example, for the "Internet" attribute, customers may also use the words like "internet browser", "web browser", "web" and so on. A dictionary which defines the words related to each attribute can be constructed first. The dictionary can then be used to identify the product attributes that customers describe in their reviews.

Eliciting customer data: After mapping customer reviews into product attributes, the next step is to elicit customer opinions in a numerical fashion. Previous work [64] elicited customer opinion data according to the adjective words used. A grade was assigned from –3 (strongly negative) to +3 (strongly positive) to the reviews according to different adjective words. For example, "horrible" is more negative than "bad". However, the grading process is a subjective process and may introduce bias into customer data.

In this dissertation, two indicators for each attribute are defined as in Equations (3.1) and (3.2) – one for pros and one for cons. For convenience, the two indicators of each attribute are called as attribute ratings in the rest of this dissertation.

$$\text{Attribute\_pros} = \begin{cases} 1 & \text{The attribute is classified into pros} \\ 0 & \text{The attribute is not classified into pros} \end{cases} \tag{3.1}$$

$$\text{Attribute\_cons} = \begin{cases} 1 & \text{The attribute is classified into cons} \\ 0 & \text{The attribute is not classified into cons} \end{cases} \tag{3.2}$$

For instance, if the attribute "screen" is only classified into pros, the two indicators for "screen" are screen_pros = 1 and screen_cons = 0. If "call quality" is not classified into pros or cons by a customer; then the values of the two indicators are both equal to 0. This elicitation does not depend on the adjective words used by customer so it can avoid any possible bias.

3.3.2 Constructing Hierarchical Customer Preference Models

In general, customers can provide the textual reviews of a product and its numerical rating. Some websites ask customers to provide their numerical ratings of product attribute aggregates as well. Here, it is presumed that the customer data from online reviews are in a hierarchical structure as shown in Figure 3.5. The top level gives the overall product ratings. The intermediate levels include the attribute aggregate ratings and attribute ratings. Attribute aggregate ratings are not always available and it is thus dash-lined in the pyramid structure of Figure 3.5. In order to model customer preferences from online reviews for product design selection, the product specifications that are publicly available at the bottom level of Figure 3.5 are collected. Product specifications are information about the objective attributes and features of the product supplied by the product vendor. The consideration of publicly available product specifications can be beneficial in two ways: (i) product specifications are easy to collect; and (ii) product specifications are useful links between customer preferences and product design.

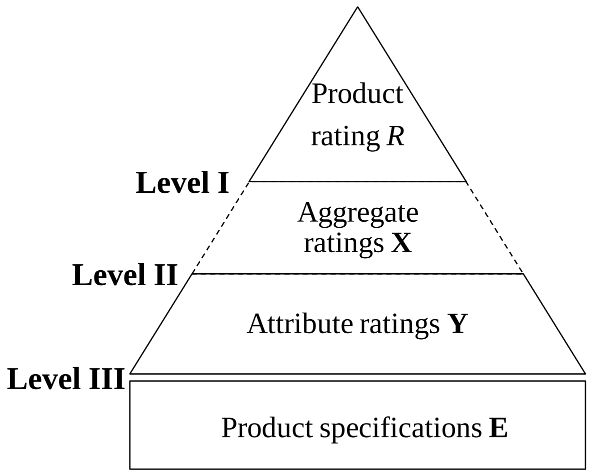


Figure 3.5 Data structure

Based on the data structure, as shown in Figure 3.5, it is proposed to formulate customer preference models in a hierarchical fashion. The hierarchical customer preference model is formulated using a bottom-up approach that predicts the ratings at an upper level using those at a lower level(s). The hierarchical customer preference models are formulated as follows.

Level I – a product rating R is a function of attribute aggregate ratings $\mathbf{X}$

$$R = \boldsymbol{\beta}_{\mathrm{I},n} \cdot \mathbf{X} + \varepsilon_{\mathrm{I}} \tag{3.3}$$

Level II – aggregate ratings $\mathbf{X}$ is a function of attribute ratings $\mathbf{Y}$

$$\mathbf{X} = \boldsymbol{\beta}_{\mathrm{II},n} \cdot \mathbf{Y} + \boldsymbol{\varepsilon}_{\mathrm{II}} \tag{3.4}$$

Level III – attribute ratings $\mathbf{Y}$ is a function of product specifications $\mathbf{E}$

$$\mathbf{Y} = \boldsymbol{\beta}_{\mathrm{III},n} \cdot \mathbf{E} + \boldsymbol{\varepsilon}_{\mathrm{III}} \tag{3.5}$$

where R, $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{E}$ are the product rating, aggregate ratings, attribute ratings and product specifications; $\boldsymbol{\beta}_{\mathrm{I},n}$, $\boldsymbol{\beta}_{\mathrm{II},n}$, and $\boldsymbol{\beta}_{\mathrm{III},n}$ are the model parameters, varying with customer n; ε_{I}, $\boldsymbol{\varepsilon}_{\mathrm{II}}$, and $\boldsymbol{\varepsilon}_{\mathrm{III}}$ are the error terms, which are correlated in the hierarchical models.

For the case without the attribute aggregate ratings, the hierarchical models can be reduced into two levels as:

Level I – a product rating R is a function of attribute ratings $\mathbf{Y}$

$$R = \boldsymbol{\beta}_{\mathrm{I}} \cdot \mathbf{Y} + \varepsilon_{\mathrm{I}} \tag{3.6}$$

Level II – attribute ratings $\mathbf{Y}$ is a function of product specifications $\mathbf{E}$

$$\mathbf{Y} = \boldsymbol{\beta}_{\mathrm{II}} \cdot \mathbf{E} + \boldsymbol{\varepsilon}_{\mathrm{II}} \tag{3.7}$$

The parameters at the different levels can be estimated using the finite mixture regression model introduced in Chapter 2 in a hierarchical fashion.

3.3.3 Product Design Selection

First, in this task, design alternatives are generated given the engineering constraints. For each design alternative i, the product rating $R_i(\mathbf{E}_i, \boldsymbol{\xi})$ can be predicted in a statistical fashion using the estimated models in Section 3.3.2, where $\boldsymbol{\xi}$ denotes all the parameters in the model – including $\boldsymbol{\beta}$'s and so on.

In order to predict the demand and profit of each alternative, the relationship between the product rating R_i and the probability of purchase P_i must be known. However, due to the lack of customer choice data in online reviews, the relationship cannot be estimated. Thus, in this dissertation, different relationships are presumed and the selected design alternative should be insensitive to the change of the relationships. The relationship is assumed to have three forms as representatives – linear, binary logit, and nonlinear (exponential) – as defined in Equations (3.8) to (3.10).

$$P_{(1),i}(\mathbf{E}_i, \boldsymbol{\xi}) = a_1 \times R_i(\mathbf{E}_i, \boldsymbol{\xi}) + c_1 \tag{3.8}$$

$$P_{(2),i}\left(\mathbf{E}_i, \boldsymbol{\xi}\right) = \frac{\exp\left(a_2 \times R_i\left(\mathbf{E}_i, \boldsymbol{\xi}\right) + c_2\right)}{\exp\left(a_2 \times R_i\left(\mathbf{E}_i, \boldsymbol{\xi}\right) + c_2\right) + 1} \tag{3.9}$$

$$P_{(3),i}(\mathbf{E}_i, \xi) = a_3 \times \exp(R_i(\mathbf{E}_i, \xi)) + c_3 \tag{3.10}$$

where $P_{(1),i}$ represents the probability of purchase for the design alternative i under the first relationship (linear), $P_{(2),i}$ represents the probability of purchase for the design alternative i under the second relationship (binary logit), $P_{(3),i}$ represents the probability of purchase for the design alternative i under the third relationship (exponential), a_1, c_1, a_2, c_2, a_3 and c_3 represent the constants in each relationship.

In order to bound the probability of purchase P in the range of [0,1], two constraints are applied to the relationships: $P(R_{\min}) = 0$ and $P(R_{\max}) = 1$, where $R_{\min} = 1$ and $R_{\max} = 5$. By applying the two constraints, the constants a's and c's can be found out. The three relationships are shown in Figure 3.6.

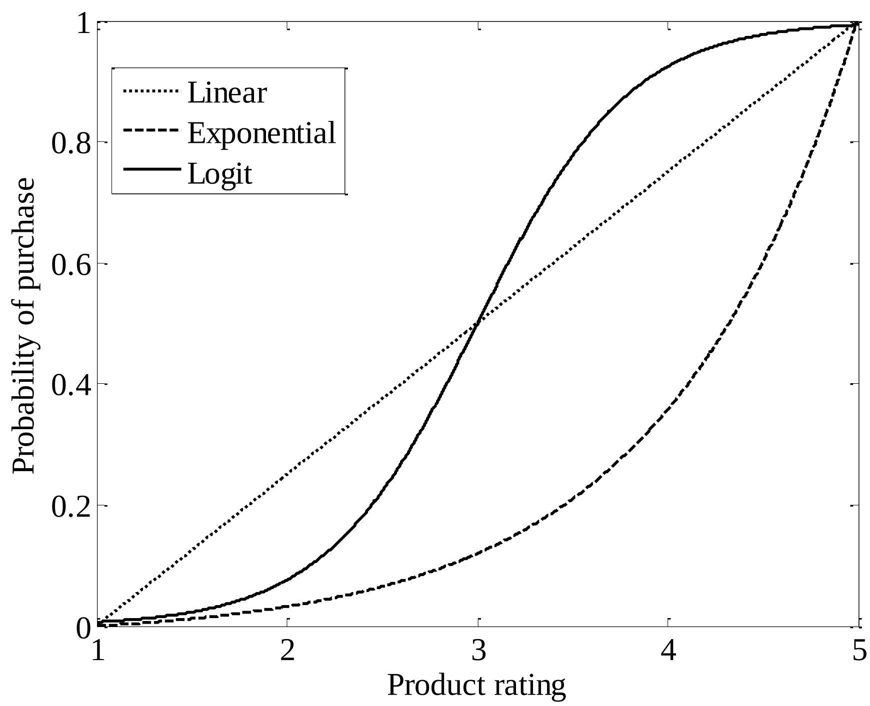


Figure 3.6 Three relationships between probability of purchase (*P*) and product rating (*R*)

Given the assumption that the demand is linearly dependent on the probability of purchase, the demand and profit can be estimated for the three forms. The design alternative – which satisfies (i) the maximum expected profit and (ii) the most insensitive to different relationships between the probability of purchase P and the product rating R – will be selected as the desired design. The second objective is considered because the actual relationship is unknown. We can only select the alternative which is insensitive over relationships. The alternative insensitive to different relationships means that the variation of the normalized profit over relationships for this alternative is small. The demand D_i and the expected profit PF_i for alternative i are obtained by

$$D_{(.),i}(\mathbf{E}_i, \boldsymbol{\xi}) = P_{(.),i}(\mathbf{E}_i, \boldsymbol{\xi}) \times MS \tag{3.11}$$

$$PF_{(.),i} = \int D_{(.),i}\left(\mathbf{E}_i, \boldsymbol{\xi}\right) \times \left(PC - C_i\right) \mathrm{d}\boldsymbol{\xi} \tag{3.12}$$

where the subscript (.) represents any relationship between the probability of product purchase and the product ratings (Equations (3.8)-(3.10)), MS is the potential market size, PC is the product price, C_i is the cost of the i^{th} design alternative.

As in Equation (3.12), the expected profit depends on product specifications $\boldsymbol{E}$ as well as the relationship between purchase probability P and the product rating R. To eliminate the effects of the relationships, the expected profit is replaced by the profit rank RK, which is equivalent to the normalized profit and used for design selection. In order to find the profit rank RK, the expected profit PF is sorted from the maximum to the minimum. The rank of the expected profit for each alternative i is recorded as RK_i. The alternative with larger profit has a smaller rank. For example, assume there are 30 design alternatives in total, the 5th alternative has the maximal profit and the 26th alternative has the minimal profit. Then $RK_5 = 1$ and $RK_{26} = 30$. For a design alternative i, there are three profit ranks because of three different relationships, $RK_{(1),i}$, $RK_{(2),i}$, and $RK_{(3),i}$. The mean and standard deviation of the ranks for each design alternative can be calculated as in Equations (3.13) and (3.14).

$$\mu_i\left(RK\right) = \left(RK_{(1),i} + RK_{(2),i} + RK_{(3),i}\right)/3 \tag{3.13}$$

$$\sigma_i\left(RK\right) = \left\{\left[\left(RK_{(1),i} - \mu_i\left(RK\right)\right)^2 + \left(RK_{(2),i} - \mu_i\left(RK\right)\right)^2 + \left(RK_{(3),i} - \mu_i\left(RK\right)\right)^2\right]/3\right\}^{1/2} \tag{3.14}$$

The design objectives – finding the maximal profit and the least sensitive to different relationships – are equivalent to minimize $\mu_i(RK)$ and $\sigma_i(RK)$. The bi-objective optimization problem is converted to a single objective problem as minimizing ($\mu_i(RK)$ + $\sigma_i(RK)$). The product design selection problem can be formulated as in Equation (3.15). The engineering constraints can be reliability, weight constraints and so on.

$$\begin{cases} \underset{i}{Minimize}\ \ \mu_i\left(RK\right) + \sigma_i\left(RK\right) \\ Subject\ to:\ \text{Engineering constraints} \end{cases} \tag{3.15}$$

3.4 Case Study

In this section, the proposed approach is applied to a smartphone design selection problem.

3.4.1 Online Customer Reviews of Interest

The public website considered for this case study is www.bestbuy.com. The customer reviews on this website (from which an example is shown in Figure 3.4) include product ratings, attribute aggregate ratings, general comments and pros/cons.

The four types of reviews are highlighted in Figure 3.7. The ratings are given in five scales from 1 to 5. The attribute aggregates are pre-determined by the website, with four aggregates – value for price, features, ease of use and battery life.

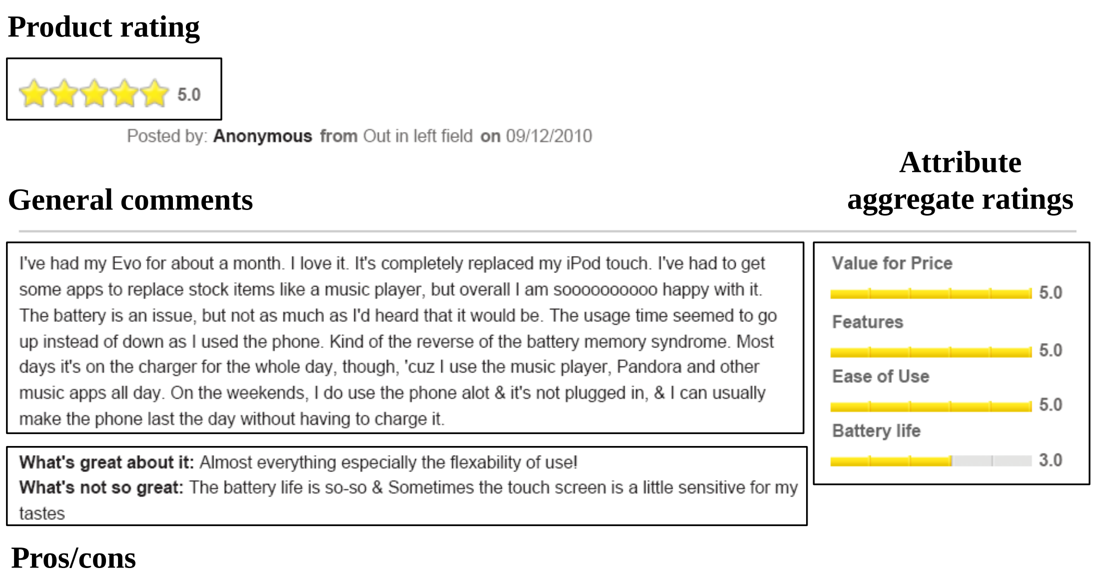


Figure 3.7 Highlighted customer reviews

Multiple observations from a single customer are collected for individual customers. Multiple observations indicate the ratings from the same customer for multiple (different) smartphones. On www.bestbuy.com, each customer who leaves the reviews has a user profile page, listing all the reviews the customer has performed. The user profile pages allow us to check whether the customers have written reviews for multiple smartphones. In total, 305 customer reviews were collected and downloaded from 143 customers. The reviews were written from June, 2010 to June, 2011.

3.4.2 Product Attributes Identification

The first step of the proposed methodology is to identify important product attributes and elicit customer ratings data. The pros and cons from the 305 customer reviews were divided into phrases as the inputs for text mining. A MATLAB toolbox – Text to Matrix Generator (TMG) [89] – was employed for this step. A list of the stop words is defined including those listed by MySQL – a popular open source database [43] plus the lists of standard adjective and adverb words [90,91]. After removing the stop words, only noun and verb words were left in the customer reviews. Next, the remaining words were stemmed and the frequent words were identified. A popular algorithm was used – association mining algorithm [42,52,92] – to identify frequent words as words/phrases appearing in more than 1% of the review phrases.

From the pros, nineteen frequent words were identified as candidate product attributes. Using the independency pruning rule [42], the candidate "*life*" was removed. From the cons, twelve frequent words were identified and two words 'life' and 'data' were pruned. In total, the retained frequent words identified from pros and cons are merged and identified 19 product attributes, as listed in Figure 3.8. After identifying product attributes, customer attribute ratings can be elicited by following the procedure in Section 3.3.1. Two ratings "attribute_pros" and "attribute_cons" for each attribute were elicited in two levels – 0 and 1. 0 indicates the attribute is not specified and 1 indicates the attribute is specified as a pro or con.

Applications	Battery	Network	Internet	Keyboard
Call quality	Camera	Design	Email	Feature
Feeling	Appearing	Memory	Price	Processor
Video	Quality	Screen	Size	Texting
Wi-Fi	HDMI	Spearker	OS	Ease of use

Figure 3.8 Identified product attributes

3.4.3 Model Estimation Results

The 305 sets of multiple observations data were used to estimate the customer preference models in a hierarchical structure. One data set consists of the ratings for one smartphone from one customer, including an overall rating R for the smartphone, attribute aggregate ratings $\mathbf{X}$ for four aggregates and attribute ratings $\mathbf{Y}$. The ordinal ratings R and $\mathbf{X}$ are treated as continuous variables. The attribute rating vector $\mathbf{Y}$ is a binary variable. As described in Section 3.3.2, the hierarchical structure of the model has three levels. The finite mixture regression for estimating the model was employed. For the n^{th} customer (n = 1, 2, …, 143), there is

$$\text{Level I: } R_n = \boldsymbol{\beta}_{g[n],\mathbf{X}} \cdot \mathbf{X}_n + \varepsilon_{\mathrm{I}} \tag{3.16}$$

$$\text{Level II: } \mathbf{X}_n = \boldsymbol{\beta}_{g[n],\mathbf{Y}} \cdot \mathbf{Y}_n + \boldsymbol{\varepsilon}_{\mathrm{II}} \tag{3.17}$$

$$\text{Level III: } \mathbf{Y}_n = \boldsymbol{\beta}_{g[n],\mathbf{E}} \cdot \mathbf{E}_n + \boldsymbol{\varepsilon}_{\mathrm{III}} \tag{3.18}$$

where $\varepsilon_{\mathrm{I}} \sim \text{Normal}(0, \sigma_{\mathrm{I}}^2)$ and $\boldsymbol{\varepsilon}_{\mathrm{II}} \sim \text{Multivariate Normal}(0, \boldsymbol{\sigma}_{\mathrm{II}}^2)$ represent the error terms – ε_{I} is a scalar, $\boldsymbol{\varepsilon}_{\mathrm{II}}$ is a 4×1 vector, Each element in $\boldsymbol{\varepsilon}_{\mathrm{II}}$ represents the measurement errors for each element in $\mathbf{X}$ respectively. $\boldsymbol{\varepsilon}_{\mathrm{III}}$ is a statistical error term – independent and identically distributed extreme values. $g[n]$ represents segment indicators, following a categorical distribution with the parameter vector $\mathbf{w} = \{w_s\}$ – the weights of segments, $\sum_s w_s = 1$.

The software – GLIMMIX – was employed for the segmentation purpose. GLIMMIX is designed for finite mixture regression modeling using an expectation-maximization algorithm [93]. Given the number of segments, GLIMMIX is able to provide the probability estimates for each segment, the probabilities of a respondent belonging to a segment, and other parameters estimates. In order to find the appropriate number of segments, a number of segments from 1 to 6 were considered and compared with respect to the performance of the model for different number of segments. Figure 3.9 plots the statistics of information criteria values (CAIC, MAIC and AIC) against the number of segments. According to Section 2.3, a better-fitted model has a smaller information criterion value. Therefore, the model with four segments outperforms because all the information criteria values are smallest when the number of segments is four. It was thus decided to assign four segments for the models. The weights for each segment are w_1=0.28, w_2=0.22, w_3=0.28 and w_4=0.22.

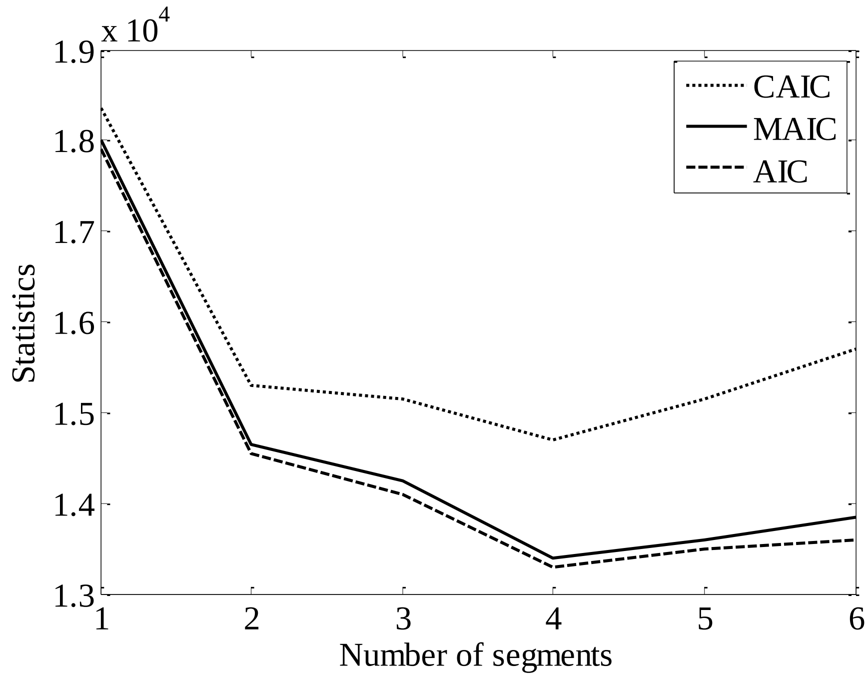


Figure 3.9 Plot of statistics against the number of segments

The 305 sets of multiple observations data were used to estimate the customer preference models in a hierarchical structure. WinBUGS [94], a widely-used software tool for Bayesian Markov Chain Monte Carlo (MCMC) is used for estimating the parameters. The criterion to select the attributes follows the widely used criterion in the literatures [16,41] – according to the frequency of the attributes. The ten most frequently mentioned attribute_pros and attribute_cons are selected for the case study. The product specifications are selected from the common product specifications available at www.bestbuy.com and www.phonescoop.com. Twenty-one product specifications are chosen for this case study. Table 3.1 lists the detailed descriptions for attribute aggregates **X**, attributes **Y** and product specifications **E** that were considered.

Table 3.1 Description of attributes and specifications

Attribute aggregates **X**:	
X_1 – Value for price;	X_2 – Features;
X_3 – Ease of use;	X_4 – Battery life;
Attributes **Y**:	
Attribute_Pros	Attribute_Cons
Y_1 – OS	Y_{11} – OS
Y_2 – Applications	Y_{12} – Applications
Y_3 – Battery life	Y_{13} – Battery life
Y_4 – Camera	Y_{14} – Call quality
Y_5 – Ease of use	Y_{15} – Camera
Y_6 – Features	Y_{16} – Keyboard
Y_7 – Keyboard	Y_{17} – Price
Y_8 – Screen	Y_{18} – Screen
Y_9 – Processor	Y_{19} – Processor
Y_{10} – Design	Y_{20} – Quality
Product specifications **E**:	
E_1 – Network variable (1=4g, 0=Not 4g);	
E_2 ~ E_5 – dummy variables for OS ([1,0,0,0]=OS1, [0,1,0,0]=OS2, [0,0,1,0]=OS3, [0,0,0,1]=OS4, [0,0,0,0]=OS5);	
E_6 – height (inch);	E_7 – width (inch);
E_8 – depth (inch);	E_9 – weight (ounce);
E_{10} – display size (the diagonal length of a display screen);	
E_{11} – total pixel resolution (defined $w_r \times h_r$, w_r and h_r are width and height resolution in pixel respectively);	
E_{12} – touch screen (1=Yes, 0=No);	
E_{13} – battery capacity (mAh)	
E_{14} – camera resolution (mega-pixel);	
E_{15} ~ E_{16} – video variables ([1,0] = high-definition video, [0, 1] = regular definition video, and [0, 0] = no video);	
E_{17} – processor variable (1= "processor speed $\geq$ 800 MHz", 0= "processor speed < 800 MHz");	
E_{18} – memory variable (1= "memory $\geq$ 1 GB", 0="memory < 1 GB").	
E_{19} – phone form (1="slide form", 0="bar form").	
E_{20} – physical keyboard (1=Yes, 0=No).	
E_{21} – Wi-Fi variable (1=Yes, 0=No).	

The parameters for each segment were estimated using WinBUGS. Table 3.2–Table 3.4 list the statistical results (means (standard deviations)) for the parameters estimated for the three levels in the model. The 95% confidence intervals (CIs) of the parameters are calculated. If the bounds of 95% CIs are on the same side of zero (both positive or negative), then the variables have purely positive or negative effects on the dependent variable, which are called significant variables and highlighted as bold in the tables.

For Level I, Table 3.2 lists the parameters for all four segments. The positive means of the parameters suggests that all the four attribute aggregates have positive effects on the product rating. Not all the effects are significant. For Segment 2, only the attribute aggregates “value for price” and “ease of use” have significant positive effects. For Segment 3, the aggregates X_1 “value for price”, “features” and “ease of use” affect the product rating R significantly. For Segments 1 and 4, all the four aggregates are significant. The most important attribute aggregates (the one with the largest parameter means) for each segment are “value for price/battery life”, “value for price”, “value for price”, and “value for prices”, respectively. Different parameter estimates in different segments imply the existence of different customer preferences.

Table 3.2 Model estimated results for Level I

Level I – parameter estimates for product rating *R*				
Attribute aggregates	Segment 1	Segment 2	Segment 3	Segment 4
Value for price	**0.40(0.08)**	**0.63(0.11)**	**0.36(0.08)**	**0.55(0.08)**
Features	**0.33(0.07)**	0.07(0.09)	**0.33(0.08)**	**0.32(0.08)**
Ease of use	**0.28(0.08)**	**0.35(0.10)**	**0.30(0.10)**	**0.29(0.09)**
Battery life	**0.40(0.06)**	0.08(0.07)	0.08(0.06)	**0.27(0.06)**
Constant	**-1.76(0.36)**	**-0.57(0.34)**	-0.37(0.26)	**-1.71(0.37)**

mean(standard deviation)

Table 3.3 for Level II lists the results for the attribute aggregate "Features" for all the four segments as an example. Overall, the parameters for attribute_pros are positive – indicating that more people specify the attributes as pros, the higher rating for "Features". In contrast, the parameters for attribute_cons are generally negative – representing that more people specify the attributes as cons, the lower rating for "Features". Most attributes have consistent effects across different segments and follow the general trends, such as "Processor_pros", "App_cons". However, some attributes have conflicting effects for different segments. For example, the parameters for "call quality_cons" are significantly negative for Segment 2 but significantly positive for Segment 3. The conflicting effects for different segments can be explained by customer preference heterogeneity. For customers in Segment 2, "call quality" is an attribute affecting the attribute aggregate "Features". If they dislike the call quality, then the rating for "Feature" is low. For customers in Segment 3, "call quality" does have not positive effects on the attribute aggregate "Features". Although they dislike the call quality, the rating for "Features" can be still high. Additionally, in the points of view from the customers in Segment 3, the smartphone with good features tends to have poor call quality.

Table 3.3 Model estimated results for Level II

Level II – parameter estimates of significant variables for the attribute aggregate "Features"

Attributes	Segment 1	Segment 2	Segment 3	Segment 4
OS_pros	-0.05(0.39)	**0.69 (0.35)**	**1.07 (0.38)**	0.12(0.43)
Apps_pros	**0.91 (0.26)**	0.60 (0.35)	0.10(0.32)	-0.78(0.90)
Battery_pros	**0.76 (0.34)**	-0.44(0.50)	**0.69 (0.33)**	0.32(0.46)
Ease of use_pros	0.31(0.25)	**0.99 (0.30)**	-0.67(0.40)	0.39(0.39)
Features_pros	**0.63 (0.30)**	0.49(0.33)	**1.13(0.52)**	**1.23 (0.47)**
Screen_pros	0.41(0.27)	**0.83 (0.32)**	0.54 (0.31)	0.24(0.38)
Processor_pros	0.46(0.37)	**0.82 (0.36)**	**1.03 (0.30)**	0.34(0.54)
Design_pros	0.41(0.37)	-1.08(0.62)	**0.67 (0.30)**	0.14(0.44)
App_cons	**-1.44 (0.45)**	-0.70(0.53)	-0.46(0.55)	-1.53(0.96)
Battery_cons	**-0.53(0.25)**	-0.04(0.32)	**1.23(0.35)**	0.77(0.42)
Call quality_cons	-0.14(0.49)	**-1.78 (0.63)**	**1.16(0.58)**	-0.24(0.68)
Price_cons	-0.47(0.68)	0.05(0.52)	-1.82 (0.94)	0.55(0.63)
Screen_cons	-0.57(0.54)	0.91(0.51)	-0.72(0.40)	-0.17(0.44)
Processor_cons	-0.18(0.40)	0.79(0.50)	**-1.37 (0.58)**	**-1.19 (0.58)**
Quality_cons	-0.07(0.44)	-0.39(0.49)	0.79(0.46)	-0.83(0.70)
Constant	**3.58 (0.28)**	**3.40 (0.30)**	**2.84 (0.35)**	**3.42 (0.38)**

mean(standard deviation)

The results for Level III in Table 3.4 are for the attribute Y_2 "Applications_pros" for all the segments. The application of a smartphone is mainly determined by the operating systems. The results suggest that at least one OS has a significant effect on the attribute "application_pros" ratings except Segment 1. For the OS specification, OS 5 ([0,0,0,0]) is a base value. The parameters indicate that compared to the base OS (OS 5), customers in Segment 2 dislikes the applications of OS 1 but like that of OS 2; customers in Segments 3 like the applications of OS 1 to 3; and customers in Segment 4 dislike the applications of OS 2. Customers across the segments seem to be indifferent between the applications of OS 4 and OS 5 because OS 4 is not significant to any segment (as a result, OS 4 is not listed as a significant variable in Table 3.4). Some other specifications have effects on "Application_pros" as well, such as height, Wi-Fi and so on.

Table 3.4 Model estimated results for Level III

Level III – parameter estimates for the attribute "Application_pros" Y_2

Specifications	Segment 1	Segment 2	Segment 3	Segment 4
OS 1	0.28(2.51)	**-3.95(2.08)**	**6.56(2.30)**	0.94(3.40)
OS 2	3.78(2.34)	**6.45(2.36)**	**6.27(2.49)**	**-6.57(2.87)**
OS 3	-2.01(5.05)	-3.16(3.60)	**6.30(1.94)**	-4.61(4.01)
Height	**-3.86(2.00)**	0.10(2.41)	**6.71(2.05)**	0.02(3.59)
Width	4.42(3.09)	**-5.41(2.84)**	**-5.40(2.18)**	**-3.78(2.10)**
Display size	**-5.32(2.28)**	1.45(1.67)	0.77(1.39)	3.14(2.34)
Resolution	0.09(0.78)	0.81(3.39)	**8.08(1.60)**	-0.28(4.06)
Touch	6.41(2.43)	**0.95(0.52)**	**-3.65(1.47)**	0.30(1.79)
Battery capacity	5.89(3.53)	**-6.21(3.04)**	**5.65(2.57)**	0.26(5.61)
video	-0.22(4.87)	**1.62(0.76)**	-1.12(0.65)	**-2.61(1.23)**
video	0.95(4.59)	**6.03(3.01)**	3.92(4.52)	1.57(5.02)
Processor	1.61(1.70)	3.25(3.18)	2.82(3.62)	**-6.64(2.50)**
Memory	1.98(1.64)	**-8.01(1.61)**	-1.82(1.68)	-0.42(5.51)
Phone form	**-5.49(2.40)**	-2.10 (2.08)	0.16(2.14)	4.25(2.62)
keyboard	3.46(2.81)	-1.20(3.72)	**-7.39(1.74)**	**-6.68(2.67)**
Wi-Fi	**4.27(1.79)**	-1.97(3.73)	**7.00(2.08)**	**-5.82(2.90)**

mean(standard deviation)

The variance σ^2_{I} for the error terms ε_{I} and the covariance matrix $\boldsymbol{\sigma}^2_{\mathrm{II}}$ for the error term $\boldsymbol{\varepsilon}_{\mathrm{II}}$ are listed in Equations (3.19) to (3.20). The diagonal elements of the covariance matrix represent the variances of each element (attribute aggregate) and the off-diagonal elements represent the covariance between two different elements. The variances of error terms $\boldsymbol{\varepsilon}_{\mathrm{II}}$ (diagonal elements of $\boldsymbol{\sigma}_{\mathrm{II}}^2$) are greater than the variance of the error term ε_{I} (σ_{I}^2), indicating a larger statistical error in Level II than that in Level I. Given the values of off-diagonal elements of $\boldsymbol{\sigma}^2_{\mathrm{II}}$ are from 0.21 to 0.36, the values implies that the covariance between any two attribute aggregates is relatively small.

$$\sigma_{\mathrm{I}}^2 = 0.30 \tag{3.19}$$

$$\boldsymbol{\sigma}_{\mathrm{II}}^{2}=\begin{bmatrix}0.87 & 0.36 & 0.35 & 0.24\\ 0.36 & 0.72 & 0.31 & 0.21\\ 0.35 & 0.31 & 0.64 & 0.30\\ 0.24 & 0.21 & 0.30 & 0.88\end{bmatrix} \tag{3.20}$$

The pseudo-r^2 values, *RMSE* and *NRMSE,* are calculated to quantify the predictive capability of the models at three levels, as listed in Table 3.5.

Table 3.5 Model evaluation for each level

	Level I	Level II	Level III
Pseudo-r^2	0.81	0.55	0.35
RMSE	0.54	0.76	0.29
NRMSE	0.14	0.19	0.29

The pseudo-r^2 value for the model of Level I is fairly high while the pseudo-r^2 values for Levels II and III are relatively lower. The pseudo-r^2 values for Levels II and III are considered to be acceptable considering the nature of subjective data and the mappings from textual reviews into numerical values. Especially Level III is a binary logistic regression, the pseudo-r^2 value for such regression is normally low and any value between 0.2 to 0.4 is usually considered as a good fit [95,96]. In short, the pseudo-r^2 values for the models indicate that the models developed from online reviews can explain customer preferences for smartphones reasonably well. It should be noted that the ratings for the three levels are not in the same scale. For levels I and II, the ratings are in a five-point scale (1, 2, …, 5); for level III, the ratings are binary (0 and 1). Instead of an *RMSE* measure, an *NRMSE* measure must be used for understanding the correctness degree of model predictability. It is observed that, as the level of the model is lower, a pseudo-r^2 value decreases while an *NRMSE* increases. It makes sense well that a pseudo-r^2 value is inversely proportional to a model prediction error.

A set of 50 reviews which are not used for estimation is used for model validation. The reviews are processed by following the procedures in Section 3.3. The product specifications are used to predict the attribute ratings, attribute aggregate ratings and product ratings. The errors between the predicted ratings and the actual ratings are quantified in the following forms – the mean absolute percentage error *MAPE*, mean absolute error *MAE* and *RMSE*. For Level III, the actual attribute rating y is a binary variable equal to 0 or 1. As defined in Equation (2.21) for *MAPE*, the actual rating y is a denominator. For the case that y = 0, *MAPE* is meaningless. Thus, *MAPE* is not calculated for Level III**.** The validation results are listed in Table 3.6.

Table 3.6 Model validation for each level

	Level I	Level II	Level III
MAPE	0.13	0.13	N/A
MAE	0.58	0.54	0.28
RMSE	0.64	0.68	0.31

The *MAPE* and *MAE* values are fairly acceptable. The *RMSE* values from validation are close to *RMSE* from estimation for each level, which implies that the model is validated using the set of out-of-sample data.

3.4.4 Product Design Selection

In this section, it is shown how the customer preference models from the online customer reviews are used for a smartphone design selection problem. The first step is to define design variables and generate design alternatives. The 21 discrete design variables are defined in Table 3.7.

Table 3.7 Design variables

Design variables	Physical meanings & Possible values
E_1	4g network (= 1); No 4g network (=0)
E_2 to E_5	OS 1(=[1,0,0,0]), OS 2(=[0,1,0,0]), OS 3 (=[0,0,1,0]), OS 4 (=[0,0,0,1]), OS 5 (=[0,0,0,0])
$[E_6, E_7, E_8, E_9, E_{10}]$	Large size (=[4.7, 2.4, 0.5, 5.1, 3.7]) Small size (=[4.3, 2.2, 0.5, 4.3, 2.8])
E_{11}	High resolution (=800×480) Low resolution (=480×320)
E_{12}	Touch-sensitive (= 1); Non touch-sensitive (= 0)
E_{13}	Battery capacity (1420, 1200 mAh)
E_{14}	Camera resolution (5, 3.2-Megapixel)
E_{15}	High-definition video (= 1); otherwise (= 0)
E_{16}	Regular-definition video (= 1); otherwise (= 0)
E_{17}	Processor speed on video12(= 1); otherwise (= 0)
E_{18}	Memory ≥ 1 GB(= 1); otherwise (= 0)
E_{19}	Slide form (= 1); Bar form (= 0)
E_{20}	W/ physical keyboard (= 1); W/o physical keyboard (= 0)
E_{21}	W/ Wi–Fi (= 1); W/o Wi–Fi (= 0)

Among the nine engineering design constraints considered for this problem, two are: LCD resolution (1st constraint) in Equation (3.21) and smartphone weight (2nd constraint) in Equation (3.22) [97]. The remaining seven are the logical decision constraints in Equations. (3.23) to (3.29): 3rd constraint – sliding phone must have a physical keyboard; 4th constraint – only the camera with 5 mega-pixel lens can provide the option of high-definition video; 5th constraint – high-definition video and regular definition video cannot exist at the same time; 6th constraint – a phone can only equip with one operating system; 7th constraint – a sliding phone is 1 inch deeper than a bar phone without a physical keyboard in the same dimension; 8th and 9th constraints – a bar phone with a physical keyboard has a 1 inch shorter display size and 0.05 inch deeper than the bar phones without a physical keyboard in the same dimension because the physical keyboard on the same surface makes the screen smaller. Note that the difference of the dimension in the 7th to 9th constraints are estimated based on the dimension changes in the smartphones existing on the market.

$$E_{11} \leq 7.37 \times 10^{4} \times E_{10}^{2} \tag{3.21}$$

$$5.1 \times 10^{-4} \times E_6 \times E_7 \times E_8 \leq E_9 \tag{3.22}$$

$$E_{19} = 1 \text{ iff } E_{20} = 1 \tag{3.23}$$

$$E_{15} = 1 \text{ iff } E_{14} = 5 \tag{3.24}$$

$$E_{15} \times E_{16} = 0 \tag{3.25}$$

$$E_2 + E_3 + E_4 + E_5 \leq 1 \tag{3.26}$$

$$E_8 = E_8 + 1 \text{ if } E_{19} = 1 \tag{3.27}$$

$$E_8 = E_8 + 0.05 \text{ if } E_{19} = 0 \text{ and } E_{20} = 1 \tag{3.28}$$

$$E_{10} = E_{10} - 1 \text{ if } E_{19} = 0 \text{ and } E_{20} = 1 \tag{3.29}$$

The next step is to estimate the cost for each design. The cost is estimated as the summation of the component costs. The assembly cost is ignored. The component costs are estimated according to References [98-102] listed in Table 3.8. The estimated minimum cost is $98.65 and the maximum cost is $179.91.

Table 3.8 Smartphone component cost estimation ($)

Components		Cost estimation ($)
Network	4g	15
	No 4g	10
OS	1	15
	2	12
	3	12
	4	12
	5	10
Shell		$2.2\times E_7\times E_8\times E_9$
Screen		$0.0000339\times E_{11}+10$ $0.000339\times E_{11}+30$ if touch sensitive
Battery		$0.0045\times E_{13}–6.9$
Camera lens		$2.77\times E_{14}-3.8$
Processor		$3\times E_{17}+15$
Memory		$15\times E_{18}+10$
Phone form		$10\times E_{19}$
Physical keyboard		$5\times E_{20}$
Wi–Fi		$8\times E_{21}$

The final step is to calculate the profit for each alternative and select the design satisfying the objectives. Based on the design constraints, in total, 6400 design alternatives can be generated. The parameters estimated from online customer reviews are used to predict the product ratings for each design alternative. Assuming there are three prices possible for the design alternatives – $99.99, $199.99, and $299.99. The three prices are set according to the low-, middle- and high-end smartphone markets. The potential market size (*MS*) for this design is assumed to be 1,000,000. The customers from www.bestbuy.com are assumed to be representative of the whole market and thus it is assumed that the potential market has the same four segments as the customers from Bestbuy.com (w_1=0.28, w_2=0.22, w_3=0.28 and w_4=0.22). As described in Section 3.3.3, three relationships between the probability of purchase and the predicted product ratings are assumed as – linear, logit and exponential. The expected profit for each alternative can be calculated and sorted in a descending order to find the profit ranks. The means and standard deviations of the profit ranks for each alternative are calculated as in Equations (3.13) and (3.14). The figures of the means and standard deviations of the profit ranks are plotted under each price in Figure 3.10. It can be observed that the figures are in a spindle shape. The shape suggests that the variations for the ranks at the two ends are small – implying that the design alternatives with highest or lowest ranks, that is highest or lowest profits, are insensitive to different possible relationships. This is because the probabilities of purchase for the three relationships are close to each other when the ratings are high or low, as represented in Figure 3.6, the three curves are close at two ends. It indicates that higher/lower ratings yield to larger/smaller probabilities of purchase regardless of relationships. Since profit/demand is linearly dependent on the probability of purchase, the profit is also insensitive to relationships when the ratings are high or low.

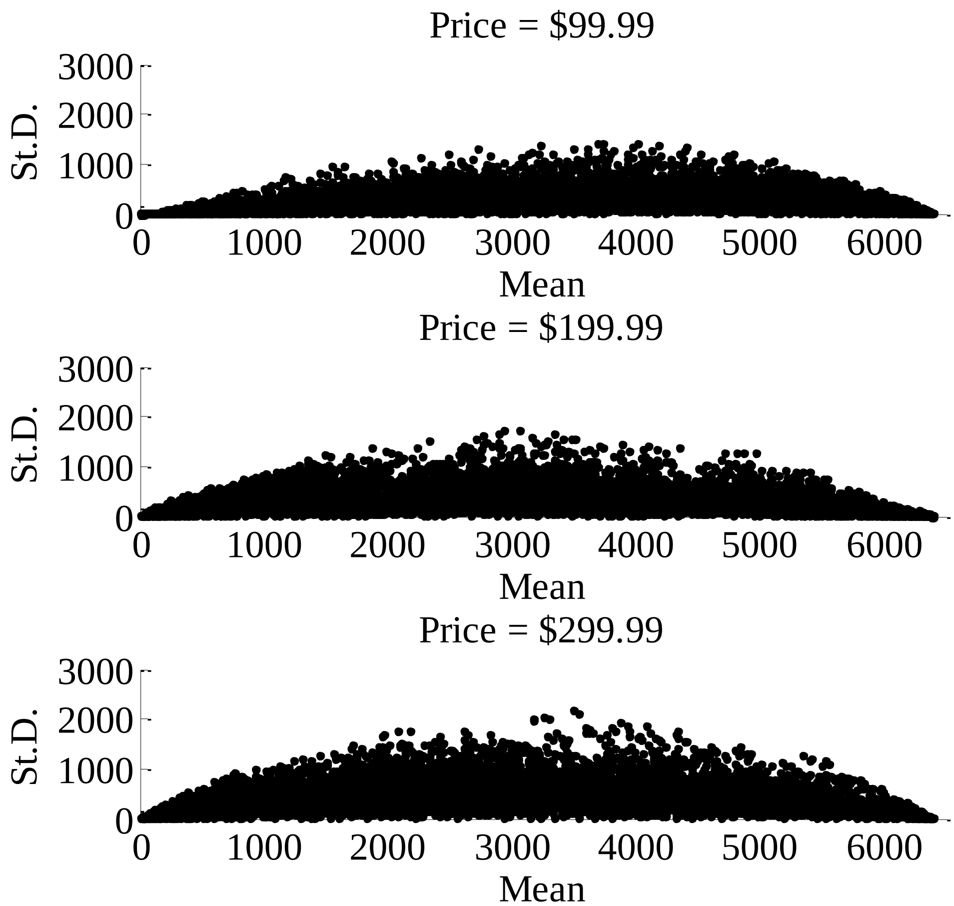


Figure 3.10 Mean and standard deviations of profit ranks

The design alternative with the smallest mean of profit ranks has the smallest deviation of ranks as well. Therefore, the alternative is selected as the desired design. The design selection results for each price are listed in Table 3.9. Given different prices, the design results are consistent in most design variables. As the price increases, the network is suggested to be upgraded from non-4g to 4g. The battery capacity increases and Wi-Fi is added as well. The change of design implies that the smartphone with the advanced equipment (e.g., 4g, Wi-Fi) for high-end markets yields to larger profit.

Table 3.9 Design results

	Price	$99.99	$199.99	$299.99
Design variables	Network (E_1)	No 4g	4g	4g
	OS (E_2 ~E_5)	5	2	2
	Size (E_6 ~E_{10})	Small	Small	Small
	Screen resolution (E_{11})	Low	Low	Low
	Touch (E_{12})	Non-touch	Non-touch	Non-touch
	Battery capacity (E_{13})	1200	1420	1420
	Camera resolution (E_{14})	3.2	3.2	3.2
	Video (E_{15} ~ E_{16})	Regular	Regular	Regular
	Processor (E_{17})	>800MHz	>800MHz	>800MHz
	Memory (E_{18})	<1GB	<1GB	<1GB
	Phone form (E_{19})	Bar	Bar	Bar
	Physical keyboard (E_{20})	Yes	Yes	Yes
	Wi-Fi (E_{21})	No	No	Yes

3.5 Summary

In this chapter, a new methodology is proposed for customer-driven product design selection by using web-based online customer reviews. In the methodology, the existing text mining techniques are extended in order to identify product attributes and elicit customer preference data from customer reviews. The finite mixture regression model was employed for modeling the customer data from customer reviews. The use of the finite mixture regression enables modeling unobserved customer preference heterogeneity in customer data from customer reviews. Finally, the customer model developed from customer reviews is used for product design selection problem – select a product design alternative that maximizes the profit and is relatively insensitive to different possible relationships between the probability of purchase and the product rating.

The work presented in this chapter makes contributions by: (i) overcoming a major limitation in existing customer-driven product design selection methods which can significantly decrease the cost and time required in the acquisition of customer data; (ii) extending and integrating existing techniques to overcome the disadvantages of customer reviews and ultimately use customer reviews for product design selection. For demonstration, the proposed method was applied to a smartphone case study. The results show that by making use of web based customer reviews, the proposed method can elicit product attributes of customers' interests; develop the customer preference model from web-based customer reviews while accounting for the heterogeneity of customer preferences; and finally use the model for customer-driven product design selection. The entire process is purely a customer-driven process and based on a free data source – web-based customer reviews.

In next chapter, online customer reviews from multiple websites will be processed and integrated for product design selection.

Chapter 4: Product Design Selection Using Online Customer Reviews from Multiple Websites with Website Heterogeneity

In the last chapter, an approach in product design selection using online reviews collected from a single website was presented. However, online customer reviews from a single website may not be a good representative of customer data in a target market. The consideration of online reviews from multiple websites for product design selection is the subject of this chapter. The material in this chapter can be beneficial in two ways: (i) online reviews from multiple websites is more representative of the market compared to the reviews from a single website; and (ii) multiple websites might be necessary to get sufficient amount of data especially for product aggregates that have limited data from just a single website. Furthermore, the heterogeneity of online reviews across different websites is too significant to be ignored. Motivated by these reasons, this chapter proposes an approach of eliciting and processing online customer data from multiple websites, and integrating customer data by using a meta-analysis technique.

Section 4.1 gives an introduction. Section 4.2 defines the problem and describes the assumptions. Section 4.3 describes the proposed methodology for product design using online reviews from multiple websites. Section 4.4 presents a smartphone case study that demonstrates the applicability of the proposed method. Section 4.5 gives some concluding remarks.

4.1 Introduction

This chapter makes the following two contributions for elicitation and integration of customer reviews.

The first contribution is a text classification method to elicit product attributes and customer data from multiple websites regardless of their own formats. At the early stage, online customer reviews were freely written textual reviews only (called general comments in this chapter). As the popularity of online customer reviews increased, the format of customer reviews improved significantly. Instead of general comments, some public websites request customers to summarize the pros/cons for a product. Current research [16,41,42,47-50] effort has been made to identify product attributes from either general comments or the pros/cons from a single website. In fact, some researchers have found [16,41] that the text mining results from general comments are significantly worse than the results obtained using the pros/cons. Nevertheless, the pros/cons summary of a product is not a standard format across different websites. Due to the lack of pros/cons on some websites, current methods focusing on the pros/cons summary from a single website cannot be extended to multiple websites. This chapter proposes a text classification method to elicit product attributes and customer attribute ratings from multiple websites in three steps. First, product attributes and customer attribute ratings are elicited from the websites with the pros/cons: The procedure here is the same as our previous work, as described in Chapter 3 [103]. Second, product attributes and customer attribute ratings are elicited from general comments from the websites without the pros/cons by using the supervised classification method SVM (support vector machine). The attribute sentences are elicited and classified into the positive group (pros) and the negative group (cons) by using the pros/cons from the websites with pros/cons. Here, the difference between our proposed approach and previous literatures [47-50] is the choice of training datasets when using SVM. As discussed in the previous literature [47-50], researchers manually classify general comments into pros or cons, as training datasets. In our approach, the pros/cons from the websites with pros/cons are used as training datasets. Thus, using these pros/cons as training sets can significantly reduce the requirement of human work. In the final step, the product attributes considered for product design are identified from all the websites of interest.

The second contribution is the use of a meta-analysis technique for integrating customer review data from multiple websites. This technique was first proposed in the area of health sciences [70] and has been widely applied for psychological [104], educational [105], social science [106], and marketing research fields [107] as well as others [71-73]. The meta-analysis technique is a statistical process to combine results from multiple studies and has not been applied for combining online reviews from multiple websites. It should be noted that different websites may require different procedures to obtain customer reviews with their dissimilar review formats. The heterogeneity of websites makes it difficult to combine customer review data from multiple websites. The meta-analysis technique can provide a feasible solution to integrate customer data from multiple websites and reconcile potential differences among them.

4.2 Problem Definition and Assumptions

As shown in Figure 4.1, a designer seeks a design for a consumer durable product by using online reviews from multiple websites. The desired design should satisfy the engineering constraints and meanwhile maximize the profit for a manufacturer as well. Two types of heterogeneity may exist in the customer data from multiple websites. The first is called website heterogeneity in this dissertation. It is mainly caused by different formats of the websites. The other is called customer preference heterogeneity, which is caused by different customer preferences for the products. In order to keep the heterogeneity simple and clearly observe its effect, in this chapter customer preference heterogeneity is ignored. Other assumptions made in this chapter include: (i) online reviews from multiple websites is a good representative of customer voices; and (ii) multiple observations from a single customer are regarded as independent.

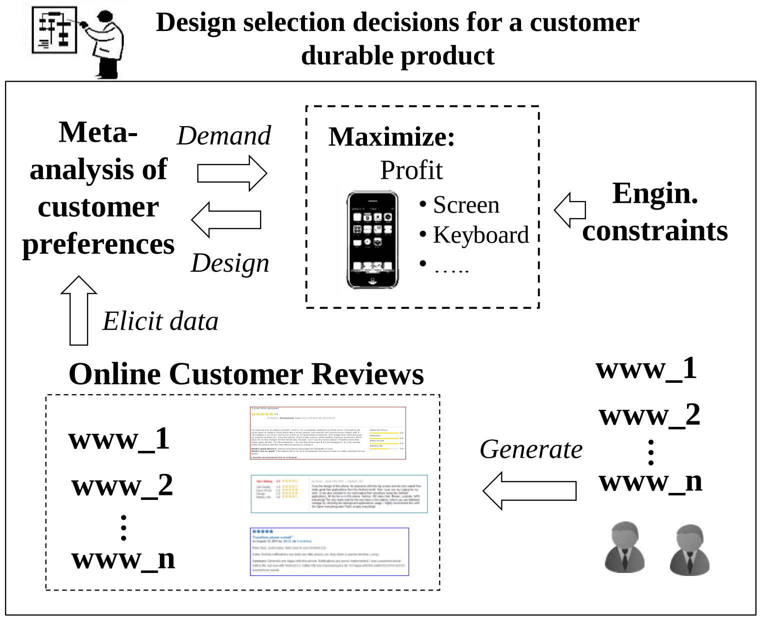


Figure 4.1 Problem definition

4.3 Approach

The proposed method, as shown in Figure 4.2, involves four tasks: (Task 1) collecting online reviews from multiple websites; (Task 2) eliciting product attributes and customer data from multiple websites; (Task 3) meta-analysis of hierarchical customer preference models; and (Task 4) selecting a product design. In Task 2, a text classification method is proposed to elicit product attributes and customer data from multiple websites. The attribute-related information is processed and elicited from the pros/cons from available websites. The information is later used for eliciting product attributes and customer data from general comments in other websites. To take into account the website heterogeneity, Task 3 applies a meta-analysis technique for integrating customer data from multiple websites and constructing customer preference models. In Task 4, a profit function is formulated using the customer preference models. The profit is incorporated as a design objective for product design selection. The four tasks are detailed in the next four subsections. As reviewed in Section 2.4.2, there are two main differences among multiple websites – attribute aggregates and pros/cons. The differences in pros/cons are to be handled during text mining (Task 2) and the differences of attribute aggregates during model estimation (Task 3).

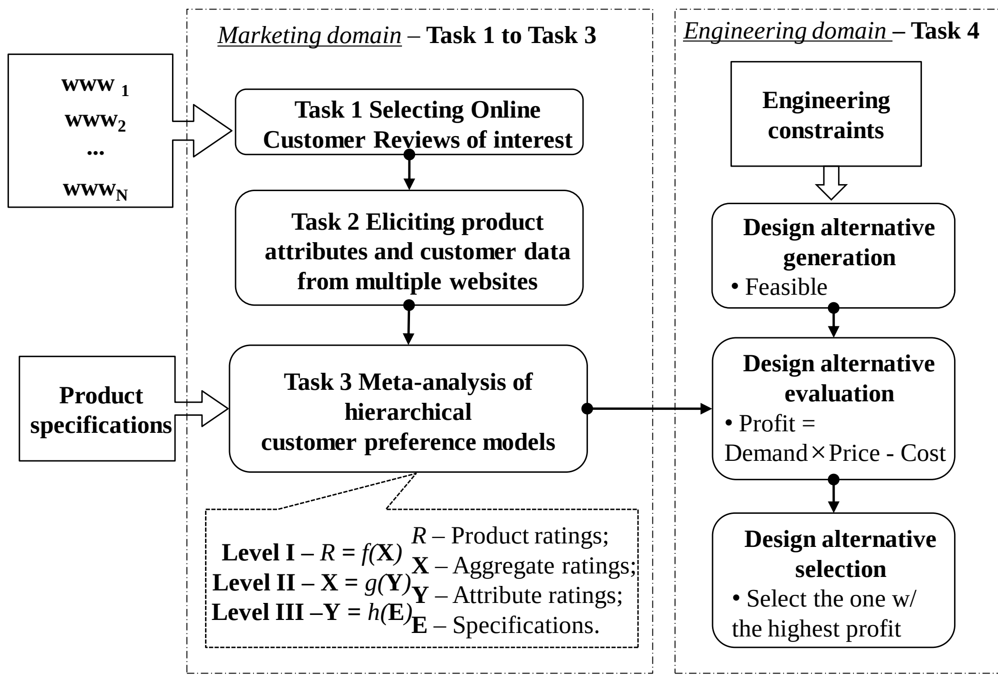


Figure 4.2 Overall framework of the proposed methodology

4.3.1 Collecting Online Customer Reviews of Interest

Recall Section 2.6 where meta-analysis was described as a technique to integrate customer review data from multiple medical clinics. In this chapter, a clinic can be thought of as a public website. Customer reviews are collected from a clinic (website). Treatments are equivalent to the different product specifications being rated by customers and a treatment effect refers to the effect of different product specifications as to how customers rate the product. The first step of meta-analysis is to compare and select the clinics (websites). Recall Section 2.4.2, customer reviews from the public websites are very similar to each other. For example, they include both numerical ratings (product ratings) and textual reviews (general comments). These similarities among the websites make the applications of meta-analysis feasible.

4.3.2 Eliciting Product Attributes and Customer Data from Multiple Websites

The formats of customer reviews from different websites are similar but not identical. As stated in Section 2.4.2, there exist non-ignorable differences among different websites, such as whether the websites collect attribute aggregate ratings and whether the websites collect pros/cons. Based on the differences of the pros/cons, multiple websites are grouped into three sets. The first set is denoted as Website I in which the pros/cons are freely written by customers. The second set (Website II) is websites in which they provide a checklist for pros/cons. Thus the pros/cons in Website II are guided by the websites, not freely expressed by customers. The third set (Website III) refers to those websites that are without pros/cons.

The goal of this task is to elicit product attributes and customer data from multiple websites. Product attributes and their orientation (positive (pros) or negative (cons) evaluation) can be elicited with a least amount of manual work by using the pros/cons data. As shown in Figure 4.3, Task 2 contains three major steps – identifying attribute candidates, determining attribute orientations and the attributes from multiple websites. The first step identifies a pool of product attributes (or product attribute candidates) from Website I, which include the product attributes used for product design. Two kinds of dictionaries, to be constructed from Website I, are used to elicit attributes and customer data from Websites II and III. Dictionary I is called attribute dictionary including the words used to indicate product attributes, such as "screen", "quality". Dictionary II is called attribute support dictionary including the words used to evaluate attributes such as "great", "poor", "like". The next step is to determine the orientations of attributes – whether the attributes are evaluated as pros or cons. The orientations of attributes in Websites I and II can be determined according to whether the customers specify attributes as pros or cons. The orientations of the attributes in Website III are determined by a supervised classification (or orientation) of the attributes. In the final step, the most frequently mentioned attributes among the pool of the attributes are determined as the final product attributes used for product design selection. These steps are detailed in the subsections below.

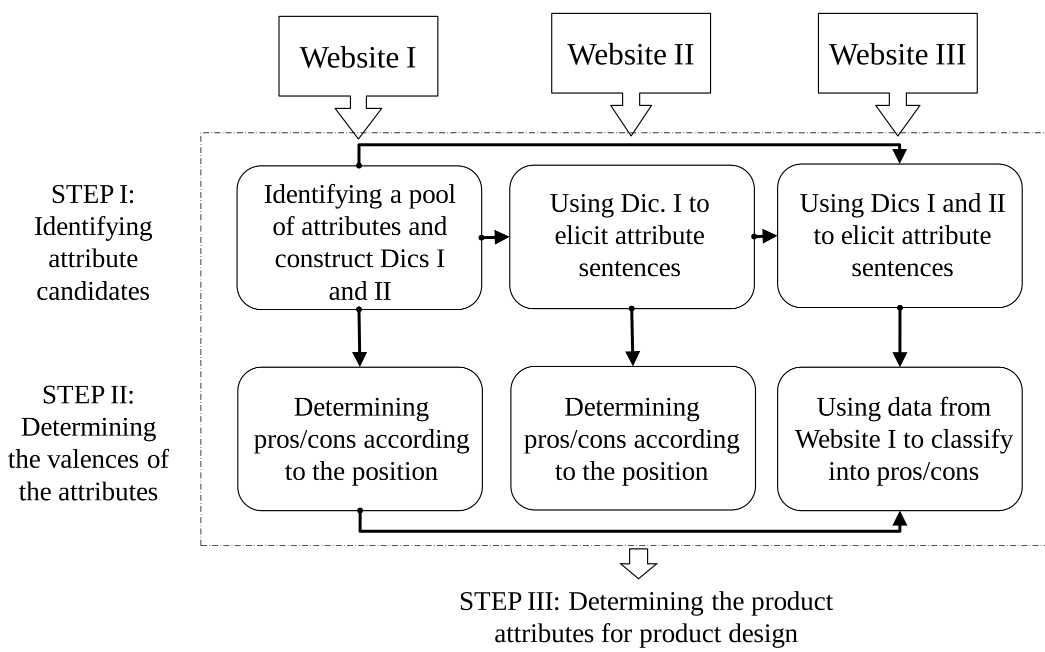


Figure 4.3 Steps in Task 2

4.3.2.1 Eliciting Product Attribute Candidates

Based on the previous literature [16,41,42,47-56], product attributes can be identified from either pros/cons or general comments. It is known that all the websites have general comments. Thus, the most direct way to identify attributes should be ignoring pros/cons and identifying attributes from general comments in multiple websites. However, this work may require a lot of human work, prior-knowledge and sophisticated text mining techniques. Meanwhile, the pros/cons statements in the customer reviews cannot be ignored because they summarize comments with customer-specified orientations (pros – positive orientation; cons – negative orientation). It is observed that not all the pros/cons are freely expressed by customers, i.e., a checklist of pros/cons pre-specified by the websites as in Website II. Since the pros/cons summary in Website II are not customer specified language, only the pros/cons in Website I should be employed for identifying attributes and constructing dictionaries. As shown in Figure 4.3, this step includes the work of identifying product attributes from Website I and eliciting attributes from Websites II and III, which are explained in the following.

Identifying product attribute candidates from Website I: A pool of attributes are identified from Website I following the same procedure described in Section 3.3.1 – identifying frequent words and pruning the words which are not attributes. The attributes are elicited from the pros and the cons separately and merged together as the attributes identified from Website I. The attributes identified from Website I are not the attributes identified from multiple websites but the candidate attributes from multiple websites. Later in Section 4.3.2.3, a way to determine the attributes from multiple websites will be given.

Eliciting attribute sentences from Websites II and III: The pros/cons from Websites I and II and general comments from Website III are divided into review sentences by standard separators including commas, slashes, and semicolons. The review sentences which evaluate the attributes are called attribute sentences. The attribute sentences are identified from the review sentences by using Dictionaries I and II. Dictionary I – attribute dictionary – includes the words used to indicate attributes, for example, “screen”, “quality”. Dictionary II – attribute support dictionary – includes the words used to evaluate attributes, such as “great”, “poor”, “like”. Dictionary I should be able to take into account the variations in the use of language because customers may use different words when describing an identical thing. For example, for the “Internet” attribute, customers may also use the words like “internet browser”, “web browser”, “web” and so on. Dictionary I is constructed by manually processing the review sentences in Website I which are not identified as attribute sentences. Dictionary II is constructed automatically. The software Text to Matrix Generator (TMG) is used to process the review sentences in Website I by removing the stop words and the words in Dictionary I. After the removing step, the words retained in the reviews are mainly the words evaluating the attributes, as candidate words for Dictionary II. For example, for a review sentence “I like the screen”. The words “I” and “the” are removed as stop words and “screen” is removed as the words in Dictionary I. Only the word “like” is retained as a candidate word for Dictionary II. The candidate words are sorted according to the occurrences of their presence in the reviews. The most frequently presented words are selected and placed in Dictionary II. Manual work is necessary to check the words selected for Dictionary II are the words evaluating the attributes.

The review sentences in Website II including words in Dictionary I are elicited as attribute sentences. Eliciting attribute sentences from general comments (Website III) is more complicated than eliciting attribute sentences from pros/cons (Websites I and II). General comments contain a lot of noise data, such as customers' stories about their experience and actions. Their stories contain attribute words (words in Dictionary I) but the stories do not really evaluate the attributes, for instance, "I turned off apps". The real attribute sentences from general comments should contain attribute words (words in Dictionary I) as well as attribute support words (words in Dictionary II), for example, "The apps are awesome". Thus, the review sentences in Website III can be identified as attribute sentences if they include the words in Dictionaries I and II

4.3.2.2 Determining the Orientations of Product Attributes

The orientations of the product attributes can be defined in two levels – positive (pros) or negative (cons). It is relatively easy to determine the orientations of the attributes from the pros or cons data acquired in Websites I and II. However, it is very difficult to determine the orientations from the general comments in Website III because of a great deal of variations in customer language.

In this task, a supervised classification method – support vector machine (SVM) – will classify the attribute sentences elicited from general comments into two classes – pros and cons. The SVM has two steps – producing a classifier and predicting the classes for new data. In the first step, the pros/cons from Website I are processed as the training datasets to produce a classifier. The benefit of using pros/cons is that the classes (pros and cons) have already been assigned by customers. After the pre-processing, the attribute sentences from the pros/cons in Website I are transferred into numerical vectors in the domain of Dictionary II. The dimension of the vectors is equal to the number of the words in Dictionary II. The number of the words in

Dictionary II is pre-set, e.g., 300 [69] due to the consideration of computation efficiency. The component values in a vector are 0 or 1, indicating whether a corresponding word in Dictionary II appears in the sentence (1 = Yes, 0 = No). For example, assuming that Dictionary II includes five words – "great", "nice", "stupid", "small" and "better" and there are two attribute sentences – one from pros ("Great value for service") and one from cons ("Stupid walkie talkie button"). The numeric vector for the first sentence is [1, 0, 0, 0, 0], indicating the first word in Dictionary II – "great" – appears in the sentence but the others do not. The vector for the second sentence is [0, 0, 1, 0, 0], indicating only the third word "stupid" appears in the sentence. Two training datasets are $\mathbf{v}_1 = [1, 0, 0, 0, 0]$, $u_1 = 1$ (pros) and $\mathbf{v}_2 = [0, 0, 1, 0, 0]$, $u_2 = -1$ (cons). By solving the optimization problem in Equation (2.16), the parameters $\mathbf{W}$ and b can be found. The next step is to classify attribute sentences (testing datasets) from the general comments in Website III into the pros or cons class. The attribute sentences are also transformed into the vectors in the domain of Dictionary II. As described in Section 2.5, to classify a vector, the value $(\mathbf{W} \cdot \mathbf{v} - b)$ is calculated. If $(\mathbf{W} \cdot \mathbf{v} - b) \geq 0$, the vector will be classified into pros; otherwise, cons. For example, there is one attribute sentence (testing dataset) elicited from general comments – "Great call quality on both ends". The vector for this sentence is $\mathbf{v}_c = [1, 0, 0, 0, 0]$. If the value $(\mathbf{W} \cdot \mathbf{v}_c - b) > 0$, then this sentence is classified into the pros class.

After all the attribute sentences from multiple websites are classified into pros/cons, the pros/cons are modeled as customer attribute ratings following the procedure in Section 3.3.1.

4.3.2.3 Determining the Product Attributes for Product Design

The final step is to determine the product attributes for product design from the attribute candidates identified from Website I. In general, the attributes which are the most frequently referred can be selected and determined as the attributes for the websites of interest. The frequency of attributes is the occurrences of the attributes being specified as pros or cons across the websites.

4.3.3 Integrating Customer Preferences from Multiple Websites

As described in Section 3.3.2, customer data are in a hierarchical structure as in Figure 3.5. In order to construct the hierarchical models under website heterogeneity, a random-effects meta-analysis technique is applied. The hierarchical customer preference models are formulated as follows.

Level I – a product rating R is a function of attribute aggregate ratings $\mathbf{X}$

$$R = \boldsymbol{\beta}_{\mathrm{I}} \cdot \mathbf{X} + \boldsymbol{\alpha} + \varepsilon_{\mathrm{I}} \tag{4.1}$$

Level II – aggregate ratings $\mathbf{X}$ is a function of attribute ratings $\mathbf{Y}$

$$\mathbf{X} = \boldsymbol{\beta}_{\mathrm{II}} \cdot \mathbf{Y} + \boldsymbol{\gamma} + \boldsymbol{\varepsilon}_{\mathrm{II}} \tag{4.2}$$

Level III – attribute ratings $\mathbf{Y}$ is a function of product specifications $\mathbf{E}$

$$\mathbf{Y} = \boldsymbol{\beta}_{\mathrm{III}} \cdot \mathbf{E} + \boldsymbol{\theta} + \boldsymbol{\varepsilon}_{\mathrm{III}} \tag{4.3}$$

where R, $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{E}$ are the product rating, attribute aggregate ratings, attribute ratings and product specifications; $\boldsymbol{\beta}_{\mathrm{I}}$, $\boldsymbol{\beta}_{\mathrm{II}}$, and $\boldsymbol{\beta}_{\mathrm{III}}$ are the model parameters; ε_{I}, $\boldsymbol{\varepsilon}_{\mathrm{II}}$, and $\boldsymbol{\varepsilon}_{\mathrm{III}}$ are the error terms; $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\theta}$ are random effects due to website heterogeneity, representing the variation over websites. The random effects in the three levels are expressed as $\boldsymbol{\alpha}=[\alpha_1, .., \alpha_l, ..., \alpha_Q]$, $\boldsymbol{\gamma} =[\gamma_1, .., \gamma_l, ..., \gamma_Q]$, and $\boldsymbol{\theta} =[\theta_1, .., \theta_l, ..., \theta_Q]$, where α_l, γ_l, and θ_l represent the random effect in the l^{th} website (l=1, 2, ..., Q), usually following the normal distributions [108] $\alpha_l \sim \mathrm{Normal}(0, \tau_{\mathrm{I}}^2)$, $\gamma_l \sim$ Normal(0,

τ_{II}^2) and $\theta_l \sim$ Normal(0, τ_{III}^2). Q denotes the number of websites in total. τ_{I}^2, τ_{II}^2, and τ_{III}^2 are the between-website variance component in Levels I, II and III respectively.

It should be noticed that attribute aggregates determined by each website are distinguishable. Assume there is a set of attribute aggregates considered for meta-analysis, some websites do not carry entire attribute aggregates or part of the aggregates. The attribute aggregate which is not collected in a website can be regarded as missing data of that website. For example, assume there are five attribute aggregates considered for the meta-analysis, $\mathbf{X} = [X_1, X_2, X_3, X_4, X_5]$, where X_1= "value for price", X_2= "features", X_3= "ease of use", X_4 = "design", and X_5 = "battery life". Website A does not collect attribute aggregate ratings, thus, $\mathbf{X}_{\mathrm{A}}$ is regarded as missing data of Website A. Website B carries part of the attribute aggregates, say, it carries X_1 ("value for price"), X_2 ("features"), X_3 ("ease of use") and X_5 ("battery life"). Then X_4 ("design") can be regarded as missing data of Website B. The missing data can be imputed by WinBUGS [94]. The basic idea of the imputation is that the missing data (missing aggregates) can be estimated by borrowing information from known data (known aggregates from other websites) as well as by regression as in Equation (4.2).

After constructing customer preference models, the product design selection process is performed which is the same as the selection process in Section 3.3.3.

4.4 Case Study

Same as Section 3.4, a smartphone design selection problem is considered as an example.

4.4.1 Selecting Online Customer Reviews of Interest

This case study involves eight public websites listed in Table 4.1. All websites have product ratings and general comments. Among them, five websites (www.bestbuy.com, www.epinions.com, www.att.com, www.samsung.com and www.tmobile.com) deal with attribute aggregate ratings while six websites collect the pros/cons (three websites with the pros/cons checklists). Therefore, depending on the pros/cons data existence and format, the eight websites can be divided into three groups of Website I, II and III. Website I includes the websites with freely written pros/cons – part of the websites from Groups I and III (see Section 2.4.2). Website II includes the websites with guided pros/cons – part of the websites from Groups I and III (see Section 2.4.2). Website III includes the websites without pros/cons – the websites from Groups II and IV (see Section 2.4.2). Table 4.2 displays the attribute aggregates that the eight websites deal with. The first column lists all the attribute aggregates appearing in the eight websites and the second column presents the number of the appearances of the attribute aggregates. It can be seen that some attribute aggregates appear in multiple websites, such as battery life, whereas other aggregates is recognized in one website only, e.g. clarity.

Table 4.1 Website groups

	Websites	Product ratings	Attribute aggregate ratings	General comments	Pros/cons	Pros/Cons guided by websites
Web. I	bestbuy.com	Y	Y	Y	Y	N
	cnet.com	Y	N	Y	Y	N
	epinions.com	Y	Y	Y	Y	N
Web. II	att.com	Y	Y	Y	Y	Y
	samsung.com	Y	Y	Y	Y	Y
	tmobile.com	Y	Y	Y	Y	Y
Web. III	amazon.com	Y	N	Y	N	–
	phonescoop.com	Y	N	Y	N	–

Table 4.2 Attribute aggregates that the eight websites deal with

Attribute aggregates	Replications	8 websites				
		bestbuy	epinions	att	samsung	tmobile
Features	4	√		√	√	√
Battery life	4	√	√	√		√
Ease of use	3	√		√		√
Value for price	2	√			√	
Design	2			√	√	
Call quality	1					√
Performance	1				√	
Display	1			√		
Durability	1		√			
Clarity	1		√			
Portability	1		√			

√ indicates that attribute aggregates appear in that website

Similar to Section 3.4, multiple observations data are collected. In total, 932 customer reviews were collected and downloaded from 380 customers from the eight websites. The reviews were written from September 2009 to March 2011.

Table 4.3 The number of reviews from the eight websites

Websites	No. of customers	No. of reviews
bestbuy.com	61	128
cnet.com	51	123
epinions.com	42	132
att.com	57	121
samsung.com	52	106
tmobile.com	44	96
amazon.com	39	97
phonescoop.com	34	129
Total	380	932

4.4.2 Product Attributes Identification

The first step of the proposed methodology is to identify important product attributes and elicit customer data from the eight websites. Firstly, frequent words are identified individually from the pros and cons from Website I. Twenty-four frequent words and eighteen frequent words were identified from pros and cons respectively. Using the pruning rules, “data”, “life” and “lot” were removed. Then the retained words from pros and cons were merged together into the twenty-five attribute candidates identified from Website I (www.bestbuy.com, www.cnet.com and www.epionions.com), which are listed in Figure 4.4.

Applications	Battery	Network	Internet	Keyboard
Call quality	Camera	Design	Email	Feature
Feeling	Appearing	Memory	Price	Processor
Video	Quality	Screen	Size	Texting
Wi-Fi	HDMI	Spearker	OS	Ease of use

Figure 4.4 Attribute candidates identified from Website I

Two dictionaries are constructed in the way explained in Section 4.3.2.1. Dictionaries I and II are used to elicit attribute sentences and corresponding attribute ratings in Websites II and III. The attribute sentences in Website II are elicited using Dictionary I whereas those in Website III are elicited using Dictionaries I and II. Attribute ratings (pros and cons indicators) can be easily elicited from the pros and cons list in Websites I and II, whereas the attribute sentences in Website III are classified into pros or cons by the SVM. For validation of the classification result, a manual classification of a hundred attribute sentences into pros and cons can be used as the reference. The classification result from the SVM is compared with the reference and the classification precision can be calculated as

$$\text{Precision} = \frac{\text{number of sentences being correctly classified}}{\text{total number of sentences}} \tag{4.4}$$

Table 4.4 lists the classification precision of attribute sentences being classified into pros or cons for two websites – amazon.com and phonescoop.com. As listed in the table, the classification precision is similar for the two websites (0.84 and 0.85), which is higher than 0.76 for cell phone, and comparable to the average precision 0.84 across several product categories [42]. Plus the proposed method overcomes the limitation – predicting the orientations (positive or negative) by using adjective words only.

Table 4.4 Classification precision

	Amazon.com	Phonescoop.com	Cell phone	Average
Precision	0.84	0.85	0.76	0.84

Finally, the attributes from the eight websites are selected from the twenty-five attribute candidates identified from Website I. The attributes with the top ten pros/cons frequency are selected from the eight websites, as listed in Table 4.5. Eight attributes appear in both pros and cons sides. The pros side has two non-common attributes "Text" and "Ease of use" and so do the cons – "Size" and "Internet". Table 4.5 also lists the frequency values of the pros or cons indicator. The total number of reviews is 932. The indicator with the highest frequency is "applications_pros", as high as 218.

Table 4.5 Attributes selected from the eight websites

Attribute_pros	Frequency	Attribute_cons	Frequency
Applications	218	Applications	108
Battery	180	Battery	237
Keyboard	181	Keyboard	136
Camera	274	Camera	92
Processor	172	Processor	85
Quality	206	Quality	86
Screen	378	Screen	83
OS	191	OS	64
Text	178	Size	59
Ease of use	195	Internet	54

4.4.3 Model Estimation, Results, Interpretations and Comparisons

In this section, the model employed for this case study is firstly introduced. Then, the estimated results are represented. Finally, to understand the website heterogeneity, the meta-analysis results are interpreted and compared with results from individual websites.

4.4.3.1 Model Estimation

The data from the eight websites are used for estimation of model parameters using a meta-analysis approach. To conduct meta-analysis of multiple websites, random-effects linear regression is employed for Levels I and II [108]. Random-effects binary logistic regression is employed for Level III [108]. The details of each model are introduced in the following.

- Level I – the random-effects linear regression

The product rating (R_l) in the l^{th} website can be modeled as

$$R_l = \boldsymbol{\beta}_{\text{I}} \cdot \mathbf{X}_l + \alpha_l + \varepsilon_{\text{I}} \tag{4.5}$$

where $\mathbf{X}_l$ is the attribute aggregate ratings in the l^{th} website; α_l represents the random effects for the website $l = 1, \ldots, 8$, representing the variation over websites, $\alpha_l \sim \text{Normal}(0, \tau_{\text{I}}^2)$, where τ_{I}^2 is the between-website variance component for Level I; ε_{I} is a statistical error term, $\varepsilon_{\text{I}} \sim \text{Normal}(0, \sigma_{\text{I}}^2)$; $\boldsymbol{\beta}_{\text{I}}$ are model parameters in Level I.

- Level II – the random-effects linear regression

The attribute aggregate ratings ($\mathbf{X}_l$) in the l^{th} website can be modeled as

$$\mathbf{X}_l = \boldsymbol{\beta}_{\text{II}} \cdot \mathbf{Y}_l + \boldsymbol{\gamma}_l + \boldsymbol{\varepsilon}_{\text{II}} \tag{4.6}$$

where $\mathbf{Y}_l$ are the attribute ratings in the l^{th} website; $\boldsymbol{\gamma}_l$ represents random effects for the attribute aggregate ratings $\mathbf{X}_l$ in the website $l = 1, \ldots, 8$; random effects $\boldsymbol{\gamma}_l$ are website-specific and identical for any aggregate ratings. Thus, the components of $\boldsymbol{\gamma}_l$

are equal to γ_l ~ *Normal* (*0,* τ_{II}^2) and τ_{II}^2 is the between-website variance component for Level II; a statistical error term $\boldsymbol{\varepsilon}_{\text{II}}$ ~ multivariate Normal(0, $\boldsymbol{\sigma}_{\text{II}}^2$); $\boldsymbol{\beta}_{\text{II}}$ are model parameters in Level II. Attribute aggregate ratings **X** are distinct across different websites. The missing **X**s are imputed in Level II. They are imputed based on the regression in Level II and also the correlations between different aggregates. The attribute aggregates **X** are assumed to be multivariate-normally distributed.

- Level III – the random-effects binary logistic regression

The j^{th} observation set of attribute ratings ($\mathbf{Y}_{j,l}$) and its probability ($\boldsymbol{\phi}_{j,l}$) in the l^{th} website can be modeled as

$$\mathbf{Y}_{j,l} = \boldsymbol{\beta}_{\text{III}} \cdot \mathbf{E}_{j,l} + \boldsymbol{\theta}_l + \boldsymbol{\varepsilon}_{\text{III}} \tag{4.7}$$

$$\text{logit}(\boldsymbol{\phi}_{j,l}) = \boldsymbol{\beta}_{\text{III}} \cdot \mathbf{E}_{j,l} + \boldsymbol{\theta}_l \tag{4.8}$$

where $\boldsymbol{\phi}_{j,l}$ denotes the probability of $\mathbf{Y}_{j,l}$ equal to 1; $\mathbf{E}_{j,l}$ are the j^{th} observation set of product specifications in the l^{th} website; $\boldsymbol{\theta}_l$ represents random effects for the attribute ratings $\mathbf{Y}_{j,l}$ in the website l = 1, …, 8; random effects $\boldsymbol{\theta}_l$ are website-specific and identical for any attribute ratings. Thus, the components of $\boldsymbol{\theta}_l$ are equal to θ_l ~ *Normal* (*0,* τ_{III}^2) and τ_{III}^2 is the between-website variance component; $\boldsymbol{\varepsilon}_{\text{III}}$ is a statistical error term – independent and identically distributed extreme values; $\boldsymbol{\beta}_{\text{III}}$ are model parameters in Level III.

For computational efficiency, conjugate Bayes models are used to update the statistics of the following parameters: σ_{I}, τ_{I}, τ_{II} and τ_{III} ~ inverse Gamma distribution; $\boldsymbol{\sigma}_{\text{II}}$ ~ inverse Wishart distribution and; $\boldsymbol{\beta}$s ~ Normal distribution.

Table 4.6 lists the detailed descriptions for attribute aggregates **X**, attributes **Y** and product specifications **E**. Note that for this case study five attribute aggregates were chosen because they are provided at least by two websites. The missing **X**s are imputed by borrowing the information from the known **X**s. More known data can assure more accurate imputation. That is the reason for selecting aggregates appearing in at least two websites. The product specifications are the same as that in Section 3.4.

Table 4.6 Description of attributes and specifications

Attribute aggregates **X**:	
X_1 – Value for price;	X_2 – Features;
X_3 – Ease of use;	X_4 – Battery life;
X_5 – Design;	
Attributes **Y**:	
Attribute_Pros	Attribute_Cons
Y_1 – Applications	Y_{11} – Applications
Y_2 – Battery	Y_{12} – Battery
Y_3 – Keyboard	Y_{13} – Keyboard
Y_4 – Camera	Y_{14} – Camera
Y_5 – Processor	Y_{15} – Processor
Y_6 – Quality	Y_{16} – Quality
Y_7 – Screen	Y_{17} – Screen
Y_8 – OS	Y_{18} – OS
Y_9 – Text	Y_{19} – Size
Y_{10} – Ease of use	Y_{20} – Internet
Product specifications **E**:	
E_1 – Network variable (1=4g, 0=Not 4g);	
E_2 ~ E_5 – dummy variables for OS ([1,0,0,0]=OS1, [0,1,0,0]=OS2, [0,0,1,0]=OS3, [0,0,0,1]=OS4, [0,0,0,0]=OS5);	
E_6 – height (inch);	E_7 – width (inch);
E_8 – depth (inch);	E_9 – weight (ounce);
E_{10} – display size (the diagonal length of a display screen);	
E_{11} – total pixel resolution (defined $w_r \times h_r$, w_r and h_r are width and height resolution in pixel respectively);	
E_{12} – touch screen (1=Yes, 0=No);	
E_{13} – battery capacity (mAh)	
E_{14} – camera resolution (mega-pixel);	
E_{15} ~ E_{16} – video variables ([1,0] = high-definition video, [0, 1] = regular definition video, and [0, 0] = no video);	
E_{17} – processor variable (1= = "processor speed $\geq$ 800 MHz", 0= "processor speed < 800 MHz");	
E_{18} – memory variable (1= "memory $\geq$ 1 GB", 0="memory < 1 GB").	
E_{19} – phone form (1="slide form", 0="bar form").	
E_{20} – physical keyboard (1=Yes, 0=No).	
E_{21} – Wi-Fi variable (1=Yes, 0=No).	

4.4.3.2 Model Results

The parameters were estimated using WinBUGS. Table 4.7 – Table 4.9 list the statistical results (mean values μ_β, standard deviations σ_β) for the parameters of the selected models at three levels. The significant variables as defined in Section 3.4.3 are marked with an asterisk (*) in the tables. For Level I, Table 4.7 lists the parameters for all the five attribute aggregates plus the constant. The mean values for the βs of the four aggregates (Value for price, Features, Ease of use and Battery life) are significantly positive. The positiveness indicates that the four aggregate ratings have significantly positive effects on the overall product rating R. The larger value β is; the more effects the attribute aggregate has. Therefore, the attribute aggregate "Value for price" has the largest effects on the product rating. This conclusion coincides with the conclusion in Section 3.4.3 – "value for price" has the largest effects across different segments.

Table 4.7 Model estimated results for Level I

Level I – parameter estimates for product rating R;

Attribute aggregates	μ_β	σ_β
Value for price*	0.47	0.04
Features*	0.32	0.04
Ease of use*	0.26	0.06
Battery life*	0.11	0.04
Design	-0.07	0.06
Constant*	-0.52	0.20

For Level II, Table 4.8 lists the parameters of the attributes for the attribute aggregate "Features". Five attribute_pros with an asterisk (*) are significantly positive, indicating that the attributes being specified as pros positively affect ratings for the attribute aggregate "Features". The parameters for the five attribute_cons with an asterisk (*) are signficantly negative, indicating that the attributes being specified as cons negatively affect ratings for the attribute aggregate "Features". By comparing the absolute values of βs for the ten attributes, it can be seen that the attribute cons tend to have a larger absolute value of βs, that is, the attribute cons tend to have more effects on the attribute aggregate ratings than the attribute pros.

Table 4.8 Model estimated results for Level II

Level II – parameter estimates for the aggregate "Features" X_2 (only significant attributes listed)

Attributes	$\mu_\beta(\sigma_\beta)$	Attributes	$\mu_\beta(\sigma_\beta)$
Applications_pros*	0.55(0.10)	Applications_cons*	-0.32(0.13)
Battery_pros	-0.13(0.11)	Battery_cons	-0.02(0.09)
Keyboard_pros	0.03(0.11)	Keyboard_cons*	-0.53(0.18)
Camera_pros	-0.04(0.10)	Camera_cons	-0.17(0.12)
Processor_pros*	0.26(0.12)	Processor_cons	-0.07(0.14)
Quality_pros	0.08(0.09)	Quality_cons	-0.27(0.18)
Screen_pros	0.15(0.09)	Screen_cons*	-0.59(0.13)
OS_pros*	0.32(0.12)	OS_cons*	-0.62(0.14)
Text_pros*	0.25(0.11)	Size_cons	-0.24(0.15)
Ease of use_pros*	0.26(0.10)	Internet_cons*	-0.61(0.16)
		Constant*	4.17(0.15)

For Level III, Table 4.9 lists the parameters of the product specifications for the attribute Y_1 – "Application_pros". Level III employs the binary logistic regression models. For the OS specification, OS 5 ([0,0,0,0]) is a base value. The parameter for OS 1 and OS 2 is significantly positive, which means when OS changes from OS 5 (the base value) to OS 1 or 2, the probability of the attribute "Application" being specified as pros become larger. In other words, compared to customers using OS 5, customers using OS 1 and 2 are more likely to specify the attribute "application" as pros. This can be observed in Section 3.4.2 as well – customers seem to like the applications of OS 1 and 2. For all other product specifications, in short, the smartphones, which are shorter, narrower, in a bar form, and with physical keyboard and Wi-Fi equipment, tend to have a higher probability with the attribute “application” being specified as pros.

Table 4.9 Model estimated results for Level III

Level III – parameter estimates for the attribute “Application_pros” Y_1 (only significant *E*s listed)

Specifications	$\mu_\beta(\sigma_\beta)$	Specifications	$\mu_\beta(\sigma_\beta)$
Network	-0.24(0.30)	Touch screen	0.18(0.53)
OS 1*	0.96(0.44)	Battery capacity	-0.39(0.68)
OS 2*	0.96(0.41)	Cam resolution	-0.17(0.09)
OS 3*	-1.34(0.75)	Video (HD)	1.35(0.95)
OS 4	0.41(0.29)	Video (regular)	0.65(0.85)
Height*	-1.67(0.54)	Processor	0.03(0.34)
Width *	-3.93(1.09)	Memory	-0.03(0.27)
Depth	2.07(1.75)	Phone form*	-1.73(0.50)
Weight	0.19(0.19)	Keyboard*	1.11(0.53)
Display size*	2.15(0.65)	Wi-Fi*	1.13(0.37)
Total resolution	-0.19(0.12)	Constant*	5.73(2.52)

The variance σ_{I}^2 for the error terms ε_{I} and the covariance matrix $\boldsymbol{\sigma}_{\mathrm{II}}^2$ for the error terms $\boldsymbol{\varepsilon}_{\mathrm{II}}$ are listed in Equations (4.9) to (4.10). The variances of error terms $\boldsymbol{\varepsilon}_{\mathrm{II}}$ (diagonal elements of $\boldsymbol{\sigma}_{\mathrm{II}}^2$) are larger than the variance of the error term ε_{I} (σ_{I}^2), indicating a larger error in Level II and that in Level I.

$$\sigma_{\mathrm{I}}^2 = 0.26 \tag{4.9}$$

$$\boldsymbol{\sigma}_{\mathrm{II}}^2 = \begin{bmatrix} 1.21 & 0.50 & 0.56 & 0.36 & 0.39 \\ 0.50 & 0.73 & 0.40 & 0.21 & 0.37 \\ 0.56 & 0.40 & 0.84 & 0.34 & 0.45 \\ 0.36 & 0.21 & 0.34 & 0.71 & 0.02 \\ 0.39 & 0.37 & 0.45 & 0.02 & 0.69 \end{bmatrix} \tag{4.10}$$

The pseudo-r^2 values, *RMSE*, and *MAE* are calculated to quantify the predictive capability of the models, as listed in Table 4.10.

Table 4.10Model evaluation

	Level I	Level II	Level III
Pseudo-r^2	0.74	0.42	0.27
RMSE	0.35	0.88	0.29
MAE	0.25	0.65	0.31

The pseudo-r^2 value for the model of Level I is fairly high while the pseudo-r^2 values for Level II are relatively lower. The pseudo-r^2 value for Levels II is considered to be acceptable considering the nature of subjective data and the mappings from textual reviews into numerical values. The errors in Level I are smaller however the errors in Level II seems to be high, especially *RMSE*. The larger errors in Level II indicate model fitting in Level II is worse than that of Level I. Larger difference between *RMSE* than *MAE* in Level II indicates that there is a large variance existing in the individual errors in the data because given the definitions in Section 2.7, *MAE* is a linear evaluation of errors but *RMSE* is a quadratic evaluation of errors, which is exaggerated by large individual errors. For Level III – binary logistic regression, as stated in Section 3.4, the pseudo-r^2 value is normally low and can be accepted if the value is between 0.2 – 0.4 [95,96]. In short, the pseudo-r^2 values for the models indicate that the models developed from customer reviews can explain customer preferences for smartphones reasonably well.

Besides pseudo-r^2, the *DIC* value is also calculated for Level III. The *DIC* value for the Level III model is compared with that for a null model. The null model is the Level III model with the intercept only. If the *DIC* value for the full model is significantly smaller than that for the null model, it can be concluded that at least one predictor in the full model is significant. The *DIC* values for the full and null models are listed in Table 4.11. It can be noted that the *DIC* values for Level III is noticeably smaller, which indicates at least one predictor in Level III is significant and the goodness-of-fit of the Level III is acceptable.

Table 4.11 DIC values for Level III

	Level III	The null model of Level III
DIC	1.74E+05	1.10E+06

4.4.3.3 Model Interpretations and Comparisons

The meta-analysis results are compared with the estimated results for individual websites. The comparison result for Level I is demonstrated in Table 4.12. The box-plots of parameters for two attribute aggregates are shown in Figure 4.5. The comparisons for Levels II and III are omitted due to similarity. The model results for individual websites are estimated using the models defined in Equations (4.5) to (4.8) except that the random-effects are removed.

By comparing the mean values of the parameters, it can be seen that the individual results are close to each other and the meta-analysis results are nearly the average of the individual results but slightly smaller than the average. The reason of being smaller than the average is that the number of independent variables (attribute aggregates) for meta-analysis (five aggregates for meta-analysis) is greater than the number from individual websites (equal to or less than four aggregates from individual websites). More variables will induce fewer weights on each variable. That explains the parameters for meta-analysis tend to be smaller than the average of individual results. By comparing the standard deviations of the parameters, it is found that the standard deviations of meta-analysis results are smaller than the standard deviations of individual results. The smaller deviation is credited to more samples involved in the meta-analysis.

Table 4.12 Comparison of model estimated results for Level I

Attribute aggregates	Meta-analysis		Bestbuy		Epinions		Att		Samsung		Tmobile	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Value for price	0.47	0.04	0.43	0.05	–	–	–	–	0.58	0.07	–	–
Features	0.32	0.04	0.38	0.06	–	–	0.37	0.09	0.40	0.09	0.46	0.09
Ease of use	0.26	0.06	0.18	0.06	–	–	0.43	0.07	–	–	0.36	0.09
Battery life	0.11	0.04	0.18	0.04	0.22	0.08	0.21	0.05	–	–	0.27	0.07
Design	-0.07	0.06	–	–	–	–	0.05	0.09	0.12	0.08	–	–
Constant	-0.52	0.20	-0.61	0.22	2.90	0.36	-0.17	0.26	-0.54	0.31	-0.57	0.27

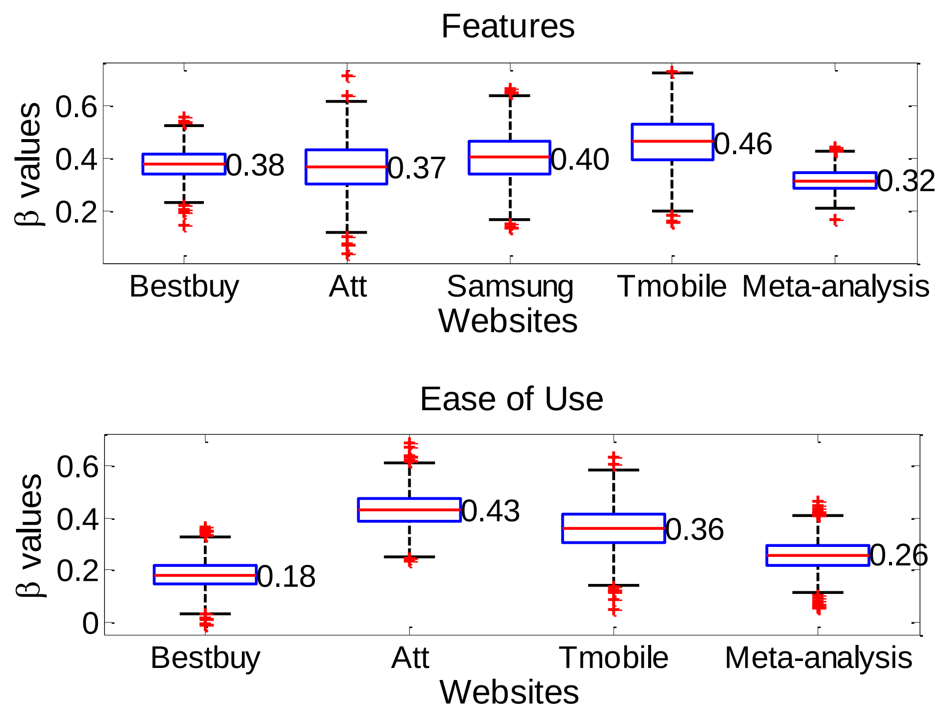


Figure 4.5 Box-plot comparison of the meta-analysis results

The random effects for the three levels are compared to analyze the possible website heterogeneity across the eight websites. Website heterogeneity observed can be caused by a lot of factors such as the differences of website formats, the effects of different text mining process, etc. The box-plots of the random effects are shown in Figure 4.6. To better understand the random effects in the websites, they are sorted into the three groups as before. Website I includes www.bestbuy.com, www.cnet.com and www.epinions.com. Website II includes www.att.com, www.samsung.com and www.tmobile.com. Website III includes www.amazon.com and www.phonescoop.com. The format differences for each group are summarized in Table 4.13. This paragraph mainly discusses the random effects caused by the format differences. For Website I, the pros/cons are obtained in the same procedure but the attribute aggregates are different. The website 'www.bestbuy.com' includes four attribute aggregates, 'www.cent.com' does not have any aggregates, and 'www.epinions.com' only includes one aggregate. It is known that in the hierarchical models, the pros/cons affect the results of Level III. Thus, the random effects of the three websites for Level III should be similar, which can be observed in the bottom figure of Figure 4.6. Meanwhile, due to the differences of attribute aggregates, the random effects of the three websites for Levels I and II should be different, which can be observed in Figure 4.6 as well. For Website II, the pros/cons are all guided. However, the guided pros/cons are only similar for 'www.att.com' and 'www.tmobile.com' but different from 'www.samsung.com'. As shown in Figure 4.7, the pros/cons checklists for 'www.att.com' and 'www.tmobile.com' are very specific and lengthy but the checklist for 'www.samsung.com' is abstract. All the three websites have attribute aggregates and similar. As shown in Figure 4.6, it can be found that the random effects for 'www.att.com' and 'www.tmobile.com' are close across all the three levels, but different from 'www.samsung.com'. This observation can be explained by their similarity and dissimilarity in pros/cons and attribute aggregates. For Website III, 'www.amazon.com' and 'www.phonescoop.com' are similar to each other. They do not have attribute aggregates and pros/cons. Thus, their random effects for all three levels are similar to each other as observed in Figure 4.6. It is concluded that the format differences can explain the random effects reasonably. However, as stated early in this paragraph, website heterogeneity is caused by various factors. The format difference is one of the main factors. Other factors, such as different text mining processes, may have impacts as well. Recall in Section 4.3.2, attribute ratings (**Y**) are elicited through different text mining processes for the three groups – Websites I, II and III. The different processes should also be a reason of the differences of random effects in Levels II and III for the three groups of websites.

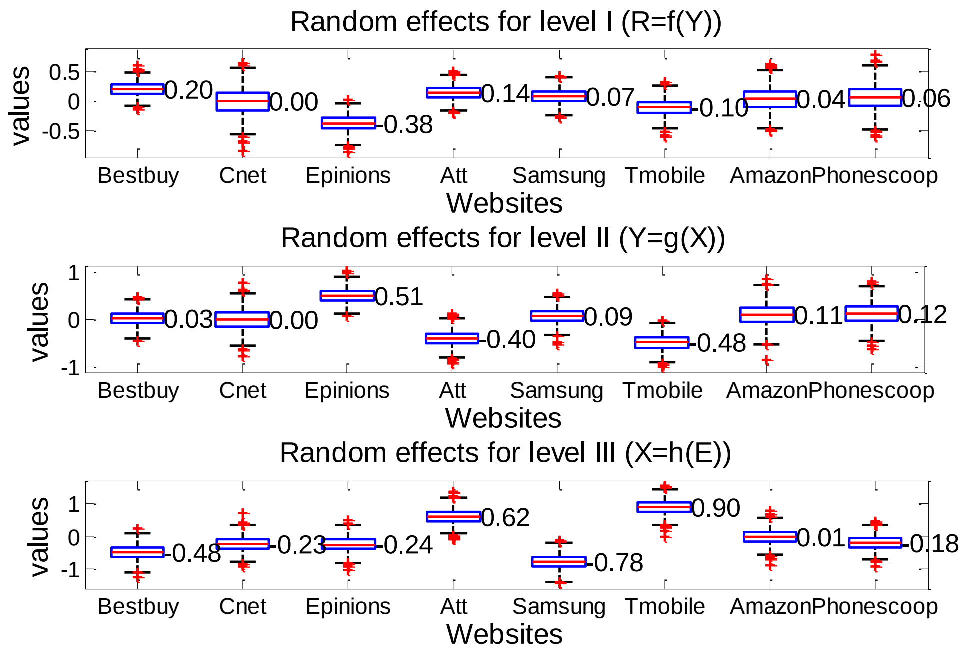


Figure 4.6 Box-plots of Random effects

Table 4.13 Summary of random effects in three website groups

		Attribute aggregate ratings	Pros/cons	Random effects
Website I	Bestbuy	Features Battery life Ease of use Value for price	Freely written	Random effects for Level III are similar.
	Cnet	NA	Freely written	
	Epinions	Battery life	Freely written	
Website II	Att	Features Battery life Ease of use Design	Guided	Random effects for Att and Tmobile are similar.
	Samsung	Features Value for price Design	Guided*	
	Tmobile	Features Battery life Ease of use	Guided	
Website III	Amazon	NA	NA	Random effects are similar.
	Phonescoop	NA	NA	

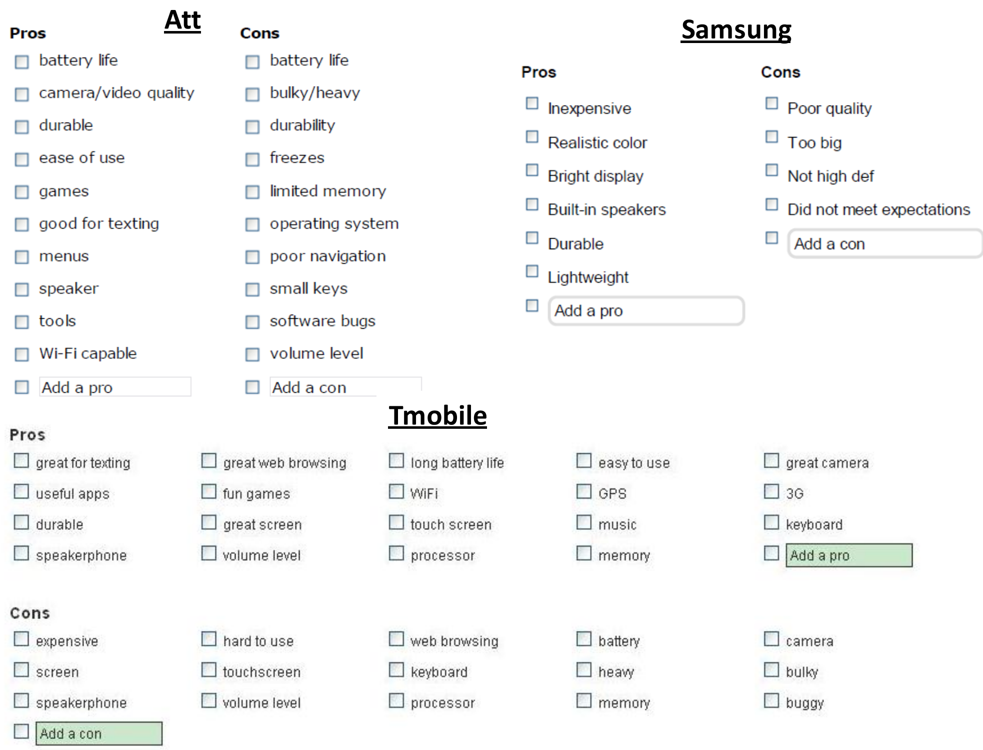


Figure 4.7 The checklists of pros/cons for Website II

4.4.4 Product Design Selection

The estimated preference model is applied for the same design problem in Section 3.4. The design selection results are listed in Table 4.14. Meanwhile, the best design alternatives are compared with the alternative with the best product ratings.

Given different prices, the design results are consistent in most design variables. As the price increases, the phone OS is suggested to be changed from OS 5 to OS 1 and the size is changed from small to large as well. It indicates that the OS and size variables should be the most important variables to customers. Among all the four best design alternatives, it can be seen that a smartphone with smaller battery capacity and a smaller storage memory seems related to higher phone ratings. This seems conflicting with the real world. However, it is probable that there may be other specifications which are missed in this case study as being more important. For example, customers should prefer the phones with a longer battery life. But the battery life is not affected by battery capacity only. It depends on OS, phone usage and other variables as well. For the memory, regardless of the memory storage, a lot of smartphones are equipped with memory slots, which allows customers to expand the memory by themselves. In that sense, memory storage may not be the only factor affecting the memory. However, other than the two variables, the design selection results are reasonable. The design alternative with the best rating tells that customers tend to rate higher for the phones with advanced equipments, e.g., 4g network, larger size, higher resolution screen and so on. Recall the design results in Table 3.9 in Chapter 3 from online customer reviews from a single website, the optimal design alternatives are similar to the ones in Table 4.14 from multiple websites, such as bar form, processor, memory and so on.

Table 4.14 Design selection results

	Price	NA	$99.99	$199.99	$299.99
Design variables	Network (E_1)	4g	No 4g	No 4g	No 4g
	OS (E_2 ~E_5)	1	5	1	1
	Size (E_6 ~E_{10})	Large	Small	Small	Large
	Screen resolution (E_{11})	High	Low	Low	Low
	Touch (E_{12})	0	0	0	0
	Battery capacity (E_{13})	1200	1200	1200	1200
	Camera resolution (E_{14})	5	3.2	3.2	3.2
	Video (E_{15} ~ E_{16})	HD	Regular	Regular	Regular
	Processor (E_{17})	>800MHz	>800MHz	>800MHz	>800MHz
	Memory (E_{18})	<1GB	<1GB	<1GB	<1GB
	Phone form (E_{19})	Slide	Bar	Bar	Bar
	Physical keyboard (E_{20})	Yes	Yes	Yes	Yes
	Wi-Fi (E_{21})	Yes	No	No	No

4.5 Summary

This chapter proposed a new methodology for customer-driven product design selection by integrating customer reviews from multiple websites. This methodology employed a text classification method to develop the pros/cons classifier from the websites with pro/cons data and use the classifier to classify general comments from the websites without pros/cons. The meta-analysis technique is then employed for integrating customer data and constructing customer preference models from multiple websites. The use of the meta-analysis technique can integrate customer data under website heterogeneity. Finally, the customer preference models developed using online customer reviews from multiple websites are used for product design selection problem – select a product design alternative that maximizes the profit.

This chapter makes two main contributions: (i) reducing the amount of human work for the content analysis in the text mining process by using the pros/cons classifier to classify general comments; (ii) accounting for website heterogeneity by integrating customer data from multiple websites through the meta-analysis technique.

The smartphone case study was used to demonstrate the feasibility of the proposed methodology. The case study considered eight public websites. The results suggested that meta-analysis captured website heterogeneity well. The design selection results yielded to reasonable solutions.

Two important assumptions made in this chapter, no customer preference heterogeneity and independence of multiple responses on a single customer. These two assumptions are going to be relaxed in next chapter.

Chapter 5: Product Design Selection Using Online Customer Reviews from Multiple Websites with Website and Customer Preference Heterogeneity

5.1 Introduction

In Chapter 4, product design selection using customer reviews from multiple websites were presented under two main assumptions. The two assumptions were: (i) no customer preference heterogeneity was considered and (ii) the correlations among multiple observations by a single customer were ignored. This chapter relaxes these two assumptions. For comparison, the model in Chapter 4 is called Model I.

The heterogeneity in customer preferences can be modeled using various types of models. The model (Model II) employed in this chapter is based on a mixture model. In this model, parameters for customer preferences ($\boldsymbol{\beta}$s) are assumed to be randomly distributed across customers.

The second assumption was the independence of multiple observations from a single customer. Multiple observations from a single customer indicate multiple reviews collected from a single customer. One review includes two types of ratings (product rating R, attribute aggregate ratings $\mathbf{X}$) and one type of decision data (attribute ratings $\mathbf{Y}$). Note that attribute ratings $\mathbf{Y}$ are essentially decision variables (specifying or not specifying an attribute as a pro or con). As reviewed in Section 2.3, multiple observations from a single customer should be correlated to some extent because the errors of the observations are caused by the customer's biases due to their inherent habits, cultural backgrounds and so on. In addition, two models were introduced in Section 2.3 to model the correlations among multiple observations – one for rating data and the other for decision data. However, no literature exists for the error correlation between rating data and decision data. Thus, in this chapter, three models are proposed to investigate the correlation of multiple observations in the hierarchical preference models – (Model III) assuming that no correlations across levels (correlations within each level); (Model IV) correlations across the top two levels (correlations between R and $\mathbf{X}$); and (Model V) correlations across all three levels (correlations between R, $\mathbf{X}$ and $\mathbf{Y}$).

Different hierarchical models for customer preference are employed to investigate relaxing the two assumptions considered in the last chapter. The first model takes into account both website heterogeneity and customer preference heterogeneity. The second to fourth models account for the two types of heterogeneity plus the correlations of multiple observations. To investigate the validity of the four hierarchical models, estimation results from the four models are compared with those from the model in Chapter 4 using the same case study. Additionally, a set of out-of-sample data was used to compare and validate the models where the out-of-sample data were collected from three websites, not used for the model estimation.

This chapter is structured as follows. Section 5.2 introduces the three hierarchical models for the customer preference. In Section 5.3, the four models are applied to the same case study considered in Chapter 4. The estimation results and performance of the models are compared. Section 5.4 attempts to validate the models by employing the test data and comparing the results with that from the model in Chapter 4.

5.2 Approach

This section proposes four different models (Models II to V) for hierarchical customer preference by relaxing two assumptions sequentially.

5.2.1 Considering both Customer Preference Heterogeneity and Website Heterogeneity (Model II)

Model II takes into account customer preference heterogeneity along with website heterogeneity. Since Chapter 4 considered website heterogeneity only, it was assumed that the parameters (**β**s) are identical over different individual customers. In this chapter, the customer preference heterogeneity is considered and expressed using the parameters (**β**s), which follow probability distributions over heterogeneous customers. The proposed model is detailed as follows.

For n^{th} customer in l^{th} website (e.g., $l = 1, \ldots, 8$, as in the smartphone example), the hierarchical model for customer preference is modeled as

$$R_{n,l} = \boldsymbol{\beta}_{\mathrm{I},n} \cdot \mathbf{X}_{n,l} + \alpha_l + \varepsilon_{\mathrm{I}} \tag{5.1}$$

$$\mathbf{X}_{n,l} = \boldsymbol{\beta}_{\mathrm{II},n} \cdot \mathbf{Y}_{n,l} + \boldsymbol{\gamma}_l + \boldsymbol{\varepsilon}_{\mathrm{II}} \tag{5.2}$$

$$\mathbf{Y}_{n,l} = \boldsymbol{\beta}_{\mathrm{III},n} \cdot \mathbf{E}_{n,l} + \boldsymbol{\theta}_l + \boldsymbol{\varepsilon}_{\mathrm{III}} \tag{5.3}$$

where $R_{n,l}$ is the product rating for n^{th} customer in l^{th} website; $\mathbf{X}_{n,l}$ are attribute aggregate ratings for n^{th} customer in l^{th} website; $\mathbf{Y}_{n,l}$ represents attribute ratings for n^{th} customer in l^{th} website; $\mathbf{E}_{n,l}$ are product specifications for n^{th} customer in l^{th} website; α_l represents the random effects of the product rating over the eight websites, $\alpha_l \sim \text{Normal}(0, \tau_{\mathrm{I}}^2)$, where τ_{I}^2 is the between-website variance for Level I; $\boldsymbol{\gamma}_l$ represents random effects of the attribute aggregate ratings $\mathbf{X}_{n,l}$ over the eight websites; random effects $\boldsymbol{\gamma}_l$ are identical within one aggregate rating because they are website-specific. Thus, the components of $\boldsymbol{\gamma}_l$ can be modeled as $\gamma_l \sim \textit{Normal}\ (0, \tau_{\mathrm{II}}^2)$, where τ_{II}^2 is the between-website variance for Level II; $\boldsymbol{\theta}_l$ represents random effects of the attribute ratings $\mathbf{Y}_{n,l}$ over the eight websites; random effects $\boldsymbol{\theta}_l$ are identical within one attribute rating since they are website-specific. Thus, the components of $\boldsymbol{\theta}_l$ can be modeled as θ_l ~ *Normal* $(0, \tau_{III}^2)$, where τ_{III}^2 is the between-website variance; $\boldsymbol{\beta}_{I,n}$, $\boldsymbol{\beta}_{II,n}$ and $\boldsymbol{\beta}_{III,n}$ are model parameters for n^{th} customer in Levels I, II and III, respectively. $\boldsymbol{\beta}_{I,n}$ ~ Multivariate Normal $(\boldsymbol{\mu}_I, \sum_I)$, $\boldsymbol{\beta}_{II,n}$ ~ Multivariate Normal $(\boldsymbol{\mu}_{II}, \sum_{II})$, $\boldsymbol{\beta}_{III,n}$ ~ Multivariate Normal $(\boldsymbol{\mu}_{III}, \sum_{III})$; ε_I, $\boldsymbol{\varepsilon}_{II}$ and $\boldsymbol{\varepsilon}_{III}$ are statistical error terms, ε_I ~ Normal$(0, \sigma_I^2)$; $\boldsymbol{\varepsilon}_{II}$ ~ multivariate Normal$(0, \boldsymbol{\sigma}_{II}^2)$; $\boldsymbol{\varepsilon}_{III}$ is independent and identically distributed extreme values.

5.2.2 Considering the Correlations of Multiple Observations across Model Hierarchy (Models III, IV and V)

As reviewed in Section 2.3, multiple observations on a single customer may be correlated. Three models are proposed to investigate the correlations of multiple observations in our model. The first model (Model III) assumes that the observations in each level are correlated, and the observations across different levels are independent. The second model (Model IV) assumes that the observations in the top two levels are correlated and the observations in the bottom level are correlated, but the error terms of rating data and the error terms of decision data are independent. The third model (Model V) assumes the errors across all the three levels as in Equations (5.1) to (5.3) are correlated. Model V is an investigation of the error correlation between rating data and decision data.

For the t^{th} observation of n^{th} customer in l^{th} website, the hierarchical model for customer preference is modeled as

$$R_{n,t,l} = \boldsymbol{\beta}_{I,n} \cdot \mathbf{X}_{n,t,l} + \alpha_l + \varepsilon_{I,n,t} \tag{5.4}$$

$$\mathbf{X}_{n,t,l} = \boldsymbol{\beta}_{II,n} \cdot \mathbf{Y}_{n,t,l} + \boldsymbol{\gamma}_l + \boldsymbol{\varepsilon}_{II,n,t} \tag{5.5}$$

$$\mathbf{Y}_{n,t,l} = \boldsymbol{\beta}_{III,n} \cdot \mathbf{E}_{n,t,l} + \boldsymbol{\theta}_l + \boldsymbol{\varepsilon}_{III,n,t} \tag{5.6}$$

$$\varepsilon_{\mathrm{I,n,t}} = \varsigma_{\mathrm{I,n}} + e_{\mathrm{I,n,t}} \tag{5.7}$$

$$\boldsymbol{\varepsilon}_{\mathrm{II,n,t}} = \boldsymbol{\varsigma}_{\mathrm{II,n}} + \mathbf{e}_{\mathrm{II,n,t}} \tag{5.8}$$

$$\boldsymbol{\varepsilon}_{\mathrm{III,n,t}} = \boldsymbol{\varsigma}_{\mathrm{III,n}} + \mathbf{e}_{\mathrm{III,n,t}} \tag{5.9}$$

where ς_{n} is a customer-specific error term; $\mathbf{e}_{n,t}$ is random error term across customers and observations.

In the first model (Model III) where there are no correlations across levels, $\varsigma_{\mathrm{I,n}}$ and $\boldsymbol{\varsigma}_{\mathrm{II,n}}$ are independent. The second model (Model IV) assume that there is a correlation between the errors in the top two levels, which is captured by letting $\varsigma_{\mathrm{I,n}}$ equal to the component in $\boldsymbol{\varsigma}_{\mathrm{II,n}}$. That is, the customer-specific error term is identical for the top two levels.

The third model (Model V) is inspired by the extant literature on modeling scale usage heterogeneity [40]. It is assumed that the error correlation across different levels is captured by λ_n. λ_n is a scale parameter for n^{th} customer across the three levels. For the t^{th} observation of n^{th} customer in l^{th} website, the hierarchical model for customer preference is modeled as

$$R_{n,t,l} = \boldsymbol{\beta}_{\mathrm{I},n} \cdot \mathbf{X}_{n,t,l} + \alpha_l + \lambda_n\, \varepsilon_{\mathrm{I}} \tag{5.10}$$

$$\mathbf{X}_{n,t,l} = \boldsymbol{\beta}_{\mathrm{II},n} \cdot \mathbf{Y}_{n,\,t,l} + \boldsymbol{\gamma}_l + \lambda_n\, \boldsymbol{\varepsilon}_{\mathrm{II}} \tag{5.11}$$

$$\mathbf{Y}_{n,\,t,l} = \boldsymbol{\beta}_{\mathrm{III},n} \cdot \mathbf{E}_{n,\,t,l} + \boldsymbol{\theta}_l + \lambda_n\, \boldsymbol{\varepsilon}_{\mathrm{III}} \tag{5.12}$$

The prior distribution of λ_n is $\ln(\lambda_n) \sim \mathrm{N}(0, \tau_\lambda^2)$.

5.3 Case Study

The smartphone case study used in Chapter 4 is employed for demonstration of the two models. The four hierarchical models (Model II, Model III, Model IV and

Model V) for customer preferences are built using customer reviews from the eight websites considered in the Chapter 4 and compared with the model (Model I) from Chapter 4.

5.3.1 Estimation Results from Model II

The posterior mean values $\bar{\mu}_\beta$ of parameter statistics estimated for Level I, Level II ("Features" only) and Level III ("Application_pros"only) are listed in Table 5.1 to Table 5.3. The posterior standard deviations $\bar{\sigma}_\beta$ are shown in the tables as an indication of customer preference heterogeneity. The significant variables are marked with an asterisk (*). As shown in Table 5.1, the effect of the four attribute aggregates on product ratings decreases from "value for price" to "design". This observation is consistent with the results of Model I in Table 4.7. The model estimates in Table 5.2 indicate that in general, attribute pros have positive effects on the attribute aggregate rating "features" and attribute cons have negative effects on the attribute aggregate rating "features." Noticeably, the weights of attribute cons are generally larger than those of attribute pros, indicating a larger effect of attribute cons on the attribute aggregate rating "features". This yields to the same conclusion drawn from Table 4.8. As shown in Table 5.3, several product specifications have significant effects on the rating for "applications pros." The quantity and quality of applications should be mainly determined by the operating system of a smartphone, which are observed in this model. The results suggest that customers like the applications of OS 1 and OS 2, but not the applications of OS 4. The results are similar to the results from Model I in Table 4.9 as well. Note that the number of significant variables of Levels II and III is smaller than that of Levels II and III in Model I. This is because by adding customer preference heterogeneity, the uncertainty of parameters gets larger.

Table 5.1 Model estimated results for Level I of Model II

Level I – parameter estimation for product rating R

Attribute aggregates	$\bar{\mu}_{\beta}$	$\bar{\sigma}_{\beta}$
Value for price*	0.37	0.03
Features*	0.29	0.04
Ease of use*	0.27	0.04
Battery life*	0.16	0.04
Design	0.06	0.03
Constant*	-0.71	0.20

Table 5.2 Model estimated results for Level II of Model II

Level II – parameter estimations for the aggregate "Features" X_2

Attributes	$\bar{\mu}_{\beta}\left(\bar{\sigma}_{\beta}\right)$	Attributes	$\bar{\mu}_{\beta}\left(\bar{\sigma}_{\beta}\right)$
Applications_pros*	0.49(0.14)	Applications_cons	-0.31(1.23)
Battery_pros	-0.06(1.24)	Battery_cons	0.04(0.41)
Keyboard_pros	0.04(0.48)	Internet_cons	-0.65(0.99)
Camera_pros	-0.03(0.13)	Keyboard_cons	-0.12(0.34)
Processor_pros	0.26(0.19)	Camera_cons	-0.09(0.21)
Quality_pros	0.10(0.11)	Processor_cons	-0.32(0.24)
Screen_pros	0.13(0.46)	Quality_cons	-0.59(0.35)
Text_pros	0.28(0.35)	Screen_cons	-0.62(0.40)
OS_pros	0.25(0. 21)	Size_cons	-0.27(0.15)
Ease of use_pros*	0.26(0. 11)	OS_cons	-0.64(0.52)
		Constant*	4.13(0.12)

Table 5.3 Model estimated results for Level III of Model II

Level III – parameter estimates for the attribute "Applications_pros" Y_1

Specifications	$\bar{\mu}_\beta\left(\bar{\sigma}_\beta\right)$	Specifications	$\bar{\mu}_\beta\left(\bar{\sigma}_\beta\right)$
Network	-0.48(0.43)	Touch screen	0.89(0.38)
OS 1*	1.59(0.49)	Battery capacity	-0.13(0.33)
OS 2	1.26(0.72)	Cam resolution	-0.24(0.20)
OS 3*	-1.56(0.45)	Video (HD)	1.31(0.54)
OS 4	0.28(0.38)	Video (regular)	0.57(0.42)
Height	-0.70(0.19)	Processor	0.11(0.53)
Width *	-3.37(0.23)	Memory	-0.16(0.44)
Depth	1.29(0.44)	Phone form*	-2.03(0.44)
Weight	0.29(0.19)	Keyboard*	1.50(0.38)
Display size*	2.03(0.22)	Wi-Fi*	1.38(0.36)
Total resolution	-0.32(0.24)	Constant	-0.89(0.35)

The variance σ_{I}^2 for the error terms ε_{I} and the covariance matrix $\boldsymbol{\sigma}_{\mathrm{II}}^2$ for the error terms $\boldsymbol{\varepsilon}_{\mathrm{II}}$ are listed in Equations (5.12) to (5.13). The meaning of the variance and the covariance matrix is the same as the ones in Model I. The values of the variance and covariance matrix are close to the values in Model I in Equations (4.9) and (4.10).

$$\sigma_{\mathrm{I}}^2 = 0.29 \tag{5.12}$$

$$\boldsymbol{\sigma}_{\mathrm{II}}^2 = \begin{bmatrix} 1.24 & 0.45 & 0.54 & 0.33 & 0.37 \\ 0.45 & 0.60 & 0.41 & 0.21 & 0.30 \\ 0.54 & 0.41 & 0.72 & 0.35 & 0.34 \\ 0.33 & 0.21 & 0.35 & 0.58 & 0.16 \\ 0.37 & 0.30 & 0.34 & 0.16 & 0.57 \end{bmatrix} \tag{5.13}$$

5.3.2 Estimation Results from Models III and IV

Same as before, the parameters and posterior standard deviations are listed in Table 5.4 to Table 5.6. The results from Models III and IV are close to each other and thus only the results from Model IV are listed here. The number of significant variables is close to that of Models II. The error terms for this model are estimated in the customer-level, thus not listed here.

Table 5.4 Model estimated results for Level I of Model IV

Level I – parameter estimates for product rating *R*;

Attribute aggregates	$\bar{\mu}_\beta$	$\bar{\sigma}_\beta$
Value for price	0.14	0.10
Features*	0.22	0.08
Ease of use*	0.19	0.09
Battery life*	0.20	0.09
Design	0.10	0.10
Constant*	0.73	0.16

Table 5.5 Model estimated results for Level II of Model IV

Level II – parameter estimates for the aggregate "Features" X_2

Attributes	$\bar{\mu}_\beta(\bar{\sigma}_\beta)$	Attributes	$\bar{\mu}_\beta(\bar{\sigma}_\beta)$
Applications_pros*	0.59(0.17)	Applications_cons	-0.41(0.23)
Battery_pros	-0.19(0.23)	Battery cons	-0.01(0.23)
Keyboard_pros	0.08(0.23)	Internet_cons*	-1.04(0.30)
Camera_pros	-0.01(0.23)	Keyboard_cons	-0.27(0.33)
Processor_pros	0.02(0.26)	Camera_cons	0.02(0.36)
Quality_pros	-0.01(0.20)	Processor_cons*	-0.64(0.24)
Screen_pro	0.14(0.21)	Quality_cons	-0.68(0.60)
Text_pros	0.25(0.21)	Screen_cons	-0.68(0.34)
OS_pros	0.16(0.21)	Size_cons	-0.50(0.28)
Ease of use_pros	0.20(0.19)	OS_cons	-0.66(0.52)
		Constant*	3.82(0.18)

Table 5.6 Model estimated results for Level III of Model IV

Level III – parameter estimates for the attribute "Applications_pros" Y_1

Specifications	$\bar{\mu}_\beta(\bar{\sigma}_\beta)$	Specifications	$\bar{\mu}_\beta(\bar{\sigma}_\beta)$
Network	-0.71(0.43)	Touch screen	1.17(0.38)
OS 1	1.91(0.49)	Battery capacity	-0.38(0.32)
OS 2	1.33(0.70)	Cam resolution	-0.24(0.20)
OS 3*	-1.61(0.45)	Video (HD)*	2.55(0.54)
OS 4	0.71(0.37)	Video (regular)	1.20(0.42)
Height*	-1.85(0.19)	Processor	-0.18(0.52)
Width*	-2.45(0.23)	Memory	-0.37(0.44)
Depth	0.41(0.44)	Phone form *	-2.29(0.44)
Weight	0.54(0.19)	Keyboard*	1.75(0.38)
Display size*	1.68(0.22)	Wi-Fi	1.00(0.36)
Total resolution	-0.23(0.24)	Constant	0.54(0.36)

5.3.3 Estimation Results from Model V

Same as Section 5.3.1, the parameters and posterior standard deviations are listed in Table 5.7 to Table 5.9. Generally speaking, the number of significant variables in Levels I and II of Model V is close to that of Model II. However, the value of the error terms is larger than that in Models I and II.

Table 5.7 Model estimated results for Level I of Model V

Level I – parameter estimates for product rating R;

Attribute aggregates	$\bar{\mu}_\beta$	$\bar{\sigma}_\beta$
Value for price	0.07	0.10
Features*	0.41	0.09
Ease of use*	0.25	0.09
Battery life*	0.27	0.09
Design	0.00	0.10
Constant	0.02	0.17

Table 5.8 Model estimated results for Level II of Model V

Level II – parameter estimates for the aggregate "Features" X_2

Attributes	$\bar{\mu}_\beta(\bar{\sigma}_\beta)$	Attributes	$\bar{\mu}_\beta(\bar{\sigma}_\beta)$
Applications_pros*	0.49(0.20)	Applications_cons	-0.38(0.46)
Battery_pros	-0.07(0.28)	Battery_cons	0.11(0.21)
Keyboard_pros	0.08(0.24)	Internet_cons	-0.60(0.44)
Camera_pros	0.00(0.20)	Keyboard_cons	-0.33(0.38)
Processor_pros	0.15(0.22)	Camera_cons	-0.06(0.68)
Quality_pros	0.06(0.20)	Processor_cons	-0.71(0.51)
Screen_pro	0.01(0.24)	Quality_cons	-0.74(0.55)
Text_pros	0.11(0.22)	Screen_cons	-0.82(1.06)
OS_pros	0.23(0.22)	Size_cons	-0.23(0.55)
Ease of use_pros	0.17(0.19)	OS_cons*	-0.68(0.33)
		Constant*	4.16(0.25)

Table 5.9 Model estimated results for Level III of Model V

Level III – parameter estimates for the attribute "Applications_pros" Y_1

Specifications	$\bar{\mu}_\beta(\bar{\sigma}_\beta)$	Specifications	$\bar{\mu}_\beta(\bar{\sigma}_\beta)$
Network	0.22(0.64)	Touch screen*	2.19(0.41)
OS 1*	1.80(0.42)	Battery capacity	-0.50(0.39)
OS 2*	2.76(0.37)	Cam resolution	-0.37(0.26)
OS 3	-1.18(0.49)	Video (HD)	1.90(0.47)
OS 4	0.76(0.51)	Video (regular)	1.01(0.51)
Height*	-0.97(0.25)	Processor	-0.17(0.57)
Width*	-3.50(0.25)	Memory	-0.72(0.68)
Depth	1.11(0.36)	Phone form *	-2.13(0.33)
Weight	0.40(0.25)	Keyboard*	1.25(0.34)
Display size*	1.89(0.28)	Wi-Fi	1.74(0.62)
Total resolution	-0.26(0.32)	Constant	-1.14(0.38)

The variance for the error terms $\lambda\ \varepsilon_{\mathrm{I}}$ and the covariance matrix for the error terms $\lambda\ \boldsymbol{\varepsilon}_{\mathrm{II}}$ are listed in Equations (5.14) to (5.15). The variances of error terms are generally greater than those in Models I and II especially a large value at the first component of

$\boldsymbol{\sigma}_{\mathrm{II}}$. The larger variance of the attribute aggregate “value for price” may explain why this variable is not significant in Model V. The individual error correlation $\tau_\lambda^{\,2}$ = 1.455.

$$\sigma_{\lambda\varepsilon}^{\,2} = 0.68 \tag{5.14}$$

$$\boldsymbol{\sigma}^2_{\lambda\boldsymbol{\varepsilon}_{\mathrm{II}}} = \begin{bmatrix} 1.82 & 0.40 & 0.52 & 0.32 & 0.72 \\ 0.40 & 0.96 & 0.52 & 0.20 & 0.44 \\ 0.52 & 0.52 & 1.36 & 0.40 & 0.52 \\ 0.32 & 0.20 & 0.40 & 0.96 & 0.28 \\ 0.72 & 0.44 & 0.52 & 0.28 & 1.60 \end{bmatrix} \tag{5.15}$$

5.3.4 Model Comparison

The performances of the four preference models are compared in terms of pseudo-r^2, RMSE, and DIC values, as shown in Table 5.10. Due to the nature of the regression in different levels, the pseudo-r^2 and RMSE values are employed for the comparison of the linear regression in the top two levels. The DIC values are employed for the comparison of the binary logit regression in the bottom level. The pseudo-r^2 value in Level I decreases from Model I to Model V while the RMSE value increases. This trend indicates that the model fitting quality becomes worse from Model I to Model V in Level I. However, the performance of Models I to IV is close to each other. Unlike Level I, from Model I to Model V in Level II, the model fitting becomes better as the pseudo-r^2 values increase and RMSE values decrease. The performance of Models III to V is similar and slightly better than the performance of Models I and II. For Level III, the DIC values are close to each other, the differences are not significant. Models III and IV in Level III are identical thus the DIC values are the same. More analysis of the model performance comparisons will be given with the validation results in next section.

Table 5.10 Model Performance Comparisons

	Models	I	II	III	IV	V
Level I	Pseudo-r^2	0.74	0.71	0.68	0.69	0.55
	RMSE	0.35	0.36	0.37	0.36	0.52
Level II	Pseudo-r^2	0.42	0.48	0.54	0.53	0.54
	RMSE	0.88	0.74	0.68	0.69	0.74
Level III	*DIC* ($\times 10^5$)	1.8	1.78	1.74	1.74	1.71

5.4 Model Comparison and Validation

A set of out-of-sample data are used to validate the proposed two hierarchical preference models (Models II and III). The out-of-sample data are downloaded from the three websites: www.yahoo.com, www.letstalk.com and www.viewpoints.com. The characteristics of the three websites are summarized in Table 5.11. The three websites are selected to represent three different groups specified in Chapter 4. The purpose of selecting the representative websites is to validate the performance of the proposed two models based on the representative websites. Some but not significant portion of customers summarized the pros/cons in the website www.viewpoints.com. Therefore, this website is regarded as the website without pros/cons because the reviews without pros/cons are picked for this validation study. Several reviews (10-20 reviews) for five or six smartphones are randomly downloaded from each website. The reviews are processed by following the procedure given in Chapter 4. The specifications **E** for each smartphone are used to predict the attribute ratings, attribute aggregate ratings and product ratings. The prediction errors are quantified by comparing the predicted ratings with the actual ratings.

Table 5.11Websites information

websites	Product rating	Attribute aggregate rating	General comments	Pros/cons
www.letstalk.com	Y	Y	Y	N
www.yahoo.com	Y	Y	Y	Y
www.viewpoints.com	Y	N	Y	N

The validation results are listed below in Table 5.12. The mean absolute percentage error *MAPE*, mean absolute error *MAE* and root mean squared error *RMSE* values are calculated for each level. Same as in Chapter 3, *MAPE* is not calculated for Level III. For the website www.viewpoints.com, the predicted attribute aggregate ratings cannot be validated since no attribute aggregate rating is provided. The five models work equally well regardless of selection of the website type. The results in Table 5.12 are the average results of the three websites. By comparing the performance of Model I with that of Models II to V, it can be seen that the performance of Model I, from the aspects of in-sample fit (Table 5.10) and out-of-sample error (Table 5.12), is consistently worse than that of Models II to V, which implies that the model prediction performance can be improved by taking into account customer preference heterogeneity. The performance of Models II to IV is close to each other in Table 5.10 and Table 5.12 – suggesting by considering the correlation of multiple observations, the model performance is not improved

Table 5.12Validation results

	Models	I	II	III	IV	V
Level I	*MAPE*	0.14	0.09	0.10	0.08	0.12
	MAE	0.59	0.38	0.44	0.36	0.51
	RMSE	0.64	0.44	0.49	0.42	0.57
Level II	*MAPE*	0.20	0.14	0.17	0.15	0.19
	MAE	0.79	0.52	0.64	0.54	0.77
	RMSE	0.91	0.62	0.72	0.69	0.87
Level III	*MAE*	0.28	0.21	0.19	0.19	0.20
	RMSE	0.30	0.24	0.22	0.22	0.22

5.5 Summary

This chapter focuses on proposing and validating three extended hierarchical preference models. This chapter attempts to relax two assumptions from Chapter 4 – no customer preference heterogeneity and no correlations of multiple observations from a single customer. The three models are applied for the same case study in Chapter 4. The out-of-sample data set from three representative websites were employed for a validation study. Model II outperforms in both the in-sample fit and out-of-sample validation. The results suggest that the model performance has been improved by taking account of customer preference heterogeneity. However, due to the fair performances of Models II, III and IV, it is suggested that the correlations of multiple observations in online customer reviews may not strong.

Chapter 6: Conclusion

This dissertation has focused on making use of online customer reviews for product design selection. After presenting introductory material and terminology in Chapters 1 and 2, we discussed the proposed approaches: (1) customer reviews from a single website was modeled for product design selection (Chapter 3), (2) customer reviews from multiple websites were integrated and modeled for product design selection with two strong assumptions (i.e., no customer preference heterogeneity and multiple observations on a single customer are independent) (Chapter 4), and (3) an extended study of integrating customer reviews from multiple websites by relaxing the two assumptions considered in Chapter 4 is made (Chapter 5).

In this chapter, highlights and concluding remarks from the proposed models presented in chapters 3-5 are presented in Section 6.1. The main contributions of this dissertation are discussed in Section 6.2. Several main limitations are discussed in Section 6.3. Finally, Section 6.4 outlines several possible extensions and research directions.

6.1 Concluding Remarks

This section is devoted to summarize each research thrust.

6.1.1 Product Design Selection using Online Customer Reviews from a Single Website

Chapter 3 presented a new approach for making use of customer reviews for customer-driven product design. The main assumption was that customer reviews from a single website can represent the customer voice of the entire market. In the proposed approach in Chapter 3, the existing text mining techniques were modified and extended to identify product attributes and elicit customer preference data from customer reviews. The finite mixture regression model was employed for modeling the customer data from customer reviews and modeling unobserved customer preference heterogeneity in the customer data. Finally, the model developed from customer reviews was used for the product design selection problem – selecting a best product design alternative that maximizes the profit.

The proposed approach involved three major steps. In the first step, the pros and cons of online customer reviews were divided into phrases as inputs. The frequent words were identified from the pre-processed pros/cons phrases. By applying two pruning rules, product attributes were identified as the retained frequent words. Customer ratings for the identified product attributes were modeled as a discrete variable. The customer rating of an attribute is divided into two ratings: the attribute pro rating and the attribute con rating. The value for the attribute pro rating is set to 1 if the attribute is specified as pros and 0 if it is not specified as pros. Similarly, the value for the attribute con rating is set to 1 if the attribute is specified as cons; otherwise 0 if not specified.

In the second step, customer data including product ratings, attribute aggregate ratings and elicited attribute ratings, along with product specifications, were used for modeling customer preferences. The finite mixture regression model was employed to capture unobserved customer preference heterogeneity. According to the relationship of different types of data, the customer preference models were built in a hierarchical fashion, linking from product ratings to product specifications. In the design selection, the profit of a design alternative can be estimated based on the assumption of the relationship between profit/demand and product ratings. The design with the maximal profit was selected as the desired design.

The smartphone case study was used to demonstrate the overall approach. Customer reviews were online customer reviews from a public website, www.bestbuy.com, and 305 sets of customer reviews were downloaded, from which 19 product attributes were identified. The finite mixture regression model was applied to model customer preferences. In the design process, three price scenarios were assumed: $99.99, $199.99 and $299.99. The design alternatives with the maximal profit were selected for each scenario.

The takeaway of the proposed approach is (i) online customer reviews are a good alternative of customer survey data to overcome the limitation of data acquisition in product selection; (ii) online customer reviews were successfully applied for product selection by the proposed approach; and (iii) the proposed approach was developed as a systematical approach to elicit, process and model customer reviews for product design selection by extending and assembling several existing techniques in the research area of text mining, customer preference models and design selection. The takeaway from the smartphone example is that the most profitable design alternatives are the designs with a faster processor, in a bar form and with a physical keyboard regardless of the product price. As the price goes up, the smartphone design with advanced equipments, such as 4g network, yields to a larger profit.

6.1.2 Product Design Selection using Online Customer Reviews from Multiple Websites with Website Heterogeneity

Chapter 4 presented a new methodology to integrate customer reviews from multiple websites for product design selection. At the beginning of Chapter 4, the differences between different public websites were summarized. The websites were divided into four groups according to what types of customer data are included in the websites. The four types of customer data can be collected: product rating, attribute aggregate ratings, general comments and pros/cons. The websites may contain all or part of the four types of data. It is complicated to integrate customer reviews from multiple websites because of the differences in different websites.

In order to integrate customer reviews from multiple websites, several main assumptions were made that (i) customer reviews from multiple websites represent the target market well; (ii) heterogeneity in customer data is mainly caused by website heterogeneity; and (iii) the responses from the same customer are independent. In the methodology, a text classification method was suggested to develop the attribute orientation classifiers by using the pros/cons data and then classify the orientations of the general comments. The meta-analysis technique was employed for integrating customer data and constructing customer preference models from multiple websites. It was shown that the meta-analysis technique could be used to integrate customer data by allowing the differences among the websites. Finally, the best design alternative that could maximize the profit was selected where the profit was modeled by using the customer model developed from customer reviews.

A text classification technique was suggested to systematically process customer reviews from multiple websites. The basic idea of the proposed technique is to learn attribute orientation information from pros/cons and use the learned information to process general comments from the websites without pros/cons. In the first step, a pool of product attributes was first identified from the pros and cons from available websites. Two dictionaries were constructed from the pros and cons, namely, an attribute dictionary and an attribute support dictionary. The two dictionaries were used to elicit attribute sentences from customer reviews in other websites. Then the product attributes, which are most frequently mentioned in all the websites, were identified as product attributes from all the websites. In the second step, customer attribute ratings were elicited. For the websites with pros/cons, customer rating can be easily elicited. For the websites with general comments only, the attribute orientation information was learned with customer rating data elicited from the websites with pros/cons. The SVM method was used to classify the attribute sentences into pros and cons based on the learned orientation information.

The meta-analysis technique was applied to integrate customer reviews from multiple websites. This technique enables modeling customer preferences under website heterogeneity. This study employed the same design process used in Chapter 3. The design alternative with the maximal profit was selected. In the smartphone case study, customer reviews from eight websites was selected and 932 customer reviews from the eight websites were downloaded and processed. The meta-analysis results were compared with those from the individual websites. Meanwhile, website heterogeneity was observed and analyzed. Finally, the preference model developed from multiple websites was used for product design, namely, selecting a design alternative with the maximal profit for three price scenarios.

The takeaway of the proposed approach is (i) online customer reviews from multiple websites were successfully integrated and applied for design selection; (ii) the proposed approach is applicable for any websites regardless of their own formats;

and (iii) the application of the meta-analysis technique properly models and explains the website heterogeneity existing among different websites. The takeaway of the smartphone case study is that the design selection results from multiple websites are close to the results from a single website. The consistent results imply that customer preferences from multiple websites are similar to each other.

6.1.3 Product Selection using Customer Reviews from Multiple Websites with Website and Customer Preference Heterogeneity

Chapter 5 attempted to extend the study in Chapter 4 by proposing and validating the three hierarchical preference models using customer reviews from multiple websites. The study considered two aspects: (1) customer preference heterogeneity and (2) the correlation of multiple observations from a single customer. The models were compared and validated using a set of out-of-sample data.

This chapter developed the hierarchical preference models by accounting for customer preference heterogeneity. It can be expressed in the parameters ($\boldsymbol{\beta}$s), which follow a population distribution over heterogeneous customers. In Chapter 5, two models are proposed to capture the correlations of multiple observations from a single customer.

The suggested hierarchical preference models were constructed using the same data set as the model in Chapter 4 – customer reviews from the eight websites. The results of the three models were compared with that in Chapter 4 to investigate the reality of customer preference heterogeneity and error correlations. Finally, the modeling results were validated using a set of out-of-sample data. The results suggest that by taking into account customer preference heterogeneity, the model performance has been improved. However, the models to deal with the error correlations do not work well. This may imply that the correlations of multiple observations in the case study are not strong.

The takeaway of the proposed approach is (i) the consideration of customer preference heterogeneity improves the model performance, which implies the existence of customer preference heterogeneity; (ii) the correlations of multiple observations from a single customer may not be strong in online customer reviews; and (iii) the models proposed in Chapters 4 and 5 are successfully validated by a set of out-of-sample data.

6.1.4 Discussion

Two issues in this dissertation are discussed in this section. First, brand is not considered as a variable in the proposed customer preference models. In this research, customer preference models are constructed and used for product design selection. In the product design selection process, designers have no controls on the design brand. What designers can do is to determine the product features. Therefore, including a brand variable in the customer preference models for product design selection is not helpful. However, it must be noted that there may exist the brand effects in the model estimates and thus design selection results. For example, in Figure 4.6, the effects of Samsung are different from the effects of Att and Tmobile in the Website II, although the three websites have similar formats of customer reviews. Att and Tmobile sell different brands of smartphons but Samsung only carries its own brand. Therefore, the brand effects may be one reason of the differences in the effects.

The second is the insignificant product design change with the price in the results. It can be observed that in Table 3.9 and Table 4.14 (design selection results from customer reviews from a single website and multiple websites respectively), as price goes up, the design only changes in two or three features and most other design values stay same. There are two main possible reasons for this insignificant design change. Firstly, the price is fixed and not changed with each design alternative. Usually, a customer's choice heavily depends on the product price. However, from online customer reviews, no price data can be collected and modeled. Customer choice is assumed to be a function of the product rating in our dissertation. Since the product price *PC* is fixed as a constant for all the alternatives, the expected profit *PF* is mainly determined by the cost of each design alternative based on the formulation in Equation 3.12. The formulation indicates that the design selected by maximizing the profit is basically the design with the lowest cost. When the price goes up, the selected design is changed in few features to keep the cost lower and consequently higher profit. Secondly, the product rating in customer reviews is not spread in a wide range. Most of the ratings are in the range of 3 to 5. This narrow range of the product rating may not provide enough change in customer preferences in terms of different product designs. This is a limitation of customer reviews used and may be one reason of insignificant design change with the product price.

6.2 Main Contributions

The main contributions are as listed in the following:

- A new approach to use online customer reviews for product design selection was developed. This approach extended the existing text mining techniques and assembled several existing methods in the research area of text mining, customer preferences modeling and product design. The approach can elicit customer data from online customers, construct customer preference models and select a desirable design alternative.

- A new text classification approach was proposed to elicit customer data from multiple websites. This proposed approach can elicit product attributes and customer data from multiple websites regardless of their formats. This approach combines an unsupervised classification method and a supervised classification method, which help to assure the accuracy of data elicitation.
- A new approach to integrate customer reviews from multiple websites was developed by applying the meta-analysis technique.
- A new hierarchical model was proposed by incorporating a mixture model with the meta-analysis technique in order to take into account both customer preference heterogeneity and website heterogeneity.
- Two new hierarchical models were proposed to take into account the correlations of multiple observations (reviews) from a single customer.

6.3 Limitations

This section discusses three major limitations of our proposed approach.

In this dissertation, product attributes are identified and customer data are elicited from online customer reviews through some modifications of text mining techniques. One shortcoming of these techniques is the dependence on human work. An attribute dictionary and an attribute support dictionary were constructed manually. The construction of such dictionaries can boost the performance of text mining techniques, but inevitably induce human intervene into the results.

Second, some possible biases on customer reviews are ignored in this dissertation. (i) The influence that customers might have over each other is ignored. Customers who review a product may first take a look at other reviews. Other reviews will inevitably have effects on the reviews. (ii) Online reviews are often managed by some professional companies, which may unduly influence customer reviews, for example, the companies will delete some reviews they thought not appropriate.

Finally, customer preferences are assumed to be unchanged during a certain time period. The evolution of customer preference over time is ignored.

6.4 Future Research Directions

This section discusses some possible future research directions. The directions may lead to the ways to overcome the limitations of the proposed approaches or extend the applications of the approaches.

6.4.1 Improvements in Product Attributes Identification and Customer Data Elicitation

Customer ratings for product attributes are elicited from the attribute orientation data – pros or cons. However, it might be difficult to distinguish the degree of the orientation based on customers' words. For example, adjective words, say "great" and "awesome", used by customers to describe attributes are not sufficient to understand which expression gives greater positive orientation. Thus, the main question is how to quantify the difference between different adjective words. Additionally, not only adjectives but also verbs are used to evaluate attributes, for instance, "love," "dislike," etc. The difficulty is how to more precisely sense the degree of the attribute orientation described by adjectives and verbs. If the above difficulties can be overcome and the degree of the attribute orientation can be elicited from the adjectives and verbs used to evaluate the attributes, the accuracy of customer preference models can be improved.

6.4.2 Consideration of Various Data Sources

No doubt, online customer reviews are a promising data source for customer-driven product design. However, there is one critical type of data not available in online reviews – purchase price. Price is an important variable for customer preference modeling. Unfortunately, no websites collects customer purchase price. And product price varies with time. Although the time of each customer review is recorded, it is difficult to find out the price change over time. Even if such information can be determined, there are still two issues. First, the actual purchase price is not always equal to the price on the market due to possible promotions. Second, the time when customer reviews are provided may not be at the same as the time of the product purchase. It is possible that a customer writes a product review a few weeks after the purchase and usage of the product.

Another important type of customer data that is missing from customer reviews is choice data. Although most customers who provide the reviews are real product users, from the reviews it is impossible to figure out how many similar products a customer chooses a product from. And some customers provide the reviews just after testing the products in the store. Without the choice data, an assumption of the relationship between purchase decision and product ratings has to be made.

Several types of customer reviews data are not used in this dissertation.

The first type is the recommendation. Some websites ask the customers whether they want to recommend a product to their friends. The willingness of giving a recommendation may reflect customer preferences of a product to some extent.

The second type of data not considered is the measurement of helpfulness. Most websites provide an option for the people who read the reviews to evaluate the reviews and say whether it is helpful. This may be considered in the preference models as well.

A few websites acquire some personal information, such as the usage – personal use or business use; the ownership – how long the customer owned the product. These types of personal information may be used to understand customer preference heterogeneity.

Finally, the two biases on customer reviews stated in Section 6.3 should be taken into account during the text mining or modeling process as one possible future research.

6.4.3 Improvements in Customer Preference Modeling and Design Selection

In this dissertation, customer reviews is the only data source for preference modeling and design selection. Although survey data acquisition requires significant costs and time, they are accurate and targeted for specific purposes. If customer reviews and survey data can be combined in an appropriate way, the preference modeling accuracy can be improved and the cost can be reduced.

Customer preferences evolve with time and product specifications change, especially for the fast evolving product category, like smartphones, E-readers, computers. As an extension, the time factor may be added into the customer preference models to capture the change of customer preferences over time.

Current research makes use of online customer reviews from United States. However, online customer reviews are widely available across different countries and should be a good data source for capturing customer preferences in different local and global markets.

The competition between different manufacturers is neglected in the proposed design selection process as well. Customer reviews can be grouped by different manufactures. By learning and modeling customer reviews from each manufacturer, some useful information for modeling competition between different manufacturers may be elicited.

References

[1] Bradlow, E. T., 2010, "User-Generated Content: The 'Voice of the Customer' in the 21st Century", Studies in Fuzziness and Soft Computing, Volume 258/2010, 27 – 29.

[2] Georgiopoulos, P., 2003, "Enterprise-wide product design: linking optimal design decisions to the theory of the firm", D.Eng. Thesis, University of Michigan, Ann Arbor, MI.

[3] Gupta, S.K. and Samuel, A.K., 2001, "Integrating market research with the product development process: A step towards design for profit", Proceedings of the ASME Design Engineering Technical Conferences, Pittsburgh, PA.

[4] Hazelrigg, G.A., 1988, "A framework for decision-based engineering design", ASME Journal of Mechanical Design, 120(4):653–658.

[5] Gu, X., Renaud, J.E., Ashe, L.M., Batill, S.M., Budhiraja, A.S. and Krajewski, L.J., 2002, "Decision-based collaborative optimization", Journal of Mechanical Design, 124(1):1–13.

[6] Michalek, J. J., Ceryan, O., and Papalambros, P. Y., Koren, Y., 2005, "Manufacturing investment and allocation in product line design decision making", Proceedings of the 2004 ASME IDETC Conference, Long Beach, CA.

[7] Kumar, D., Chen, W., and Simpson, T. W., 2006, "A market-driven approach to the design of platform-based product families," CD-ROM Proceedings of the AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Portsmouth, VA.

[8] Wunsch-Vincent, S., and Vickery, G., 2007, “Participative web: user-created content”, Report for the organisation for economic co-operation and development’s directorate for science, technology and industry as part of the WPIE work on digital content. <http://www.oecd.org/dataoecd/57/14/38393115.pdf>, last accessed on 11/26/2010.

[9] “User-generated content”, from Wikipedia, <http://en.wikipedia.org/wiki/User-generated_content>, last accessed on 11/26/2010.

[10] “Bazaarvoice”, <http://www.bazaarvoice.com>, last accessed on 05/06/2011.

[11] Commander, C., Analyst, The CMO Group, Forrester Leadership Boards, 2006, “Customer-driven design and development”, <http://www.forrester.com/role_based/pdfs/Customer-Driven_Design_And_Development_ReportBrief.pdf>, last accessed on 11/26/2010.

[12] Kaneko, T., Nakamura, Y., and Anse, M., 2007, “A method for generating plans for retail store improvements using text mining and conjoint analysis”, Lecture Notes in Computer Science, 4558, pp.910-917.

[13] Amarchinta, H., and Grandhi, R., 2008, “Multi-attribute structural optimization based on conjoint analysis," AIAA Journal, 46(4), pp. 884-893.

[14] Michalek, J. J., Feinberg, F. M., and Papalambros, P. Y., 2005, “Linking marketing and engineering product design decisions via analytical target cascading,” Journal of Product Innovation Management, 221, pp. 42-62.

[15] Li, H., and Azarm, S., 2000, “Product design selection under uncertainty and with competitive advantage,” Journal of Mechanical Design, 122(4), pp. 411-418.

[16]Lee, T., and Bradlow, E., 2006, "Automatic construction of conjoint attributes and levels from online customer reviews," The Wharton School, University of Pennsylvania, Working Paper OPIM WP 06-08-01.

[17]Wassenaar, H. J., and Chen, W., 2006, "An approach to decision-based design with discrete choice analysis for demand modeling," Journal of Mechanical Design, 125(3), pp. 490-497.

[18]Besharati, B., Luo, L., Azarm, S. and Kannan, P. K., 2006, "Multi-objective single product robust optimization: An integrated design and marketing approach," Journal of Mechanical Design, 128(4), pp. 884-892.

[19]Bradlow, E., 2005, "Current issues and a wish-list for conjoint analysis," Applied Stochastic Models in Business and Industry, 21(4-5), pp.319-323.

[20]Green, P., and Rao, V., 1971, "Conjoint measurement for quantifying judgmental data," Journal of Marketing Research, 8(3), pp. 355-363.

[21]Netzer, O., Toubia, O., Bradlow, E., Dahan, E., Evgeniou, T., Feinberg, F., Feit, E., Hui, S., Johnson, J., Liechty, J., Orlin, J., and Rao, V., 2008, "Beyond conjoint analysis: Advances in preference measurement," Marketing Letters, 19(3-4), pp. 337-354.

[22]Kumar, D., Hoyle, C., Chen, W., Wang, N., and Gomez-levi, G., 2007, "Incorporating customer preferences and market trends in vehicle package design," ASME International Design Engineering Technical Conferences, Las Vegas, NV, DETC2007-35520.

[23]Griffin, A., and Hauser, J. R., 1993, "The voice of the customer," Marketing Science, 12, 1-27.

[24] Akao, Y., 1994, "Development history of quality function deployment", The Customer Driven Approach to Quality Planning and Deployment, ISBN 92-833-1121-3, Asian Productivity Organization, Tokyo-107, Japan.

[25] Besanko, D., Gupta, S., and Jain, D., 1998, "Logit demand estimation under competitive pricing behavior: An equilibrium framework," Management Science, 44 (1), 1533-1547.

[26] Griffin, A., and Hauser, J. R., 1992, "Patterns of communications among marketing, engineering, and manufacturing—a comparison between two new product teams," Management Science, 38(3), 360-373.

[27] Anderson, S. P., De Palma, A., and Thisse, J., 1992 "Discrete choice theory of product differentiation", The MIT Press, Cambridge, MA.

[28] Hanson, W., and Martin, K., 1996, "Optimizing multinomial logit profit functions", Management Science, 42(7), 992-1003.

[29] Luo, L., Kannan, P. K., Besharati, B., and Azarm, S., 2005, "Design of robust new products under variability: marketing meets design," Journal of Product Innovation Management, 22(2), 177-192.

[30] Jiang K., Lee, H. L., and Seifert, R. W., 2006, "Satisfying customer preferences via mass customization and mass production", IIE Transactions, 38: 25-38.

[31] Lancaster, K., 1966, "A new approach to consumer theory", J. of Political Economy. 65: 567-585.

[32] Train, K., 2002, "Discrete choice: Methods with simulation", Cambridge University Press, Cambridge.

[33] Lindsay, B.G., 1995, "Mixture models: Theory, geometry, and applications". NSF-CBMS Regional Conference Series in Probability and Statistics, Vol. 5, Institute of Mathematical Statistics, Hayward.

[34] Smith, W., 1956, "Product differentiation and market segmentation as alternative marketing strategies," Journal of Marketing, 21, 3-8.

[35] Wedel, M., and Kamakura, W. A., 1998, "Market segmentation: conceptual and methodological foundations", Dordrecht, The Netherlands: Kluwer Academic Publishers.

[36] Akaike, H., 1974, "A new look at the statistical model identification", IEEE Transaction Automatic Control, 19, 716 - 723.

[37] Sarstedt, M., 2008, "Market segmentation with mixture regression models: Understanding measures that guide model selection", Journal of Targeting, Measurement and Analysis for Marketing, 16(3):228–246.

[38] Bozdogan, H., 1987, "Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions", Psychometrika , Vol. 52 , No. 3 , pp. 345 – 370 .

[39] Ashok, K., Dillon, W., and Yuan, S., 2002, "Extending discrete choice models to incorporate attitudinal and other latent variables," Journal of Marketing Research, 39, 31–46.

[40] Rossi, P.E., Gilula, Z., and Allenby, G.M., 2001, "Overcoming scale usage heterogeneity: A Bayesian hierarchical approach," Journal of the American Statistical Association, 96 (453), 20-31.

[41] Liu, B., Hu, M., and Cheng, J., 2005, "Opinion observer: analyzing and comparing opinions on the Web", WWW 2005, The International World Wide Web Conference, Chiba, Japan.

[42] Hu, M., and Liu, B., 2004, "Mining and summarizing customer reviews", Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004), , Seattle, Washington, USA, 168–177.

[43] "MySQL: full-text stop words", <http://dev.mysql.com/doc/refman/5.1/en/fulltext-stopwords.html>, last accessed on 11/28/2010.

[44] Salton, G. and McGill, M., 1983, "Introduction to modern information retrieval", New York, McGraw-Hill.

[45] Porter, M.F., 1980, "An algorithm for suffix strip-ping." Program, 14, 130–137.

[46] Salton, G., Wong, A., and Yang, C.S., 1975, "A vector space model for automatic indexing", In Communications of the ACM 18 (11), pp. 613 – 620.

[47] Nasukawa, T. and Yi, J., 2003, "Sentiment analysis: Capturing favorability using natural language processing", Second International Conference on Knowledge CAPture (KCAP 03), Sanibel, Florida, USA.

[48] Ghani, R., Probst, K., Liu, Y., Krema M., and Fano, A., 2006, "Text mining for product attribute extraction", SIGKDD Explorations 1(8) 41–48.

[49] Lee, T., 2004, "Use-centric mining of customer reviews", Workshop on Information Technology and Systems, New Orleans, LA.

[50] Snyder, B., and Barzilay, R., 2007, "Multiple aspect ranking using the good grief algorithm", Proceedings of the Human Language Technology Conference of the

North American Chapter of the Association of Computational Linguistics (HLT-NAACL 2007), , Rochester, NY, 300–307.

[51] Ghose, A., Ipeirotis, P. and Sundararajan, A., 2006, “The dimensions of reputation in electronic markets”, NYU Center for Digital Economy Research Working Paper No. CeDER-06-02.

[52] Hu, M., and Liu, B., 2004, “Mining opinion features in customer reviews”, Proceeding of the 2004 AAAI Spring Symposium Series: Semantic Web Services, San Jose, California, 755–760.

[53] Scaffidi, C., Bierhoff, K., Chang, E., Felker, M., Ng, H., and Jin, C., 2007, “Red opal: Product-feature scoring from reviews”, Proceedings of the 8th ACM conference on Electronic commerce (EC’07), San Diego, CA, 182–191.

[54] Popescu, A., and Etzioni, O., 2005, “Extracting product features and opinions from reviews”, Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP 2005), Vancouver, B.C., Canada, 339–346.

[55] Lee, T., and Bradlow, E., 2011, “Automated marketing research using online customer reviews”, Journal of Marketing Research, Working Paper.

[56] Das, S. R., and Chen, M., 2007, “Yahoo! for Amazon: sentiment extraction from small talk on the web”, Management Science 53(9) 1375–1388.

[57] Pang, B., and Lee, L., 2002, “Thumbs up? Sentiment classification using machine learning techniques”, Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-02), Philadelphia, PA, 79–86.

[58] Turney, P. D., 2002, “Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews”, Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL 2002), Philadelphia, PA, 417–424.

[59] Chen, P., Dhanasobhon S., and Smith, M., 2007, “All reviews are not created equal: the disaggregate impact of reviews and reviewers at Amazon.com”, International Conference on Information Systems (ICIS), Montréal, Quebec, Canada.

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

[61] Dellarocas, C., Awad N. F., and Zhang, X., 2004, “Exploring the value of online product ratings in revenue forecasting: The case of motion pictures”, Proc. 24th Internat. Conf. Inform. Systems, Washington D.C.

[62] Duan, W., Gu B., Whinston, A. B., 2008, “Do online reviews matter? An empirical investigation of panel data”, Decision Support Systems, 45(4), 1007 – 1016.

[63] Forman, C., Ghose, A., and Wiesenfeld, B., 2008, “Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets”, Information Systems Research, 19(3), 291 – 313.

[64] Archak, N., Ghose, A., and Ipeirotis, P., 2008, “Deriving the pricing power of product features by mining consumer reviews”, CeDER Working paper, New York University, New York.

[65] Feldman R., Goldenberg, J., and Netzer O., 2011, “Mine your own business: market structure surveillance through text mining”, Working Paper.

[66] Messer, M., Grotepa J., and Frenzel U.K., 2009, “Towards a function-based collective innovation framework”, Proceedings of IDETC/CIE 2009, ASME 2009 International Design Engineering Technical Conferences, San Diego, CA.

[67] Cortes, C. and Vapnik, V., 1995, “Support vector networks”, *Machine Learning*, 20:273–297.

[68] Yang, Y. and Liu, X., 1999, "A re-examination of text categorization methods", In Proceedings of SIGIR-99, 22nd ACM International Conference on Research and Development in Information Retrieval, pp. 42–49.

[69] Dumais, S., Platt, J., Heckerman, D., and Sahami, M., 1998, “Inductive learning algorithms and representations for text categorization,” in Proceedings 7th International Conference on Information and Knowledge Management.

[70] Pearson, K., 1904, “Report on certain enteric fever inoculation statistics,” British Medical Journal, 3:1243–1246.

[71] Sutton, A. J., Higgins, J. P., 2007, “Recent developments in meta-analysis”, Stat Med; 27:625–659.

[72] Dominici, F., Parmigiani, G., Wolpert, R., and Reckhow, K., 1997, “Combining information from related regressions”, Journal of Agricultural, Biological, and Environmental Statistics, 2(3):294–312. Good literature review on the application of hierarchical models to meta analysis, Page 4.

[73] Hsieh, C., and Pugh, M., 1993, “Poverty, income inequality and violent crime: A meta-analysis of recent aggregate data studies”, Criminal Justice REview, 18, 182 - 202

[74]DuMouchel, W., 1990, “Bayesian meta-analysis”, in Statistical methodology in the pharmaceutical sciences, ed. D Berry. New York: Dekker.

[75]Congdon, P., 2001, “Bayesian statistical modelling”, New York: John Wiley & Sons.

[76]Hedges, L. V., and Olkin, I., 1985, “Statistical methods for meta-analysis”, San Diego, CA: Academic Press.

[77]Hedges, L., 1997, “Bayesian meta-analysis”, pp. 251-275 in Statistical analysis of medical data, eds B Everitt and G. Dunn. London: Arnold.

[78]Silliman, N., 1997, “Hierarchical selection models with applications in meta-analysis”, J. American Statistical Association 92, (439), 926 – 936.

[79]Sutton, A. J., and Abrams, K. R., 2001, “Bayesian methods in meta-analysis and evidence synthesis”, Statistical Methods in Medical Research, 10:277–303.

[80]Egger, M., Davey Smith, G. and Phillips, A., 1997, “Meta-analysis: principles and procedures”, Br. Med. J., 315 1533–1537.

[81]Gilks, W. R., Richardson, S., and Spiegelhalter D. J., 1996 “Markov Chain Monte Carlo in practice”, London: Chapman and Hall.

[82]Menard, S., 1995, “Applied logistic regression analysis”, Sage university paper series on quantitative applications in social sciences, vol. 106. Thousand Oaks, California, pp 98.

[83] Willmott, C.J., Ackleson, S.G., Davis. R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O’Donnell. J., Rowe. C.M., 1985, “Statistics for the evaluation of model performance”, Journal of Geophysical Research, vol. 90(C5):8995–9005.

[84]Chung, F. L. and Lee, T., 1992, “A node pruning algorithm for backpropagation networks”, International Journal of Neural Systems, 3, 301-14.

[85]Willmott, C. J., and Matsuura, K., 2005, "Advantages (MAE) over the root mean square error (RMSE) performance", Climate Research, 30, 79−82.

[86]Hanke, J.E., and Reitsch, A.G., 1995, "Business forecasting", 5th ed., Prentice-Hall: Englewood Cliffs, NJ.

[87]Spiegelhalter, D., Best, N., Carlin, B. and Angelika L., 2002, “Bayesian measure of model complexity and fit,” Journal of Royal Statistics Society, 64 (4), 583–639.

[88]Luo, L., Kannan, P.K., Ratchford, B. T., 2008, “Incorporating subjective characteristics in product design and evaluations”, Journal of Marketing Research, 45(2), 182-194.

[89]Zeimpekis, D., and Gallopoulos, E., 2006, “TMG: a MATLAB toolbox for generating term-document matrices from text collections”, In J. Kogan, C. Nicholas, and M. Teboulle, editors, Grouping Multidimensional Data: Recent Advances in Clustering, pp. 187–210, Springer.

[90]“Adjective list”, <http://www.keepandshare.com/doc/12894/adjective-list>, last accessed on 11/28/2010.

[91]“Adverbs -- common List in American English”, <http://www.paulnoll.com/Books/Clear-English/English-adverbs.html>, last accessed on 11/28/2010.

[92]Agrawal, R. and Srikant, R., 1994, “Fast algoritm for mining association rules”, VLDB'94, Proceedings of the 20th International Conference on Very Large Data Bases, Santiago de Chile, Chile.

[93] Wedel, M., 2001, “GLIMMIX: Software for estimating mixtures and mixtures of generalized linear models”, Journal of Classification, 18(1), 129–137.

[94] Lunn, D. J., Thomas, A., Best, N. and Spiegelhalter, D., 2000, “WinBUGS- a Bayesian modelling framework: Concepts, structure, and extensibility”, Statist. Comput. 10 325-337.

[95] Yacob, M., and Shuib, A., 2009, “Assessing the preference heterogeneity in marine ecotourism attributes by using choice experiment”, International Journal of Economics and Management, 3(2): 367-384.

[96] Hosmer, D. W., and Lemeshow, S., “Applied logistic regression”, Wiley: New York, 2000.

[97] Tucker, C. S. and Kim, H. M., 2009, “Data-driven decision tree classification for product portfolio design optimization”, Journal Computer Information Science Engineering, 9(4), 041004 (14 pages).

[98] “Mouser electronics”, <http://www.mouser.com/>, last accessed on 02/20/2011.

[99] “Future electronics”, <http://www.futureelectronics.com/en/Pages/index.aspx >, last accessed on 02/20/2011.

[100] Weinberg, B., "Designing and deploying with mobile virtualization", White Paper. İ2009 Open Kernel Labs, Inc.

[101] Keller, K., "iPhone 4 carries bill of materials of $187.51, According to iSuppli", <http://www.isuppli.com/Teardowns/News/Pages/iPhone-4-Carries-Bill-of-Materials-of-187-51-According-to-iSuppli.aspx>, last accessed on 07/12/11.

[102] Maisto, M., "Palm Pre smartphone costs $170 to produce, Report Says", <http://www.eweek.com/c/a/Mobile-and-Wireless/Palm-Pre-Smartphone-Costs-17002-to-Produce-Report-Says-113476/>, last accessed on 07/12/11.

[103] Wang, L., Youn, B., Azarm, S., and Kannan, P., 2011, "Customer-driven product design selection using web based User-Generated Content", Proceedings of the 2011 ASME IDETC Conference, Washington, DC.

[104] Barrick, M. R., and Mount, M. K., 1991, "The Big Five personality dimensions and job performance: A meta-analysis", Personnel Psychology, 44, 1-26.

[105] Glass, G. V., 1976, "Primary, secondary, and meta-analysis of research", Educational Researcher, 5, 3-8.

[106] Peterson, R.A., 2001, "On the use of college students in social science research: insights from a second-order meta-analysis", Journal of Consumer Research 28 (3), pp. 450–461.

[107] Geyskens, I., Steenkamp, J. and Kumar, N., 1999, "A meta-analysis of satisfaction in marketing channel relationships." Journal of Marketing Research 36: 223-238.

[108] Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B.,1995, "Bayesian data analysis", London: Chapman and Hall.