AN INTEGRATED MODEL OF ONLINE RATING DECISION: ROLE OF PRE-PURCHASE EXPECTATION AND POST-PURCHASE EXPERIENCE

SOUMYA MUKHOPADHYAY
NANYANG BUSINESS SCHOOL

2015
AN INTEGRATED MODEL OF ONLINE RATING DECISION: ROLE OF PRE-PURCHASE EXPECTATION AND POST-PURCHASE EXPERIENCE

SOUMYA MUKHOPADHYAY

Nanyang Business School

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the degree of Doctor of Philosophy

2015
ACKNOWLEDGEMENTS

I am deeply indebted to many people who have supported and guided me through my graduate studies. I am especially grateful to my mentor and supervisor Dr. Chung Tuck Siong for all the cogent and critical comments that he patiently provided throughout various stages of my dissertation. He has always been open to new ideas and encouraged me to think independently. I am thankful to my dissertation committee members, Dr. Zhang Xiaoling and Dr. Shi Hongyan, for their invaluable advice and suggestions that helped me to improve the quality of my research.

I am also thankful to the entire faculty and staff in the Marketing and International Business (MIB) Division, NTU, for being so wonderfully supportive, helpful and understanding. Finally, I would like to express my heartfelt gratitude to my family and friends without whose kind and active support I could not have finished this candidature.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................................................... 1

TABLE OF CONTENTS ........................................................................................................................................... 2

LIST OF FIGURES ................................................................................................................................................... 3

LIST OF TABLES .................................................................................................................................................... 4

ABSTRACT ............................................................................................................................................................. 5

CHAPTER 1: INTRODUCTION .................................................................................................................................. 6

CHAPTER 2: MODEL DEVELOPMENT ...................................................................................................................... 17

CONCEPTUAL FRAMEWORK .................................................................................................................................. 17

MATHEMATICAL FRAMEWORK ............................................................................................................................ 19

CHAPTER 3: ASSESSMENT OF THE ESTIMATION ALGORITHM ................................................................................ 24

CHAPTER 4: EXPERIMENTAL DESIGN AND COVARIATE SPECIFICATION ....................................................... 35

EXPERIMENTAL PROCESS .................................................................................................................................... 35

COVARIATE SPECIFICATION ................................................................................................................................. 40

CHAPTER 5: ESTIMATION RESULTS ...................................................................................................................... 47

CHAPTER 6: SIMULATION STUDIES .................................................................................................................... 59

CHAPTER 7: DISCUSSION AND CONCLUSION ...................................................................................................... 68

REFERENCES .......................................................................................................................................................... 76

APPENDIX ............................................................................................................................................................... 81
LIST OF FIGURES

Figure 1: Individual Level Decision Making Process..................................................18
Figure 2: Geweke Plot for Beta.................................................................................29
Figure 3: Geweke Plot for Cut Points (Gamma).........................................................30
Figure 4: Geweke Plot for Error Correlations (Rho)..............................................30
Figure 5: Gelman Plot at 0.5 Correlation Level (Beta)............................................33
Figure 6: Gelman Plot at 0.5 Correlation Level (Gamma).........................................33
Figure 7: Gelman Plot at 0.5 Correlation Level (Rho)..............................................34
Figure 8: Experimental Interface 1 (Exposure to Prior Rating Data)......................37
Figure 9: Experimental Interface 2 (Consumption)..................................................38
Figure 10: Experimental Interface 3 (Disconfirmation Measurement)....................38
Figure 11: Experimental Screenshot 4 (Incidence Decision)....................................39
Figure 12: Geweke Plots for Beta Coefficients.........................................................53
Figure 13: Geweke Plots for Cutoff Points ...............................................................54
Figure 14: Geweke Plots for Error Correlations......................................................54
Figure 15: Gelman Plots (Beta)................................................................................57
Figure 16: Gelman Plots (Cutoffs)..........................................................................58
Figure 17: Gelman Plots (Error Correlations).............................................................58
Figure 18: Evolution of Rating Environment –Moderately Low Quality (Left Panel) Vs Moderately High Quality (Right Panel).................................................................61
Figure 19: Evolution of Rating Environment –Neutral Quality.................................63
Figure 20: Evolution of Rating Environment –High Negative Quality (Left Panel) Vs High Positive Quality (Right Panel).................................................................64
Figure 21: Observed Functional Relationship (Valence and Variance vs. Product Quality).................................................................66
LIST OF TABLES

Table 1: Variable Names and Corresponding Parameter Symbols ........................................ 25
Table 2: Summary of Posterior Marginal Distributions of Parameters .................................. 27
Table 3: Geweke Diagnostic for Estimated Parameters ....................................................... 28
Table 4: Heidelberger-Welch Diagnostic .............................................................................. 31
Table 5: Gelman Diagnostics at 0.5 Correlation Level (Beta) ............................................. 32
Table 6: Gelman Diagnostics at 0.5 Correlation Level (Gamma) ........................................ 33
Table 7: Gelman Diagnostics at 0.5 Correlation Level (Rho) ............................................. 34
Table 8: Description of Various Attributes of Musical Tracks ............................................ 36
Table 9: Correlation Matrix of Acoustic Attributes ............................................................. 40
Table 10: Total Variance Explained .................................................................................... 40
Table 11: Rotated Component Matrix .................................................................................. 41
Table 12: Correlation Matrix (Rating Environment Variables and Their Interactions) .... 43
Table 13: Total Variance Explained .................................................................................... 43
Table 14: Rotated Component Matrix .................................................................................. 44
Table 15: Variable Names and Corresponding Parameter Symbols ..................................... 48
Table 16: Summary of Posterior Marginal Distributions (Beta, Gamma and Rho) ............. 48
Table 17: Geweke Diagnostic for Estimated Parameters ....................................................... 51
Table 18: Heidelberger-Welch Diagnostic ........................................................................... 55
Table 19: Gelman Diagnostics (Beta) .................................................................................. 55
Table 20: Gelman Diagnostics (Cutoffs) .............................................................................. 58
Table 21: Gelman Diagnostics (Error Correlations) ............................................................. 58
Table 22: Comparison of the modeling approach of the current research with earlier works ...................................................................................................................... 73
Table 23: Comparison of findings of the current research with earlier works ....................... 74
ABSTRACT

The current literature mostly treats the rating decisions as a function of the concurrent rating environment characteristics and product experience without explicitly accounting for the discrepancy between a consumers’ expected and realized perceptions of product quality. To fill this literature gap, I developed a three stage sequential decision model that links expectation disconfirmation with post consumption rating decisions in the online context. I empirically examine the proposed model in a controlled laboratory setting where I was able to observe how individuals consume available product related information (rating) and how this information influences their product choice and subsequent rating decisions. I find that expectation disconfirmation play an important role in dictating individual level rating or review incidence decision as well as their rating evaluation decision. Using a number of simulated studies I show that the proposed model can successfully recreate a range of unique distributional (U-shaped or J-shaped distribution) and evolutionary characteristics (temporal change of valence and variance) of online rating environment that are commonly found across various contexts and platforms.

Keywords: Online opinion consumption and generation, online rating evolutionary characteristics, disconfirmation’s effect on rating decisions, three-stage sequential decision model, Sample Selected Probit, Music
CHAPTER 1: INTRODUCTION

BACKGROUND AND MOTIVATION:

The rapid advancement of online technologies has resulted in a plethora of platforms through which today’s consumers can share their thoughts and opinions about the products available in the marketplace. This has encouraged consumers to voluntarily contribute their experiences and opinions to the marketplace and to actively seek the benefit of learning from others’ experiences. Consequently, consumers have become more and more involved in both generation and consumption of online word-of-mouth (WOM) and the Internet has emerged as a dominant and pervasive source of product information that plays a decisive role in the success or failure of any product.

Given its increasing popularity and enormous impact on consumer decision making process, online reviews and its business implications have generated a lot of interests among marketing practitioners and researchers. In the last decade, a large number of studies have examined the online rating environment from various perspectives. Early scientific contributions in this field investigated how online ratings can effectively be measured and how they can be used as an easy and cost-effective proxy for conventional WOM (Godes and Mayzlin 2004). Subsequently electronic word-of-mouth (e-WOM) has been generally accepted and used as a reliable source of customer information for various research studies (Li and Hitt 2008). A parallel and more prolific stream of studies in this domain has focused on the relationship between online reviews and sales (Chevalier and Mayzlin 2006; Moe and Trusov 2011; Zhu and Zhang 2010). Majority of these studies
indicate that consumers rely on three basic characteristics of the online rating environment, namely valence (average numerical rating), volume (number of reviews), and variance (the dispersion of positive and negative ratings) to make their purchase decisions (Berger et al. 2010; Chintagunta et al. 2010; Dellarocas et al. 2007; Duan et al. 2008b). However, different studies show varying, and sometimes contradicting, impacts for these three characteristics of the rating environment. In more recent times researchers have examined a more diverse range of related topics including impact of reviewer identity disclosure on sales (Forman et al. 2008), role of trust and value in review helpfulness (Mudambi and Schuff 2010), presence of various bias in online customer reviews (Hu et al. 2008a; Hu et al. 2008b; Li and Hitt 2008) etc. These research works have not only enriched our understanding of how online opinions influence consumption decisions but also provided interesting and practical insights for marketers (Chen and Xie 2005). However, our understanding of the underlying process of rating contribution and its linkage with the emergence of various macro level characteristics (distributional and temporal) of online rating environment is still quite limited and incomplete.

It is generally recognized that a thorough understanding of the rating decision mechanism has a lot of practical and theoretical relevance. This is of practical importance to marketers because the Internet, unlike traditional sources of consumer information such as advertising, does not offer any direct and efficient regulatory mechanism. Consequently marketers have very little control over how product information is communicated online to their end consumers. Therefore a clear description of the integrated decision process can help marketers to identify and influence the factors that affect the tone and content of online opinions about their brand/product. In addition, from a theoretical perspective, a
better understanding of the underlying process of opinion creation and dissemination can enable researchers to explain some of the distributional and temporal characteristics of online product review system that are almost universally present across a wide variety of contexts and platforms. The ability to recreate these temporal and distributional characteristics can be considered as an important indirect evidence that a proposed model is indeed representative of the underlying dynamics of the rating environment. Consequently, using the proposed framework we can examine the functional relationship between intrinsic product quality and various aggregate measures of previously posted ratings. These functional relationships along with a clear exposition of underlying the process through which the underlying quality of the product leads to the emergence of various distributional and temporal dynamics of the rating environment can enable us to address an important question that has been examined by a related stream of literature: do the online ratings accurately reflect true product quality? The findings in this domain suggest that reviews may not reflect true product quality because they are subject to heterogeneous consumer tastes (Li & Hitt, 2008) and various social influences (Hu, Zhang, & Pavlou, 2009; Moe & Schweidel, 2012). This notion is also supported by the fact that the online rating distribution often follows a distinctive J-shaped pattern instead of an expected normal curve (Hu et al., 2009). This distinctive pattern is usually attributed to two biases: Self-selection bias and Under-reporting bias (Godes & Silva, 2012; Li & Hitt, 2008). Given these evidences that the online ratings in general reflect a slightly distorted view of the true underlying product quality and the fact that online rating environment is a complex system with multiple feedback mechanism, it can be expected that the extent of these distortions across various levels of underlying product quality
might follow an asymmetric non-linear path. In other words, we posit that each aggregate measure of online rating environment might have a distinctive functional relationship with the underlying product quality. A clear exposition of this functional relationship has significant managerial as well as theoretical importance. From a practical perspective, this understanding can help the firms to take a conscious and deliberated decision about product quality after giving due consideration to the possible response pattern that might eventually emerge in the online rating environment. On the other this such a description might shade some light on how the extent of distortions in the expressed online opinions vary when the intrinsic quality level changes from very low to very high value.

In one of the early attempts to model the rating decision mechanism in the context of designing an efficient recommendation system, Ying et al. (2006) suggested a two-step mechanism that governs the individual level rating decisions. In a more recent paper, Moe and Schweidel (2012) adopted a similar approach to empirically model online posting behavior in terms of two interrelated decisions: incidence decision (whether to contribute) and evaluation decision (what to contribute). They treated these rating decisions as an isolated (but interrelated) set of actions that are predominantly dictated by the consumer's post-purchase product evaluations and rating environment characteristics. Moreover, they examined how selection and adjustment effects influence review incidence and evaluation decisions. This two stage description of individual rating decision process is simple, parsimonious and it also has the ability to explain some of the observed evolutionary patterns in the rating environment. However, a consumers’ rating decision is a post-purchase action that is equivalent to engaging in word-of-mouth in the online environment. A large body of literature indicates that individual degree of satisfaction or
dissatisfaction is the most important antecedent of product related word-of-mouth (Anderson 1998; Rogers 1962; Westbrook 1987). The main antecedents of satisfaction, as identified by past research, are: expectations, perceived quality, and disconfirmation (Anderson and Sullivan 1993; Churchill and Surprenant 1982). Consequently, it can be argued that online rating decisions are also dictated not only by post-purchase product evaluations but also by pre-purchase expectation and the expectation disconfirmation. Therefore, a more complete description of the rating decision process must examine the pre-purchase and post-purchase processes together so as to account for the impact of each of these stages on rating contribution process. In this research I propose an integrated decision model that describes how pre-purchase expectation and post-purchase evaluations relate to subsequent rating decisions. In specific, I postulate that online rating information helps the consumers to form an initial (pre-purchase) opinion about the expected performance of the focal product or service (Zhao et al. 2013). This initial opinion plays a determining role in her choice or consumption decisions (Berger et al. 2010; Chintagunta et al. 2010; Duan et al. 2008a; Moe and Trusov 2011). Subsequently, in the post consumption phase, the consumer compares the actual product experience against the pre-purchase expectation (Anderson and Sullivan 1993) and forms a disconfirmation judgment (Halstead 1999). This disconfirmation judgment (the quantum of difference between pre-purchase expectations and post-purchase experience) in turn dictates her subsequent course of action (Churchill and Surprenant 1982) in terms of opinion incidence (whether to contribute) and evaluation (what to contribute).

I recognize that such a generalization requires additional information that is not easily available from secondary dataset. Therefore, unlike the past studies that use survey
or secondary data, I assess my empirical model in a laboratory setting that allow me to observe actual product choice and product consumption together with the consumption and contribution of product opinions while controlling for other external factors that affect secondary data. With product experience data I am able to investigate the effects of consumption experience on opinion incidences and evaluations. The consumption and construction of product opinions process are linked and thus my model estimates the influence of the different factors affecting these two processes simultaneously. By incorporating rating environment characteristics, product attributes, pre-purchase expectations and post-purchase evaluations in my model I am able to better understand their effects on these two processes. Additionally, using simulated studies (based on the estimated parameter values from my experiment) I show that the proposed three stage model can explain some of defining distributional and evolutionary characteristics of online rating environment. The initial choice stage coupled with a sequential rating decision mechanism (based on a disconfirmation framework) acts as natural filtering mechanism that rules out a large proportion of users from contributing a rating. In specific, the proposed model offers a natural explanation for the apparent paucity of data (sparseness) that characterizes many online rating systems.

The main theoretical contribution of this paper lies in developing a generalized framework to understand the role of pre-purchase expectation and post-purchase experience on rating decision. My results show that the expectation disconfirmation plays a crucial role in shaping rating decisions. Methodologically I contribute by developing the proposed three stage (two binary stages followed by an ordinal stage) sequential sample selected decision model with correlated random effects. Finally, I also demonstrate that
the proposed framework can mimic the evolutionary dynamics of the online rating environment that eventually leads to various distributional characteristics of online rating environments. These demonstrations not only offer an alternative explanation of the emergence of these characteristics but also provide an indirect support to the fact that the proposed framework is a realistic representation of the underlying decision mechanism.

The rest of this paper is organized as follows. In Chapter 2, I develop my conceptual framework and the empirical model. Chapter 3 describes the assessment of the proposed estimation algorithm using simulated data. My experimental design is described in Chapter 4. In Chapter 5, I present my estimation results and discuss the related insights. Chapter 6 presents the simulation results (using the empirical estimates). Concluding remarks and future research directions are provided in Chapter 7.

**Literature Review:**

The current research is related to the stream of research that examines the process underlying the generation and consumption of WOM. In past two decades researchers have tried to understand this process from various perspectives. In one of the early works in this direction, Sundaram et al. (1998) studied the motivation behind expression of positive and negative product opinions in an offline setting where they found that people are mostly driven by the basic feelings of altruism and vengeance respectively. During the same time Anderson (1998) proposed that relationship between satisfaction and WOM inherently follows a non-linear (U Shaped) pattern. Contrary to the contemporary notion of a monotonically increasing level of WOM activity with increasing satisfaction, he suggested that the consumers are more likely to express their opinion when they are extremely satisfied or dissatisfied with the product.
At the end of the last millennium the Internet quickly emerged as a dominant platform for shopping as well as for sharing product related information. This trend in turn led to the emergence of e-WOM as the most influential form of communication among consumers across the globe. Consequently, marketing researchers have focused on understanding this new form of communication and its impact on business environment. A considerable amount of effort has also been spent in order to better understand which aggregate review measure, valence (average numerical rating), volume (number of online reviews) or variance, can better predict future sales (Chevalier and Mayzlin 2006; Chintagunta et al. 2010; Dellarocas et al. 2007; Liu 2006; Moe and Trusov 2011; Zhu and Zhang 2010). Chevalier and Mayzlin (2006) demonstrated that the differences between consumer reviews posted on Barnes & Noble and those posted on Amazon.com were positively related to the differences in book sales via the two websites. Clemons et al. (2006) reported that the variance of online rating along with the most positive quartile of reviews play an important role in predicting the sales of craft beers. Godes and Mayzlin (2004) showed in a different setting that the "dispersion" of conversations about TV shows across online consumer communities and the popularity of these TV shows were strongly related. More recently, Monic Sun (2011) examined the informational role of rating variance. Based on fixed-price market and flexible-price market data they investigated how rating variance influences demand and profit for given level of average rating.

These studies have played a significant role in establishing the linkage between different characteristics (valence, volume and variance) of online review/rating environment and product sales. However, the results are inconclusive and sometimes
contradictory. Moreover they have also indicated the inherently dynamic nature of online rating environments and consequently a number of researchers have also studied the evolutionary dynamics of online consumer product ratings (Dellarocas and Narayan 2006; Duan et al. 2008a; Godes and Silva 2012; Hu et al. 2008b; Moe and Schweidel 2012; Moul 2007). In an experimental setup, Schlosser (2005) found that the opinions expressed by others often influence how posters adjust their product evaluations. Several other researchers have shown empirically that posted product ratings and reviews evolve over time and order (Li and Hitt 2008; Moe and Trusov 2011). In a recent paper, Godes and Silva (2012) have offered a number of alternative explanations for observed sequential and temporal dynamics of online opinions. In specific, most of these studies indicate that online reviews and ratings become increasingly negative as ratings environments mature. These evidences of systematic biases have eventually directed the research attention towards understanding the individual level online rating decision making process (Ho et al. 2013; Moe and Schweidel 2012; Ying et al. 2006).

These scholarly research works have undoubtedly contributed towards a better understanding of underlying process of rating contribution, its consequent impact on product sales and temporal/distributional evolution of online rating environment. However, despite this rapid growth of scholarly work, a number of questions remain open in this field. One such issue that deserves further attention is how pre-purchase expectation and post-purchase evaluations relate to subsequent rating decisions of an individual in terms of incidence (whether to contribute) and evaluation (what to contribute). Expectation-confirmation theory posits that expectations, coupled with perceived performance, lead to post-purchase satisfaction (Oliver 1980). It suggests that
the pre-purchase evaluation provides a benchmark against which the actual product experience is compared (Anderson and Sullivan 1993). It is generally hypothesized that expectations serve as the comparison standard that consumers use to evaluate performance and form a disconfirmation judgment (Halstead 1999). Disconfirmation, in turn, affects satisfaction, with positive disconfirmation leading to satisfaction and negative disconfirmation leading to dissatisfaction (Bolton and Drew 1991; Churchill and Surprenant 1982; Oliver and DeSarbo 1988; Tse and Wilton 1988; Woodruff et al. 1983). Given this widely accepted view that the quantum of difference between pre-purchase expectations and post-purchase experience has a definitive impact on the satisfaction / dissatisfaction and hence the future course of actions of individual consumers (Churchill and Surprenant 1982), it would be logical to expect that expectation-disconfirmation would play an important role in dictating an individual’s posting behavior in terms of whether to contribute (rating incidence) and what to contribute (rating evaluation). This proposition is also supported by Moe and Schweidel (2012).

In the present work I build upon this basic idea and develop an integrated model of individual posting behavior that takes into account the prior (pre-purchase) expectations of consumers. In the process I try to assess the impact of various online rating environmental characteristics on consumer choice decisions in a controlled experimental environment. Although this does not necessarily resolve the conflicting findings reported by earlier researchers but still it offers a much cleaner description of various influences that shape the consumer choice decisions. Moreover, the proposed model offers an alternative understanding of the evolutionary and distributional dynamics of online rating environment through an integrated framework. Moreover I start by developing a
conceptual framework and the corresponding econometric representation in the next chapter. Subsequently I develop an estimation strategy for the given econometric model.
CHAPTER 2: MODEL DEVELOPMENT

CONCEPTUAL FRAMEWORK

Extant literature (Moe and Schweidel 2012; Ying et al. 2006) on online rating dynamics posits that consumers posting behavior can be described in terms of two separate yet interrelated rating decisions: incidence decision (whether to contribute) and evaluation decision (what to contribute). Moreover, these two rating decisions are assumed to be influenced predominantly by the consumer's underlying product evaluations (Moe and Schweidel 2012). However, past literature (Anderson and Sullivan 1993) suggests that the process of expression of public opinion about a product is dictated not only by the realized post-consumption utility of the product but also by the pre-purchase expectation and utility disconfirmation. Therefore, an individual’s rating decisions can also be expected to be influenced by his or her pre-consumption expectation about the product performance, post consumption product evaluation and the discrepancy between these two. In this chapter, I develop an integrated model of the proposed underlying process through which pre-consumption expectation and post-consumption evaluations lead to subsequent rating decisions. The proposed model links individuals’ consumption experiences with the rating decisions.

Figure 1 represents a schematic outline of the proposed conceptual model. In specific I posit that during the pre-purchase evaluation stage individuals form an initial perception about the expected performance of the focal product or service. This perception is influenced by the available external information regarding the quality of the product. The source of external information in the context of this paper is the prior ratings given by other individuals (i.e. the existing online ratings). I recognize that there are other sources
of information (e.g. verbal communications, advertisements, price and packaging etc.) that can influence the pre-consumption expectation (these factors are represented by a dotted box on the left top corner in figure 1). However, taking all these sources of information brings in additional complexities into the modeling framework. Moreover, there is no way it can be ensured that the factors under consideration are sufficient and adequately represent all possible sources of product information. Consequently, I design my experiment to eliminate the impact of all the external sources of information other than the rating environment. This approach not only enables us to deal with the possible confounding effect of these extraneous influences but also offers a clean setup where we can examine the impact of various rating environment variables.

![Figure 1: Individual Level Decision Making Process](image)

Once a pre-consumption perception is formed, individuals take a decision to consume (or not to consume) the product. In the post-consumption stage, they use the pre-consumption perception about the expected product performance as a benchmark to compare the realized product performance and form a disconfirmation judgment. This
disconfirmation judgment along with realized product performance plays a crucial role in dictating the rating decisions.

**MATHEMATICAL FRAMEWORK**

Structurally the model has the following components: (1) a Binary Probit to represent first level decision to consume or not to consume a product (choice decision), (2) a second Binary Probit to develop a probabilistic representation of the decision to post or not to post a rating (incidence decision) and finally, (3) and an Ordered Probit to capture the decision making process regarding what rating to post (evaluation decision). I combine these components and propose a sequential three stage framework that integrates the pre-purchase expectation and post-purchase evaluation processes. In effect this sequential framework represents a generalized experience model that can be used to understand the subsequent incidence and evaluation decisions.

In specific, the proposed model describes N sets of observations where each observational set consists of three sequential decisions for a given product i:

- Whether to consume or not to consume a specific product (choice decision)
- Whether or not to rate a consumed product (incidence decision)
- What rating to assign given a favorable incidence decision (evaluation decision)

This decision sequence for a particular product i is modeled using a three equation system. The first equation represents the *choice decision* in terms of a binary (dichotomous) response variable \(y_{i1}\). A value of 1 for this response variable indicates that the product is consumed and a value of 0 indicates otherwise. Similarly, in the second equation, another binary response variable \(y_{i2}\) reflects the rating *incidence decision* and
a value of 1 for this response variable indicates that a rating is posted and a value of 0 indicates otherwise. Moreover, the value of $y_{i2}$ is observable only when $y_{i1}$ is 1. It is to be noted that in the present experimental setup we look at one product at a time. This implies that the individual is not really making a choice among multiple product options but he or she is making a decision (to consume or not to consume) about a single product option. Finally, the dependent variable in the third equation, denoted by $y_{i3}$, captures the ordered categorical responses (ratings) for observation $i$. This ordered response is observed only when both $y_{i1}$ and $y_{i2}$ are positive (indicated by 1). Each of these three decisions is linked to a corresponding underlying continuous latent variable. The value for these three latent variables (denoted by $y_{i1}^*$, $y_{i2}^*$ and $y_{i3}^*$) can be interpreted as the perceptual or utilitarian entity that drives the respective decision stage. In case of the choice decision, $y_{i1}^*$ represents the initial (pre-choice) quality perception of product option $i$. As it is already mentioned earlier, in our experimental setup we ensure that the pre-consumption evaluation is only dependent on rating environment. In other words, no other information (apart from rating) about the product option $i$ is made available to the respondents at the pre-purchase decision stage. Consequently, the initial quality perception can be represented purely by a function of prior rating information given by other users. I use three covariates ($X_{ii}$) derived from the valence (mean), variance, and volume (total number of ratings) and their two way interaction terms to represent this prior rating information. I refer to these three variables generically as the “rating environment characteristics”. The incidence decision depends on $y_{i2}^*$, which is expressed as a function of rating environment characteristics and the amount of disconfirmation in
the product quality experienced for product option \( i \). These variables are denoted by \( X_{i2} \).

Finally, the evaluation decision is dictated by the underlying quality opinion (\( y_{i3}^{*} \)) about the product option \( i \). The third latent variable \( y_{i3}^{*} \) is influenced by a set of covariates (\( X_{i3} \)) that includes rating environment characteristics and the latent quality (defined as a function of various product characteristics) of the product and its square term. Following prior research (Dellarocas and Narayan 2006; Moe and Schweidel 2012) I incorporated the square term to represent the possibility of any non-linear impact that the post purchase evaluation might have on the decision to post a rating (i.e. an extreme quality evaluation would have more pronounced effect on the rating decision than a moderate evaluation).

Each of these underlying continuous latent variables (\( y_{i1}^{*}, y_{i2}^{*} \) and \( y_{i3}^{*} \)) is assumed to be linear functions of the corresponding covariates as follows:

\[
y_{i}^{*} = \begin{pmatrix} y_{i1}^{*} \\ y_{i2}^{*} \\ y_{i3}^{*} \end{pmatrix} = \begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \end{pmatrix} + \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \varepsilon_{i3} \end{pmatrix} = X_{i1} \beta_1 + \varepsilon_{i1}, \quad X_{i2} \beta_2 + \varepsilon_{i2}, \quad X_{i3} \beta_3 + \varepsilon_{i3} \]

\[
\text{for, } i=1...N
\]

where \( \varepsilon_{i1} \sim N(0, \Sigma) \) or, \( y_{i}^{*} = X_{i} \beta + \varepsilon_{i} \) where \( \varepsilon_{i} \sim N(0, \Sigma) \)

In equation 1 and 2, \( X_{i1}, X_{i2} \) and \( X_{i3} \) represent vectors of decision specific covariates as described earlier. For simplicity I capture the interrelationship across different decision stages only through the error correlations. While, \( \beta_1, \beta_2 \) and \( \beta_3 \) represent parameter vectors corresponding to the covariates including the intercept terms;
I further assume that error components follow a trivariate normal distribution i.e. $\epsilon_i \sim N(0, \Sigma)$ Where,

$$
\Sigma = \begin{pmatrix}
1 & \rho_{12} & \rho_{13} \\
\rho_{12} & 1 & \rho_{23} \\
\rho_{13} & \rho_{23} & 1
\end{pmatrix}
$$

The off-diagonal elements ($\rho_{cd}$) of $\Sigma$ denote the conditional error correlations between latent continuous variables $y_{ic}'$ and $y_{sd}'$. The diagonal elements of $\Sigma$ are set to 1 for identification purpose. Given these specifications of the latent continuous variables I can write the complete decision process as follows:

$$
y_{i1} = I(y_{i1} > 0)
$$

$$
y_{i2} = I(y_{i2}' > 0) \times y_{i1}
$$

$$
y_{i3} = k \times y_{i2} \text{ if } \gamma_{k-1} \leq y_{i3}' < \gamma_k \text{ for } k=1,2,...K
$$

The equations 4 to 6 hold for $i=1,2,...N$ and $I(\cdot)$ represent an indicator function that takes a value of 1 when the argument is true. The ordered responses in equation 6 are assumed to be characterized by threshold values $\{\gamma_1, \gamma_2, ..., \gamma_{K+1}\}$ where $\gamma_k$ represents the $k^{th}$ cut-off. For identification purpose I define that $\gamma_1 = -\infty$, $\gamma_2 = 0$ and $\gamma_{K+1} = \infty$. Consequently, $\gamma$ for a one-dimensional ordinal variable with $K$ categories can be written as follows:

$$
\gamma = [(\gamma_1 = -\infty), (\gamma_2 = 0), ..., (\gamma_{K+1} = \infty)]
$$

The modeling approach employed herein can be seen as a generalization of the ordered probit that views an individual’s response as an outcome of three underlying latent processes. Because of the complexity of the proposed model, I conduct inference by using Monte Carlo Markov Chain simulation methods. My data sampler is designed
based on data augmentation strategy proposed by Tanner and Wong (1987), where I augment the parameter vector with latent continuous variables $y_{11}^*, y_{12}^*$ and $y_{13}^*$. I use a Metropolis-Hastings algorithm step with proposal density confined between (-1, 1) to generate $\rho_{1,2}, \rho_{1,3}$ and $\rho_{2,3}$ simultaneously. Similarly, I use a random walk Metropolis–Hastings (RWMH) step to generate $\gamma$. The detailed description of the estimation algorithm is available in the Appendix. I now evaluate how well my estimation algorithm performs using a simulation study. The details of the simulation studies are described in the next chapter followed by a detailed description of the experimental study.
CHAPTER 3: ASSESSMENT OF THE ESTIMATION ALGORITHM

As an initial test of my model, I ran a simulation study to test how well the proposed algorithm is able to recover the parameters from a simulated dataset. In this study I used a sample of 300 artificial agents; each of these individual agents was exposed to 100 product alternatives. The true population level parameter values are set at \( \beta_1 = (\beta_{1,0}, \beta_{1,1}, \beta_{1,2}, \beta_{1,3}) = (0.89, 1.00, 0.30, -0.45) \), \( \beta_2 = (\beta_{2,0}, \beta_{2,1}, \beta_{2,2}) = (-1.12, 2.00, -0.70) \), \( \beta_3 = (\beta_{3,0}, \beta_{3,1}, \beta_{3,2}, \beta_{3,3}) = (-0.77, 1.00, 0.30, -0.45) \) and finally \( \gamma = (\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6) = (-\infty, 0, 0.80, 1.60, 3.40, \infty) \). The subscripts \((l, m)\) in the beta coefficients \((\beta_{l,m})\) denote the decision stage and the covariate number respectively. The vector of three error terms was generated from the trivariate normal distribution with its mean being a vector of zeroes and a \(3 \times 3\) correlation matrix with off diagonal elements (i.e. the error correlations across three decision stages) were set at \( \rho = (\rho_{1,1}, \rho_{1,2}, \rho_{1,3}) = [(0.3, 0.3, 0.3), (0.5, 0.5, 0.5) \text{ and } (0.8, 0.8, 0.8)] \). The subscripts \((l, m)\) in the rho parameters \((\rho_{l,m})\) denote the respective decision stages. Whenever the random-walk Metropolis algorithm was used, the acceptance rate was controlled to be between .35 and .45 (Chib and Greenberg 1995; Roberts et al. 1997). The elements of \(X_{i1}, X_{i2}\) and \(X_{i3}\) were randomly generated from the standard normal distribution. Latent variables \((y^*_{i1}, y^*_{i2}\) and \(y^*_{i3}\)) representing the underlying decision process for each of these alternatives was generated using equation 2. Subsequently, these latent variables were used to decide the values of the final decisions i.e. \(y_{i1}, y_{i2}\) and \(y_{i3}\). 300 samples of 100 decisions each were generated from the above process, consequently the Bayesian sampler were used to estimate parameters based on these generated observations. Hyper-parameters of the priors were chosen as follows:
prior mean of $\beta = (\beta_1, \beta_2, \beta_3)$ values were set at 0 and prior precisions were set as $0.001*1$

where I is an $11 \times 11$ identity matrix. Ordinal threshold parameters ($\gamma_3, \gamma_4$ and $\gamma_5$) were set at values based on the proportion of cases in each cell of the ordinal measure based on the observed rating data at the third stage. The MCMC algorithm was run for 50000 iterations with first 1/3 of the total iterations as burn in. For the ease of referencing I present the variable names and their corresponding parameter symbols in table 1 and then present the results in table 2 onwards.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Intercept 1</th>
<th>$x_{11}$</th>
<th>$x_{12}$</th>
<th>$x_{13}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
<td>$\beta_{1,0}$</td>
<td>$\beta_{1,1}$</td>
<td>$\beta_{1,2}$</td>
<td>$\beta_{1,3}$</td>
</tr>
<tr>
<td>Variable</td>
<td>Intercept 2</td>
<td>$x_{21}$</td>
<td>$x_{22}$</td>
<td>$x_{23}$</td>
</tr>
<tr>
<td>Parameter</td>
<td>$\beta_{2,0}$</td>
<td>$\beta_{2,1}$</td>
<td>$\beta_{2,2}$</td>
<td>$\beta_{2,3}$</td>
</tr>
<tr>
<td>Variable</td>
<td>Intercept 3</td>
<td>$x_{31}$</td>
<td>$x_{32}$</td>
<td>$x_{33}$</td>
</tr>
<tr>
<td>Parameter</td>
<td>$\beta_{3,0}$</td>
<td>$\beta_{3,1}$</td>
<td>$\beta_{3,2}$</td>
<td>$\beta_{3,3}$</td>
</tr>
<tr>
<td>Variable</td>
<td>Error Correlation 1</td>
<td>$\rho_{1,2}$</td>
<td>$\rho_{1,3}$</td>
<td>$\rho_{2,3}$</td>
</tr>
<tr>
<td>Parameter</td>
<td>Error Correlation 2</td>
<td>$\rho_{1,3}$</td>
<td>$\rho_{2,3}$</td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>Error Correlation 3</td>
<td>$\rho_{2,3}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Variable Names and Corresponding Parameter Symbols

The estimated values of the parameters presented in the first, second and the third blocks of table 2 represent the summary of posterior marginal distributions in terms of the mean and the standard deviations. The mean estimate of each parameter obtained through MCMC is close to the true value of the corresponding parameter. The estimation results indicate that the proposed method can provide largely unbiased estimates of the $\beta$ coefficients even when the error terms are highly correlated with each other. I also found that all recovered parameter estimates are significant as the 95% confidence intervals do
not include zero. However, the recovery of the $\rho$ and $\gamma$ increasingly deteriorates with increasing level of error correlation. In general we find that up to the error correlation level of 0.5 the recovered values of $\rho$ falls within the expected norm of ±2SD range of the true values. When the correlation is set at 0.8, the recovered $\rho$ values fail to conform to this rule of thumb. On the other hand, it is found that the recovered value of $\gamma$ parameter is consistently lower than the actual parameter value. However, this underestimation of $\gamma$ parameter does not really affect our analysis as we don’t use this parameter to draw any conclusions.

### a. Estimates at Error Correlation Level: 0.3

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\beta_{1,0}$</th>
<th>$\beta_{1,1}$</th>
<th>$\beta_{1,2}$</th>
<th>$\beta_{1,3}$</th>
<th>$\beta_{2,0}$</th>
<th>$\beta_{2,1}$</th>
<th>$\beta_{2,2}$</th>
<th>$\beta_{3,0}$</th>
<th>$\beta_{3,1}$</th>
<th>$\beta_{3,2}$</th>
<th>$\beta_{3,3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Val</td>
<td>0.89</td>
<td>1.00</td>
<td>0.30</td>
<td>-0.45</td>
<td>-1.12</td>
<td>2.00</td>
<td>-0.70</td>
<td>-0.77</td>
<td>1.00</td>
<td>0.30</td>
<td>-0.45</td>
</tr>
<tr>
<td>Estimated Mean (*)</td>
<td>0.88</td>
<td>0.99</td>
<td>0.30</td>
<td>-0.45</td>
<td>-1.12</td>
<td>1.99</td>
<td>-0.68</td>
<td>-0.77</td>
<td>0.95</td>
<td>0.29</td>
<td>-0.42</td>
</tr>
<tr>
<td>Estimated SD</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.01</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\rho_{1,2}$</th>
<th>$\rho_{1,3}$</th>
<th>$\rho_{2,3}$</th>
<th>$\gamma_3$</th>
<th>$\gamma_4$</th>
<th>$\gamma_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Val</td>
<td>0.30</td>
<td>0.30</td>
<td>0.30</td>
<td>0.80</td>
<td>1.60</td>
<td>3.40</td>
</tr>
<tr>
<td>Estimated Mean (*)</td>
<td>0.29</td>
<td>0.23</td>
<td>0.28</td>
<td>0.70</td>
<td>1.40</td>
<td>3.25</td>
</tr>
<tr>
<td>Estimated SD</td>
<td>0.08</td>
<td>0.07</td>
<td>0.05</td>
<td>0.02</td>
<td>0.03</td>
<td>0.07</td>
</tr>
</tbody>
</table>

### b. Estimates at Error Correlation Level: 0.5

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\beta_{1,0}$</th>
<th>$\beta_{1,1}$</th>
<th>$\beta_{1,2}$</th>
<th>$\beta_{1,3}$</th>
<th>$\beta_{2,0}$</th>
<th>$\beta_{2,1}$</th>
<th>$\beta_{2,2}$</th>
<th>$\beta_{3,0}$</th>
<th>$\beta_{3,1}$</th>
<th>$\beta_{3,2}$</th>
<th>$\beta_{3,3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Val</td>
<td>0.89</td>
<td>1.00</td>
<td>0.30</td>
<td>-0.45</td>
<td>-1.12</td>
<td>2.00</td>
<td>-0.70</td>
<td>-0.77</td>
<td>1.00</td>
<td>0.30</td>
<td>-0.45</td>
</tr>
<tr>
<td>Estimated Mean (*)</td>
<td>0.89</td>
<td>1.00</td>
<td>0.31</td>
<td>-0.45</td>
<td>-1.13</td>
<td>2.01</td>
<td>-0.69</td>
<td>-0.78</td>
<td>0.94</td>
<td>0.28</td>
<td>-0.43</td>
</tr>
<tr>
<td>Estimated SD</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.01</td>
<td>0.03</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Parameter</td>
<td>$\rho_{1,2}$</td>
<td>$\rho_{1,3}$</td>
<td>$\rho_{2,3}$</td>
<td>$\gamma_3$</td>
<td>$\gamma_4$</td>
<td>$\gamma_5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>True Val</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.80</td>
<td>1.60</td>
<td>3.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimated</td>
<td>0.47</td>
<td>0.40</td>
<td>0.48</td>
<td>0.64</td>
<td>1.35</td>
<td>3.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimated</td>
<td>0.06</td>
<td>0.06</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\beta_{0,0}$</th>
<th>$\beta_{1,0}$</th>
<th>$\beta_{2,0}$</th>
<th>$\beta_{1,1}$</th>
<th>$\beta_{2,1}$</th>
<th>$\beta_{1,2}$</th>
<th>$\beta_{2,2}$</th>
<th>$\beta_{1,1}$</th>
<th>$\beta_{3,1}$</th>
<th>$\beta_{3,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Val</td>
<td>0.89</td>
<td>1.00</td>
<td>0.30</td>
<td>-0.45</td>
<td>-1.12</td>
<td>2.00</td>
<td>-0.70</td>
<td>-0.77</td>
<td>1.00</td>
<td>0.30</td>
</tr>
<tr>
<td>Estimated</td>
<td>0.88</td>
<td>1.01</td>
<td>0.31</td>
<td>-0.45</td>
<td>-1.12</td>
<td>2.06</td>
<td>-0.68</td>
<td>-0.73</td>
<td>0.94</td>
<td>0.29</td>
</tr>
<tr>
<td>Estimated</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.01</td>
</tr>
</tbody>
</table>

In order to check the convergence I ran a number of diagnostic tests on the simulated chains. First, I present the Geweke diagnostic (Table 3), which compares the location of the sampled parameter on two different time intervals of the chain. If the mean values of the parameter in the two time intervals are somewhat close to each other then I can assume that the two different parts of the chain have similar locations in the state space, and it is assumed that the two samples come from the same distribution. Usually one compares the last half of the chain, which is assumed to have converged (in order for the test to make sense), against some smaller interval in the beginning of the chain. For all

(*) All Parameter Estimates are significant as the 95% confidence intervals do not include zero

Table 2: Summary of Posterior Marginal Distributions of Parameters
parameters I used Geweke diagnostic with the pre-determined settings (fraction in first window is 0.1 and the second window are 0.1 and 0.5 respectively) in the Convergence Diagnosis and Output Analysis for MCMC (CODA) package. Usually, a z-score greater than 2 indicates that the mean of the series is still drifting, and a longer burn-in is required.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\beta_{1,0}$</th>
<th>$\beta_{1,1}$</th>
<th>$\beta_{1,2}$</th>
<th>$\beta_{1,3}$</th>
<th>$\beta_{2,0}$</th>
<th>$\beta_{2,1}$</th>
<th>$\beta_{2,2}$</th>
<th>$\beta_{3,0}$</th>
<th>$\beta_{3,1}$</th>
<th>$\beta_{3,2}$</th>
<th>$\beta_{3,3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>z Score</td>
<td>-1.82</td>
<td>-1.41</td>
<td>-2.15</td>
<td>1.22</td>
<td>1.51</td>
<td>0.19</td>
<td>-0.33</td>
<td>-0.79</td>
<td>-0.52</td>
<td>0.40</td>
<td>1.60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\gamma_3$</th>
<th>$\gamma_4$</th>
<th>$\gamma_5$</th>
<th>$\rho_{1,2}$</th>
<th>$\rho_{1,3}$</th>
<th>$\rho_{2,3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>z Score</td>
<td>1.09</td>
<td>1.18</td>
<td>0.98</td>
<td>-1.12</td>
<td>0.41</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Table 3: Geweke Diagnostic for Estimated Parameters

In the present case most of the z-scores for beta coefficients are well within the stipulated limit of 2 standard deviations around zero (Figure 2). In case of cut points and error correlation there is some evidence of non-stationarity as indicated by the deviations beyond the standard 2 SD norm (Figure 3 and 4):
Figure 2: Geweke Plot for Beta
The next convergence test I use is the Heidelberger-Welch diagnostic test (Heidelberger and Welch 1983) that offers convenient procedure to test the null hypothesis that the sampled values come from a stationary distribution. Their procedure is to be applied to a single chain. This test consists of two tests: Stationarity test and Half-width test. The stationarity test is used to check if the MCMC chain comes from a process with no temporal or spatial dependency in terms of its mean or its auto-covariance (covariance stationary process). The half-width test checks the adequacy of the chain size for accurate estimation of the posterior mean values.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\beta_{1,0}$</th>
<th>$\beta_{1,1}$</th>
<th>$\beta_{1,2}$</th>
<th>$\beta_{1,3}$</th>
<th>$\beta_{2,0}$</th>
<th>$\beta_{2,1}$</th>
<th>$\beta_{2,2}$</th>
<th>$\beta_{3,0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stationarity test</td>
<td>Pass</td>
<td>Pass</td>
<td>Pass</td>
<td>Pass</td>
<td>Pass</td>
<td>Pass</td>
<td>Pass</td>
<td>Pass</td>
</tr>
<tr>
<td>Start-iteration</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>13335</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>p-values</td>
<td>0.323</td>
<td>0.699</td>
<td>0.317</td>
<td>0.738</td>
<td>0.086</td>
<td>0.526</td>
<td>0.548</td>
<td>0.852</td>
</tr>
</tbody>
</table>
The results of Heidelberger-Welch diagnostic for different parameters are presented in Table 4. The results suggest that convergence was achieved immediately for the MCMC run for most of the variables. The half-width test indicates that all iterations should yield estimates of the posterior means for alpha and beta which meet the CODA default accuracy criterion of 0.1. Moreover, the posterior sample in each case passes the half-width test, and can provide a sufficiently precise estimate of the corresponding parameter.

Finally, I use Gelman-Rubin test to assess convergence that reports the Shrink factors along with their respective 50% and 97.5% quantiles. A value close to 1.00 for the shrink factor indicates that there is little evidence of dispersion between distributions to

Table 4: Heidelberger-Welch Diagnostic

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\beta_{3,1}$</th>
<th>$\beta_{3,2}$</th>
<th>$\beta_{3,3}$</th>
<th>$\gamma_3$</th>
<th>$\gamma_4$</th>
<th>$\gamma_5$</th>
<th>$\rho_{1,2}$</th>
<th>$\rho_{1,3}$</th>
<th>$\rho_{2,3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start-iteration</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5001</td>
<td>15001</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>p-values</td>
<td>0.961</td>
<td>0.662</td>
<td>0.368</td>
<td>0.331</td>
<td>0.627</td>
<td>0.065</td>
<td>0.105</td>
<td>0.456</td>
<td>0.486</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\beta_{1,0}$</th>
<th>$\beta_{1,1}$</th>
<th>$\beta_{1,2}$</th>
<th>$\beta_{1,3}$</th>
<th>$\beta_{2,0}$</th>
<th>$\beta_{2,1}$</th>
<th>$\beta_{2,2}$</th>
<th>$\beta_{3,0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.886</td>
<td>0.998</td>
<td>0.310</td>
<td>-0.449</td>
<td>-1.127</td>
<td>2.010</td>
<td>-0.694</td>
<td>-0.785</td>
</tr>
<tr>
<td>Half-width</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>0.002</td>
<td>0.000</td>
<td>0.002</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\beta_{3,1}$</th>
<th>$\beta_{3,2}$</th>
<th>$\beta_{3,3}$</th>
<th>$\gamma_3$</th>
<th>$\gamma_4$</th>
<th>$\gamma_5$</th>
<th>$\rho_{1,2}$</th>
<th>$\rho_{1,3}$</th>
<th>$\rho_{2,3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.941</td>
<td>0.282</td>
<td>-0.431</td>
<td>0.636</td>
<td>1.344</td>
<td>3.100</td>
<td>0.460</td>
<td>0.400</td>
<td>0.451</td>
</tr>
<tr>
<td>Half-width</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>0.002</td>
<td>0.007</td>
<td>0.006</td>
<td>0.005</td>
<td>0.003</td>
</tr>
</tbody>
</table>
which the chains are converging. The tables 5 to 7 show the potential scale reduction factors are very close to 1 for almost all parameters which indicate a good rate of convergence. The Gelman plots (figures 5 to 7) also support this conclusion.

**Potential Scale Reduction Factors**

<table>
<thead>
<tr>
<th>Point Estimate</th>
<th>1.00</th>
<th>1.00</th>
<th>1.00</th>
<th>1.03</th>
<th>1.00</th>
<th>1.04</th>
<th>1.00</th>
<th>1.01</th>
<th>1.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCI</td>
<td>1.00</td>
<td>1.02</td>
<td>1.00</td>
<td>1.00</td>
<td>1.12</td>
<td>1.00</td>
<td>1.00</td>
<td>1.16</td>
<td>1.00</td>
</tr>
<tr>
<td>Multivariate PSRF</td>
<td>1.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Gelman Diagnostics at 0.5 Correlation Level (Beta)
Figure 5: Gelman Plot at 0.5 Correlation Level (Beta)

<table>
<thead>
<tr>
<th>Potential Scale Reduction Factors</th>
<th>1.04</th>
<th>1.03</th>
<th>1.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point Estimate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper Confidence Interval</td>
<td>1.13</td>
<td>1.04</td>
<td>1.11</td>
</tr>
<tr>
<td>Multivariate PSRF</td>
<td></td>
<td></td>
<td>1.04</td>
</tr>
</tbody>
</table>

Table 6: Gelman Diagnostics at 0.5 Correlation Level (Gamma)

Figure 6: Gelman Plot at 0.5 Correlation Level (Gamma)

33
Potential Scale Reduction Factors

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Point Estimate</td>
<td>1.04</td>
<td>1.03</td>
<td>1.05</td>
</tr>
<tr>
<td>Upper Confidence Interval</td>
<td>1.13</td>
<td>1.04</td>
<td>1.11</td>
</tr>
<tr>
<td>Multivariate PSRF</td>
<td></td>
<td></td>
<td>1.04</td>
</tr>
</tbody>
</table>

Table 7: Gelman Diagnostics at 0.5 Correlation Level (Rho)

In the current chapter I presented the detailed assessment of the proposed estimation algorithm. The results from the artificially simulated dataset indicate that the algorithm is capable of recovering various parameters even at a high level of error correlations across the three stages. It also performs reasonably well in terms of various diagnostic tests that I used for checking the presence of non-stationarity. In the next chapter I present the experimental process along with a detailed description of various covariates that I used for the data analysis.
CHAPTER 4: EXPERIMENTAL DESIGN AND COVARIATE SPECIFICATION

EXPERIMENTAL PROCESS

Subjects in the experiment are undergraduate students from the University of [withheld for anonymity] who participated in the experiment for course credit. The experiment took place in the month of October 2013. The 146 subjects who participated in the experiment were selected without any screening criteria. Of the subjects 83% were aged between 19 to 21 years old, while the rest were aged between 22 to 24 years old. In addition, 41% of the subjects are male, while 59% of the subjects are female.

The experiment was conducted in two phases: first phase involved a simple evaluation (rating) task for a set of music clips analysis. This rating data was used to assess the relative preference of different product attributes for each respondent. The second phase of the experiment was conducted after one week after the first experiment. The second phase of the experiment was used to create a controlled environment where the individuals (respondents) can go through the three stage decision making process: a) consumption decision; b) incidence decision and c) evaluation decision. In the first (pre-consumption) stage each respondent is presented with a product option and the corresponding rating information (contributed by prior consumers of the product). Based on this rating information each respondent takes a decision (consumption) regarding the presented product option. In case of a non-favorable consumption decision the individual moves on to the next product option and is presented with the corresponding rating data. In case of a favorable decision the respondent is given a chance to consume the product. Once consumed, the individual is given an option to take a decision whether or not to rate
the product (incidence). In accordance with the incidence decision the individual is either taken to the next product option or presented with a scale to rate the product. Each respondent goes through this three stage sequential process for a number of product options. Given the experimental requirements I needed a product concept that is affordable (cost constraint), can be consumed quickly (time constraint) and repeatedly over a reasonably short period of time. Finally the product should be well defined in terms of a set of objectively defined attributes. Keeping these constraints in mind I decided to use short music clips (30-45 seconds duration) as my product concept. In addition to being cost effective, these music clips can be clearly described by a set of acoustic characteristics as defined by Echonest (table 8).

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description (Taken from Echonest Website)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tempo</td>
<td>Represents the BPM (Beats-Per-Minute) of the song in question represented by a number between 0 and 500</td>
</tr>
<tr>
<td>Danceability</td>
<td>This attribute is derived from a number of acoustic elements (tempo, rhythm stability, beat strength, and overall regularity) and it describes how suitable a track is for dancing (a value closer to 1 indicates better suitability for dancing)</td>
</tr>
<tr>
<td>Energy</td>
<td>This attribute is a scaled floating point metric from 0 to 1 that represents a perceptual measure of intensity throughout the track. It’s derived from perceived loudness, timbre, onset rate, and general entropy etc. of the music piece.</td>
</tr>
<tr>
<td>Loudness</td>
<td>Represents overall loudness of a track in decibels (dB).</td>
</tr>
<tr>
<td>Speechiness</td>
<td>This indicates the presence of spoken words in a track. Generally, values above 0.66 indicates very high verbal content, values between 0.33 and 0.66 indicates presence of music and words, while values below 0.33 represent non-speech musicals.</td>
</tr>
</tbody>
</table>

Table 8: Description of Various Attributes of Musical Tracks
I picked these clips in such a way so that I can cover maximum possible range in terms of their acoustic attributes. Consequently, each playlist used in the experiment contained tracks belonging to a wide variety of genres, origins and singers. The experimental process allows the individuals (respondents) to go through various stages of the consumption process. In the first (pre-consumption) stage each respondent is presented with a music track (product option) and the corresponding past rating data (Figure 8).

The past rating information is presented to the respondent in terms of a histogram that clearly shows the number of people who have given a particular rating. The respondent is also informed about the mean rating and the total number of people who have rated (volume). This information is provided at the top of the plot in clearly legible
capitalized fonts. Based on this rating information each respondent decides whether or not to listen to the music track. This decision represents the first (choice or consumption) decision stage in my modeling framework.

When the individual decides to click on the “Move to next” button, the present music clip is skipped and the individual moves on to the next music track and is presented with the corresponding rating data. Alternatively, when the individual click on the “Ok, let’s hear it” button, the respondent is given a chance to play and listen (consumption) to the music track (Figure 9).

![Figure 9: Experimental Interface 2 (Consumption)](image)

Once the track is completed, the individual is asked to indicate their evaluation of the product compared to their expectation (disconfirmation judgment) on a seven point scale (ranging from “extremely inferior” to “extremely superior”). This scale is adapted from
standard disconfirmation question (Oliver 1980; Westbrook 1980) where the respondents are asked how closely the experience with a particular music clip matches with his or her expectations (Fig 10).

Subsequently, they are given an option to decide whether or not to rate the product (incidence). In accordance with the incidence decision the individual is either taken to the next product option or presented with a scale to rate the product (Fig 11).

Each respondent goes through this three stage sequential process for 42 unique music tracks (the different product options). The data representing the corresponding choice, incidence and evaluation decisions are collected from all the 146 respondents over a period of two weeks. It is to be noted that this experimental setup does not really allow the rating volume to build up from 0 at the beginning of the experiment. In other words, we take a decision snapshot for every product option by exposing the respondent to an existing rating environment with non-zero volume. The temporal growth and evolution of the rating environment cannot really be captured in the experiment because of the following reasons: first, the number of respondent is very small; second, for many products all the respondents may decide not to consume or rate; third, the complete lack of data for the initial choices, as we don’t provide any external information about the products other than the prior ratings.
COVARIATE SPECIFICATION

As mentioned earlier, I used various acoustic attributes defined by Echonest to the music clips (Table 8). These attributes have long been successfully used for musical recommendations and classifications. It will be reasonable to assume that these attributes are valid descriptors of musical characteristics of a song. These attributes are presented in the table 8. The respective descriptions above clearly indicate that some of these attributes can be expected to be correlated with each other. A pairwise correlational analysis reveals that some of acoustic attributes are indeed highly correlated (table 9) and cannot be used for further analysis without running into multi-collinearity problem.

<table>
<thead>
<tr>
<th></th>
<th>Tempo</th>
<th>Danceability</th>
<th>Energy</th>
<th>Loudness</th>
<th>Speechiness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tempo</td>
<td>1.000</td>
<td>.415</td>
<td>.396</td>
<td>.286</td>
<td>-.231</td>
</tr>
<tr>
<td>Danceability</td>
<td>.415</td>
<td>1.000</td>
<td>.441</td>
<td>.448</td>
<td>.145</td>
</tr>
<tr>
<td>Energy</td>
<td>.396</td>
<td>.441</td>
<td>1.000</td>
<td>.788</td>
<td>-.084</td>
</tr>
<tr>
<td>Loudness</td>
<td>.286</td>
<td>.448</td>
<td>.788</td>
<td>1.000</td>
<td>-.105</td>
</tr>
<tr>
<td>Speechiness</td>
<td>-.231</td>
<td>.145</td>
<td>-.084</td>
<td>-.105</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 9: Correlation Matrix of Acoustic Attributes

Consequently, I ran a factor analysis across these 5 attributes to extract a reduced set of 3 uncorrelated dimensions. The result of the factor analysis is presented in table 10 and 11.

<table>
<thead>
<tr>
<th>Component</th>
<th>Rotation Sums of Squared Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eigenvalues</td>
</tr>
<tr>
<td>M1</td>
<td>1.873</td>
</tr>
<tr>
<td>M2</td>
<td>1.365</td>
</tr>
<tr>
<td>M3</td>
<td>1.123</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis

Table 10: Total Variance Explained
Using an eigenvector score of 1 as a cutoff, I could extract three mutually orthogonal factors that collectively explain 87.23% of total variance across all 5 acoustic attributes. Based on the respective compositions of these three factors I named them as Energy (Energy and Loudness), Rhythm (Tempo and Danceability), and Verbosity (Speechiness).

<table>
<thead>
<tr>
<th>Component</th>
<th>M1 (Energy)</th>
<th>M2 (Rhythm)</th>
<th>M3 (Verbosity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tempo</td>
<td>.137</td>
<td>.889</td>
<td>-.269</td>
</tr>
<tr>
<td>Danceability</td>
<td>.396</td>
<td>.697</td>
<td>.375</td>
</tr>
<tr>
<td>Energy</td>
<td>.898</td>
<td>.250</td>
<td>-.046</td>
</tr>
<tr>
<td>Loudness</td>
<td>.941</td>
<td>.143</td>
<td>-.038</td>
</tr>
<tr>
<td>Speechiness</td>
<td>-.072</td>
<td>-.071</td>
<td>.952</td>
</tr>
</tbody>
</table>

Table 11: Rotated Component Matrix

The extracted set of factors (M1, M2 and M3) can be directly used as covariates in the third (evaluation) decision stage. However, I am not interested in estimating the relative contributions of these product specific attributes towards decision process. On the contrary, I am specifically interested in understanding the impact of overall product experience on further decision making process. Furthermore, experienced quality (EQ) as covariate enables me to examine the impact of various quality levels on the evolution of rating environment.

Consequently, I needed to transform these acoustic attributes to a surrogate measure that represents overall product quality experience. I used the data collected in the first phase experiment for these purpose. As mentioned earlier, the first phase experiment was a very simple exercise where each individual was exposed to a unique set of music clips and was asked to rate the music clip on a 5 point rating scale indicating their overall experience with the music piece. These ratings can now be modeled using an ordered
probit framework with each of the extracted acoustic dimensions (M1, M2 and M3) as covariates and the rating (overall experience) data from the first phase of my experiment as the limited dependent variables. Consequently the individual \( (i) \) level perception of product quality \( (Q_{i,m}) \) for a given music clip \( (m) \) is calculated in the following manner:

\[
Q_{i,m} = \alpha_{0,i} + \alpha_1 M_{1,m} + \alpha_2 M_{2,m} + \alpha_3 M_{3,m}
\]

Where,

\( \alpha_{0,i} = \) Individual specific intercept measured from in experiment phase 1

\( \alpha_n = \) Coefficient for the \( n^{th} \) acoustic factor measured from in experiment phase 1

\( M_{n,m} = \) \( n^{th} \) acoustic factor for \( m^{th} \) musical track used in experiment phase 2

Following this approach we were able to estimate the relative contribution coefficients \( (\alpha_1, \alpha_2 \ and \ \alpha_3) \) of each acoustic dimension towards overall product quality. These coefficients were later used to calculate latent product quality perceptions for each music clip used in second phase of the experiment. In specific, we used the value of their acoustic dimensions in conjunction with these linear coefficients to calculate the perceived quality. The standardized values of these latent quality \( (Q_{i,m}) \) estimates along with the rating environment factors \( (\text{intensity, valence and discordance}) \) were used as covariates for the third (evaluation) stage of my model.

Rating environment variables are important for our study because we want to understand how various decision stages are influenced by the information contained in the rating environment. We characterize the rating environment in terms of \text{valence, variance, and volume} of previously posted ratings and \text{their two way interaction terms}. However, some of these metrics are significantly correlated with each other. Consequently,
following Moe and Schweidel (2012), we factor analyzed the rating environment characteristics (valence, variance, and volume) and their interaction terms (variance × volume, volume × valence, valence × variance) in order to identify a smaller set of mutually unrelated covariates.

<table>
<thead>
<tr>
<th></th>
<th>Valence</th>
<th>Volume</th>
<th>Variance</th>
<th>IT 1</th>
<th>IT 2</th>
<th>IT 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence</td>
<td>1.00</td>
<td>0.01</td>
<td>-0.01</td>
<td>0.44</td>
<td>0.48</td>
<td>0.00</td>
</tr>
<tr>
<td>Volume</td>
<td>0.01</td>
<td>1.00</td>
<td>0.42</td>
<td>0.38</td>
<td>0.87</td>
<td>0.85</td>
</tr>
<tr>
<td>Variance</td>
<td>-0.01</td>
<td>0.42</td>
<td>1.00</td>
<td>0.88</td>
<td>0.36</td>
<td>0.81</td>
</tr>
<tr>
<td>IT 1</td>
<td>0.44</td>
<td>0.38</td>
<td>0.88</td>
<td>1.00</td>
<td>0.54</td>
<td>0.72</td>
</tr>
<tr>
<td>IT 2</td>
<td>0.48</td>
<td>0.87</td>
<td>0.36</td>
<td>0.54</td>
<td>1.00</td>
<td>0.74</td>
</tr>
<tr>
<td>IT 3</td>
<td>0.00</td>
<td>0.85</td>
<td>0.81</td>
<td>0.72</td>
<td>0.74</td>
<td>1.00</td>
</tr>
</tbody>
</table>

IT1 = Interaction Term Valence*Variance; IT2 = Interaction Term Valence*Volume; IT3 = Interaction Term Variance*Volume;

Table 12: Correlation Matrix (Rating Environment Variables and Their Interactions)

The standardized values of the resulting factors were used as covariates for all three stages of my model.

<table>
<thead>
<tr>
<th>Component</th>
<th>Rotation Sums of Squared Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eigenvalues % of Cumulative %</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
</tr>
<tr>
<td>F1</td>
<td>2.36  0.39  0.39</td>
</tr>
<tr>
<td>F2</td>
<td>2.27  0.38  0.77</td>
</tr>
<tr>
<td>F3</td>
<td>1.33  0.22  0.99</td>
</tr>
</tbody>
</table>

Table 13: Total Variance Explained

The factor analysis results in two underlying constructs that explain 99% of the observed variation in daily ratings environments (see Table 13). The first factor (F1) is strongly related to the volume of posted ratings and its interaction with valence and variance, whereas the second factor (F2) is purely represents valence. Finally, the third factor (F3) predominantly represents variance and its interaction with valence. I name the first factor
(F1) as opinion intensity and the third factor (F3) as opinion discordance and for the second factor I just retain the original variable name valence.

<table>
<thead>
<tr>
<th>Component</th>
<th>F1 (Intensity)</th>
<th>F2 (Valence)</th>
<th>F3 (Discordance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence</td>
<td>0.06</td>
<td>0.99</td>
<td>0.06</td>
</tr>
<tr>
<td>Volume</td>
<td>0.97</td>
<td>-0.06</td>
<td>0.21</td>
</tr>
<tr>
<td>Variance</td>
<td>0.22</td>
<td>-0.08</td>
<td>0.97</td>
</tr>
<tr>
<td>IT 1</td>
<td>0.22</td>
<td>0.38</td>
<td>0.90</td>
</tr>
<tr>
<td>IT 2</td>
<td>0.88</td>
<td>0.43</td>
<td>0.21</td>
</tr>
<tr>
<td>IT 3</td>
<td>0.73</td>
<td>-0.09</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Table 14: Rotated Component Matrix

Finally, apart from rating environment characteristics and overall product experience, I use expectation disconfirmation as my third covariate. As described earlier, the disconfirmation measure for each consumed product was collected on a scale of 1-7. Evidently, this measurement is available only when an individual decides to consume a product and consequently, for large number of products the disconfirmation values are missing. These missing values lead to a unique methodological problem because simultaneous estimation of the beta coefficients ($\beta$) and the error correlations ($\rho$) across all three stages requires that there are no missing values for any of the covariates. In order to overcome this problem, I computed the missing values following an approach that is similar to the method that I used to calculate the latent quality estimates. In specific, I impute the missing disconfirmation values via an ordered Probit equation using the raw disconfirmation values (collected in the second phase experiment for the music clips that were listened to) available to me as the limited dependent variable.
The predictors used in this Probit model are the acoustic attributes (M1, M2, and M3), that predicts product performance, and rating environment factors (F1, F2 and F3), that predicts product expectation, for the corresponding music clip. Therefore the individual (i) specific latent disconfirmation (\( \delta_{ij} \)) value for a given music clip j can be expressed as follows:

\[
\delta_{ij} = c_{0,i} + \lambda_1 M_{1j} + \lambda_2 M_{2j} + \lambda_3 M_{3j} + \zeta_1 F_{1j} + \zeta_2 F_{2j} + \zeta_3 F_{3j}
\]

Where,

\( M_{nj} \) = \( n \)th acoustic factor for \( j \)th musical track used in experiment phase 2

\( F_{nj} \) = \( n \)th rating environment factor for \( j \)th musical track used in experiment phase 2

\( \lambda_n \) = Coefficient for the \( n \)th acoustic factor

\( \zeta_n \) = Coefficient for the \( n \)th rating environment factor

\( c_{0,i} \) = Individual specific intercept

For the cases where actual disconfirmation values are available, I substitute it with the estimated latent disconfirmation measure values so as to be consistent with the imputed values. I recognize that this imputation of missing values can introduce some amount of bias in the estimation process as this process can affect the conditionality assumption of rating incidence with reference to choice. As an alternative strategy I tried to use a random coefficients model with individual level intercept where the coefficient parameters are assumed to be correlated across three decision stages. However, such a model did not converge well because of increasing scarcity of data with every sequential decision step.

In the present chapter I discussed the experimental process that I followed to collect rating decisions data. Moreover, I talked about the various covariates that have
been used for the final data analysis. Given these descriptions I now proceed to the estimation results and their interpretations.
I applied the proposed model to the data collected from the experiment. The binary probit equation for the choice decision stage had three components (intensity, valence, and discordance) extracted from the rating environment characteristics as the explanatory covariates. This is logical, given that I have deliberately chosen the products (music clips) which are completely unknown to respondent. This anonymity of the musical tracks, together with the fact that the respondents were not provided with any other information (e.g. artist, genre etc.) about the music clips at or prior to the choice stage make the choice decision completely dependent on the information contained in the prior rating environment. For the incidence stage I used latent disconfirmation estimates along with three rating environment components as the predictors. As mentioned earlier this latent disconfirmation estimates are derived by using an ordered probit model (equation 9) with raw disconfirmation measures as the dependent variable and various acoustic dimensions and rating environment factors as covariates. So in an indirect way these covariates incorporate the impact of product characteristics as well as the impact of the expectations (in terms of the prior information contained in the rating environment factors) on the incidence decision. Finally, in the evaluation stage I use the latent quality (Q) perception and its square term (SQ) along with three rating environment components (F1, F2 and F3) as covariates. As described in the covariate specification section, the latent quality (Q) for a given music track is calculated by using its acoustic attributes (M1,M2 and M3) in conjunction with the beta coefficients estimated by applying an ordered probit model (equation 8) on the rating data collected in the first phase experiment. The square term
(SQ) accounts for the possible non-linear impact that the post purchase evaluation might have on the decision to post a rating.

Stage 1
Latent Choice \( (y_{i1}^*) = \beta_{1,0} + \beta_{1,1} \cdot \text{Intensity} + \beta_{1,2} \cdot \text{Valence} + \beta_{1,3} \cdot \text{Discordance} + \text{Error Term 1} \)

Stage 2
Latent Incidence \( (y_{i2}^*) = \beta_{2,0} + \beta_{2,1} \cdot \text{Disconfirmation} + \beta_{2,2} \cdot \text{Intensity} + \beta_{2,3} \cdot \text{Valence} + \beta_{2,4} \cdot \text{Discordance} + \text{Error Term 2} \)

Stage 3
Latent Evaluation \( (y_{i3}^*) = \beta_{3,0} + \beta_{3,1} \cdot \text{Quality Experience} + \beta_{3,2} \cdot \text{Square Quality Experience} + \beta_{3,3} \cdot \text{Intensity} + \beta_{3,4} \cdot \text{Valence} + \beta_{3,5} \cdot \text{Discordance} + \text{Error Term 3} \)

<table>
<thead>
<tr>
<th>Variable Parameter</th>
<th>Error Correlation 1</th>
<th>Error Correlation 2</th>
<th>Error Correlation 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho_{1,2} )</td>
<td>( \rho_{1,3} )</td>
<td>( \rho_{2,3} )</td>
<td></td>
</tr>
<tr>
<td>Variable Parameter</td>
<td>Cut Point 1</td>
<td>Cut Point 2</td>
<td>Cut Point 3</td>
</tr>
<tr>
<td>( \gamma_{3} )</td>
<td>( \gamma_{4} )</td>
<td>( \gamma_{5} )</td>
<td></td>
</tr>
</tbody>
</table>

Table 15: Variable Names and Corresponding Parameter Symbols

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \beta_{1,0} )</th>
<th>( \beta_{1,1} )</th>
<th>( \beta_{1,2} )</th>
<th>( \beta_{1,3} )</th>
<th>( \beta_{2,0} )</th>
<th>( \beta_{2,1} )</th>
<th>( \beta_{2,2} )</th>
<th>( \beta_{2,3} )</th>
<th>( \beta_{2,4} )</th>
<th>( \beta_{3,0} )</th>
<th>( \beta_{3,1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated Mean (*)</td>
<td>0.93</td>
<td>0.98</td>
<td>0.32</td>
<td>-0.42</td>
<td>-0.80</td>
<td>0.65</td>
<td>0.30</td>
<td>0.14</td>
<td>-0.14</td>
<td>-0.74</td>
<td>0.51</td>
</tr>
<tr>
<td>Estimated SD</td>
<td>0.02</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0.06</td>
<td>0.03</td>
<td>0.05</td>
<td>0.03</td>
<td>0.03</td>
<td>0.10</td>
<td>0.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \beta_{3,2} )</th>
<th>( \beta_{3,3} )</th>
<th>( \beta_{3,4} )</th>
<th>( \beta_{3,5} )</th>
<th>( \rho_{1,2} )</th>
<th>( \rho_{1,3} )</th>
<th>( \rho_{2,3} )</th>
<th>( \gamma_{3} )</th>
<th>( \gamma_{4} )</th>
<th>( \gamma_{5} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated Mean (*)</td>
<td>-0.01</td>
<td>0.39</td>
<td>0.22</td>
<td>-0.11</td>
<td>0.44</td>
<td>0.35</td>
<td>0.29</td>
<td>0.57</td>
<td>1.24</td>
<td>3.08</td>
</tr>
<tr>
<td>Estimated SD</td>
<td>0.02</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.17</td>
<td>0.08</td>
<td>0.14</td>
<td>0.04</td>
<td>0.06</td>
<td>0.23</td>
</tr>
</tbody>
</table>

(*) All Parameter Estimates are significant as the 95% confidence intervals do not include zero

Table 16: Summary of Posterior Marginal Distributions (Beta, Gamma and Rho)

Note that all covariates were standardized before they were used in the final estimation algorithm. This is to facilitate easy comparison of the impact that various factors exert on the decision variables. I present my main findings and the briefly discuss
the convergence diagnostics for the estimation process. The parameter symbols for different variables are presented in table 15 and the main findings are presented in the table 16.

The beta coefficients estimated from the experimental data shows that in the first stage, opinion intensity of the rating environment has the highest impact (0.98) on choice decisions. This is followed by variance (-0.42) and valence (0.32). These coefficients indicate that a higher mean rating leads to higher quality perception and consequently a greater probability of choice or product consumption. A negative coefficient for discordance (F3) indicates that a lack of unanimity in the user opinions negatively impact the choice probability. On the other hand higher opinion intensity (F1) in the rating environment positively influences the choice or consumption probability of a product. The intercept value is strongly positive (0.93) indicating a strong inclination to try (consume) new products.

The coefficients for the second (incidence decision) stage reflect the impacts of the various rating environment and expectation disconfirmation variables. My coefficient estimates indicate that the strongest influencing factor at this stage is disconfirmation. The coefficient value (0.65) is positive indicating a greater inclination to post a new rating with increasing extent of disconfirmation. The impact of the three rating environment characteristics shows a similar pattern as the first stage decision process. Here, valence has a positive coefficient value of 0.14; this indicates that a more positive rating environment results in higher incidence tendency. In other words, consumers are more likely to post in a rating environment where the average opinion expressed by previous posters is high. Discordance (F3) has a negative coefficient (-0.14) implying that higher
level of disagreement among the previous posters generally leads to a lower inclination towards posting a rating. Finally, the opinion intensity apparently exerts the strongest influence (0.30) on the incidence decision. This impact is positive, which indicates an increasing predisposition for posting a rating in the presence of greater opinion intensity.

I also find that the estimated population level intercept has a value of -0.80. This indicates that in general there exists a very strong negative propensity to contribute to a rating environment.

The coefficient estimates for the third stage show a strong positive (0.51) influence of product quality on evaluation decisions. The square term of quality is found to have a very small marginally negative (-0.01) impact. I also find that in this stage all three rating environmental characteristics exerts quite substantial amount of impact on individual ratings assigned to the product. The positive sign of the coefficient for valence (0.22) indicates that with higher (lower) existing rating an individual tends to adjust the value of his or her rating upwards (downwards). For volume, the strongly positive coefficient (0.39) implies that with increasing opinion intensity, an individual tends to give a higher rating to the product. Each of these indicates the presence of a very strong network effect on the final rating decision assigned to a product by an individual.

It is also observed that the opinion discordance exerts a moderate level (-0.11) of negative influence in the final rating decision stage. The sign of the coefficient indicates that a higher disagreement among the prior opinions posted leads to a lower product evaluation and vice versa. In other words, in a more (less) unanimous environment people tend to give higher (lower) rating to a product. Once again the population level intercept
(-0.74) for evaluation shows a strong baseline negativity across the population in terms of their product evaluations.

The off diagonal values in the error correlation matrix show a moderate level of association across three decision stages. The correlation between choice and incidence propensities is 0.44. This positive correlation value indicates that individuals with higher propensity to consume a product tend to show a positive inclination towards contributing to the rating environment. A slightly lower level of correlation (0.35) exists between choice and evaluation propensities. This indicates that a higher propensity to consume a product goes hand in hand with a greater predisposition to rate higher. Finally, the correlation between the general proclivity towards incidence and evaluation is slightly lower (0.29). This indicates that a greater frequency of posting behavior in general results in a stronger likelihood to give higher rating to a consumed product.

In order to examine the convergence of the MCMC process I employed the same set of diagnostics that I used in the assessment of the estimation algorithm using simulated data. The first diagnostic is the Geweke diagnostic that is based on a test for equality of the means of the first 10% and the last 50% of a Markov chain. If the samples are drawn from a stationary distribution of the chain, then the two means are equal and Geweke's statistic has an asymptotically standard normal distribution.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\beta_{1,0}$</th>
<th>$\beta_{1,1}$</th>
<th>$\beta_{1,2}$</th>
<th>$\beta_{1,3}$</th>
<th>$\beta_{2,0}$</th>
<th>$\beta_{2,1}$</th>
<th>$\beta_{2,0}$</th>
<th>$\beta_{2,4}$</th>
<th>$\beta_{3,0}$</th>
<th>$\beta_{3,1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>z Score</td>
<td>-0.32</td>
<td>0.25</td>
<td>1.23</td>
<td>-0.07</td>
<td>1.06</td>
<td>1.95</td>
<td>-0.78</td>
<td>-0.00</td>
<td>1.19</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\beta_{3,2}$</th>
<th>$\beta_{3,3}$</th>
<th>$\beta_{3,4}$</th>
<th>$\beta_{3,5}$</th>
<th>$\gamma_{3}$</th>
<th>$\gamma_{4}$</th>
<th>$\gamma_{5}$</th>
<th>$\rho_{1,2}$</th>
<th>$\rho_{1,3}$</th>
<th>$\rho_{2,3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>z Score</td>
<td>-0.09</td>
<td>1.45</td>
<td>1.65</td>
<td>-2.07</td>
<td>1.19</td>
<td>1.41</td>
<td>-0.26</td>
<td>-0.71</td>
<td>0.83</td>
<td>-0.47</td>
</tr>
</tbody>
</table>

Table 17: Geweke Diagnostic for Estimated Parameters
Except for a few cases, where some points are outside the standard ±2 range indicating a moderate level of non-stationarity, most of the cases values are well within ±2 range. Consequently, it can assumed that a much longer might be required to achieve stationarity. Accordingly, I ran a much longer (500000 iterations with 200000 burn in) MCMC chain to eliminate any possible impact of non-stationarity on estimation results. However, even such a significantly longer estimation chain had little impact on the final estimation results.
Figure 12: Geweke Plots for Beta Coefficients
Figure 13: Geweke Plots for Cutoff Points

Figure 14: Geweke Plots for Error Correlations

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\beta_{1,0}$</th>
<th>$\beta_{1,1}$</th>
<th>$\beta_{1,2}$</th>
<th>$\beta_{1,3}$</th>
<th>$\beta_{2,0}$</th>
<th>$\beta_{2,1}$</th>
<th>$\beta_{2,2}$</th>
<th>$\beta_{2,3}$</th>
<th>$\beta_{2,4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start-iteration</td>
<td>1</td>
<td>1</td>
<td>5001</td>
<td>1</td>
<td>1</td>
<td>2501</td>
<td>1</td>
<td>2501</td>
<td>2501</td>
</tr>
<tr>
<td>p-values</td>
<td>0.62</td>
<td>0.29</td>
<td>0.05</td>
<td>0.95</td>
<td>0.13</td>
<td>0.07</td>
<td>0.13</td>
<td>0.10</td>
<td>0.06</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\beta_{3,0}$</th>
<th>$\beta_{3,1}$</th>
<th>$\beta_{3,2}$</th>
<th>$\beta_{3,3}$</th>
<th>$\beta_{3,4}$</th>
<th>$\beta_{3,5}$</th>
<th>$\gamma_{3}$</th>
<th>$\gamma_{4}$</th>
<th>$\gamma_{5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start-iteration</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5001</td>
<td>2501</td>
<td>NA</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>p-values</td>
<td>0.30</td>
<td>0.37</td>
<td>0.08</td>
<td>0.11</td>
<td>0.09</td>
<td>0.00</td>
<td>0.16</td>
<td>0.13</td>
<td>0.54</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\rho_{1,2}$</th>
<th>$\rho_{1,3}$</th>
<th>$\rho_{2,3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stationarity test</td>
<td>Pass</td>
<td>Pass</td>
<td>Pass</td>
</tr>
<tr>
<td>Start-iteration</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>p-values</td>
<td>0.25</td>
<td>0.38</td>
<td>0.18</td>
</tr>
</tbody>
</table>
In case of Heidelberger-Welch diagnostic (table 18) also I find that most parameters pass the stationarity test as well as the half-width test that indicates the sufficiency of the portion of the series for estimating the posterior mean with a certain given accuracy. In one specific case ($\beta_{1,2}$) this diagnostic test fails with a N.A output. However, when I use three chain estimates for Gelman diagnostic to check stationarity, it is found that all PSRF and MPSRF values are close to 1 indicating a lack of multi-modality and non-stationarity for all parameter values. This same evidence is also apparent in the Gelman plots.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\beta_{1,0}$</th>
<th>$\beta_{1,1}$</th>
<th>$\beta_{1,2}$</th>
<th>$\beta_{1,3}$</th>
<th>$\beta_{2,0}$</th>
<th>$\beta_{2,1}$</th>
<th>$\beta_{2,2}$</th>
<th>$\beta_{2,3}$</th>
<th>$\beta_{2,4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.93</td>
<td>1.00</td>
<td>0.31</td>
<td>-0.41</td>
<td>-0.80</td>
<td>0.64</td>
<td>0.30</td>
<td>0.14</td>
<td>-0.13</td>
</tr>
<tr>
<td>Half-width</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\beta_{3,0}$</th>
<th>$\beta_{3,1}$</th>
<th>$\beta_{3,2}$</th>
<th>$\beta_{3,3}$</th>
<th>$\beta_{3,4}$</th>
<th>$\beta_{3,5}$</th>
<th>$\gamma_3$</th>
<th>$\gamma_4$</th>
<th>$\gamma_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Half-width</td>
<td>Pass</td>
<td>Pass</td>
<td>Fail</td>
<td>Pass</td>
<td>NA</td>
<td>NA</td>
<td>Pass</td>
<td>Pass</td>
<td>Pass</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.72</td>
<td>0.51</td>
<td>-0.01</td>
<td>0.39</td>
<td>0.21</td>
<td>NA</td>
<td>0.57</td>
<td>1.24</td>
<td>3.08</td>
</tr>
<tr>
<td>Half-width</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>NA</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\rho_{1,2}$</th>
<th>$\rho_{1,3}$</th>
<th>$\rho_{2,3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Half-width</td>
<td>Fail</td>
<td>Pass</td>
<td>Pass</td>
</tr>
<tr>
<td>Mean</td>
<td>0.45</td>
<td>0.31</td>
<td>0.34</td>
</tr>
<tr>
<td>Half-width</td>
<td>0.04</td>
<td>0.03</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table 18: Heidelberger-Welch Diagnostic

Table 19: Gelman Diagnostics (Beta)
Figure 15: Gelman Plots (Beta)
<table>
<thead>
<tr>
<th>Potential Scale Reduction Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point Estimate</td>
</tr>
<tr>
<td>Upper Confidence Interval</td>
</tr>
<tr>
<td>Multivariate PSRF</td>
</tr>
</tbody>
</table>

Table 20: Gelman Diagnostics (Cutoffs)

![Gelman Plots (Cutoffs)](image)

In the next chapter I present my simulation studies to explore the evolutionary pattern of the rating environment based on the results obtained from my model. I also examine the distributional characteristics of the resulting rating environment through this simulation.

![Gelman Plots (Error Correlations)](image)

Table 21: Gelman Diagnostics (Error Correlations)

<table>
<thead>
<tr>
<th>Potential Scale Reduction Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point Estimate</td>
</tr>
<tr>
<td>Upper Confidence Interval</td>
</tr>
<tr>
<td>Multivariate PSRF</td>
</tr>
</tbody>
</table>

Figure 17: Gelman Plots (Error Correlations)
CHAPTER 6: SIMULATION STUDIES

In the simulation study an artificial environment was created to mimic the three rating decision stages that individuals go through in the lab experiment. The estimated beta coefficients for the various rating environments variables are used to predict the pre-consumption evaluation. The pre-consumption evaluation or the expected product performance enables me to simulate the individual level choice decision for a given product. Conditional on a positive choice decision, I computed the underlying latent utility (which represents the consumption experience) for the chosen product of an individual based on the product attributes. Following a difference score conceptualization (Tse and Wilton 1988) of expectation disconfirmation an individual’s consumption experience is subtracted from his or her product expectation to calculate the disconfirmation experienced. This disconfirmation component along with the various rating environmental characteristics was used to predict if a rating is posted by the individual (rating incidence). Finally given a positive incidence decision (i.e. given that a rating is posted) the latent evaluation variable is computed and translated into corresponding discrete ratings using the estimated cutoff thresholds from my estimation results. My simulation procedure thus proceeds as follows:

1. First, I simulate the decisions taken by a population of 5,000 individuals. The rating environment variables valence, volume and variance are initialized with a value of 0.
2. For each of the three decision levels I calculate the latent variables ($y_{i1}$, $y_{i2}$, and $y_{i3}$) based on the estimated population-level model parameters and a correlated error correlation matrix from my experimental results.
3. For each iteration the individual level pre-purchase evaluation of a product is sampled from $N(\mu = y_{i1}, \sigma = 1)$ and used to simulate the consumption or choice decision based on equation 4.

4. Given a positive choice decision (i.e., given that a product is consumed) I calculate the disconfirmation experienced by taking the difference between post-purchase evaluation and pre-purchase expectation (from step 3) for the product.

5. Based on the post-purchase product evaluation and the estimated disconfirmation experienced I simulate the latent incidence decision variable $y_{i2}^*$. Given $y_{i2}^*$, I use equation 5 to simulate the decision of an individual agent to post or not to post a rating for the product.

6. Given a favorable incidence decision for an individual agent I simulate the latent incidence decision variable $y_{i6}^*$ and then I use $y_{i6}^*$ to simulate the final rating posted based on equation 6.

7. If a new rating is posted, I record the accumulated ratings valence as well as the rating variance and volume. The value of these rating environments variables are used in the subsequent iteration.

8. I repeat step 3-7 for multiple iterations till the simulation reaches a stationary state where no new individual contributes to the rating environment.

Note that, in the simulation process the quality of the products (Q) is manipulated to create different scenarios. In specific I set the innate quality of the product at 5 different levels: highly negative (-4) and highly positive (4); moderately negative (-2) and moderately positive (2); and finally, neutral (0) quality level. For each quality level I track the dynamic evolution of the rating valence, variance, and volume.
The evolution of rating valence, variance and volume for a product at the moderate low quality level (a quality level of -2) are presented graphically in figure 18 (Left Panel). The rating valence shows a decreasing trend which stabilizes near a value close to 1.81.

Figure 18: Evolution of Rating Environment – Moderately Low Quality (Left Panel) Vs Moderately High Quality (Right Panel)
In this case the consumers express their negative consumption experience by consistently posting lower rating. This not only results in eventual negativity in the average rating but also results in greater agreement in the opinion about the product that gets reflected by the decreasing and stabilizing trend for variance. During the simulation the rating variance remained close to 1.31. This indicates a moderate level of consensus in the opinion across the population. The general population’s opinion of the product (in terms of the valence) is in agreement with the underlying product quality level. This is a case where the overall rating eventually turns negative and the final rating distribution becomes positively skewed.

In the case of moderately high product quality (i.e. a quality level of 2) we find a different trend for how valence evolves through time (Figure 18 – Right Panel). The valence increases over time and remain stationary in a reasonably high value close to 3.7. Interestingly, the variance in this case increases marginally over time. In addition, the average variance is somewhat higher (1.38) than the previous case indicating a slightly stronger level of disagreement in the opinions about the product. The final frequency distribution of the rating values is skewed to the left with a peak at 4.

For a product with neutral quality (i.e. quality level of 0) it is found that the artificial rating environment converges onto a rating value of around 3.12 (Figure 19). Rating valence stabilizes quickly while variance increases through time to a value close to 1.77. The variance in this case is comparatively higher than the case of moderately low or moderately high product quality. The apparent disagreement in the opinion about the product is further reflected in the final frequency distribution plot for the product rating. In
this case, a nearly bi-modal (U-Shaped) distribution (with a smaller peak at 1, a higher peak at 4) is observed.

![Graphs showing evolution of rating environment, neutral quality](image)

Figure 19: Evolution of Rating Environment – Neutral Quality

For highly negative quality products (i.e. product quality level of -4.0), the rating valence converges on to a low value (1.1) (Figure 20 – Left Panel). This value is consistent with the underlying product quality level. The rating variance is very low (.23) implying a very high level of opinion agreement. The final distribution of the rating values is almost completely skewed to the right (J Shaped). This final distribution is quite representative of the underlying low quality level of the product.
Finally, I examine the evolution of rating valence, variance and volume for product with highly positive quality (i.e. quality level of 4). The valence shows a strongly
increasing trend that finally settles down at a high value of 4.1 (Figure 20 – Right Panel). In contrast, variance shows a significantly decreasing trend before eventually converging upon a reasonably high value of 1.01. This indicates that even for product with highly positive quality it takes time for the population of individuals to figure out the product quality and when the quality is made known the population generally agree on their assessment of the product. The final distribution of the rating values is strongly skewed to the left (J Shaped). The shape of the rating distributions cannot be distinguished apart from being either left (in case of positive product quality) or right (in case of negative product quality) skewed.

The simulation results suggest that product quality is one of the key drivers that modulate the nature of the rating environment. Moreover, it is found that the three stage model can successfully reproduce some of the key distributional and evolutionary characteristics of the online rating environment that are commonly observed across various product categories. For example, it explains the emergence of both positively and negatively skewed unimodal distribution (J Shaped) patterns of rating environment. It also shows that a neutral quality can lead to a split opinion environment with a bi-modal (U Shaped) frequency distribution of average ratings. The simulations also indicate that under certain situations (when there is confusion about the of the product quality) the rating environment becomes more and more negative with time (Figure 18 and 19). This has been reported by some of the earlier literatures (Li and Hitt 2008) in this field. However, I also find that even a very bad product quality can also lead to the emergence of a mild positive trend (figure 20 – left Panel) in terms of the temporal evolution of valence. There
is no way to ascertain the reason behind this trend but this might just be an expression of non-conformity to extant public opinion.

From the simulation study I find that at the moderate quality range average product ratings almost linearly increase with product quality. However, the rate of growth falls at the higher and lower quality zone. In effect the relationship between quality and valence seems to follow a very weak sigmoid form. This apparent non-linearity can be well approximated by a quadratic functional form (figure 21 – Top Panel).

Figure 21: Observed Functional Relationship (Valence and Variance vs. Product Quality)
The relationship between the extent of disagreement (signified by the variance) in the rating environment follows a more complex pattern (figure 21 – Bottom Panel). Here it is observed that a neutral quality level usually leads to a maximum disagreement while extreme (too high or too low) product qualities usually lead to a strongly unanimous opinion. This is a quite an expected pattern because extreme quality levels are often more apparent to consumers than a neutral level of quality. However, more interestingly I observe that a highly positive product quality is associated with a much higher level of disagreement than a highly negative quality of product. It seems that there is a disagreement when it comes to the definition of what is good product compared to what is bad. However, it is to be noted that the product concept in the present case is music and the very nature of music is inherently subjective. Consequently, the apparent lack of agreement in the consumer opinion may be an outcome of this intrinsic uncertainty about the definition of good music. From the simulation data I found that the relationship between underlying product quality and the variance can be expressed quite well in terms of a cubic function.

In this section I presented the findings of the simulation studies based on the three stage model of online rating contribution. I move on to the next chapter to discuss my results and provide a conclusion to this paper.
CHAPTER 7: DISCUSSION AND CONCLUSION

In this thesis I examined the role of expectation disconfirmation in online posting behavior and attempt to understand the inter-relationship between individual’s consumption and rating decisions. Moreover, I tried to explain some of the commonly observed distribution and evolutionary characteristics of online ratings that have attracted a lot of research attention in the more current literature. In contrast to the earlier research that tries to understand the performance impact of online ratings and reviews, I focused on the individual level rating behavior. I develop a modeling framework that considers the rating decision making as a sequential three stage process that starts with the consumption decision and subsequently moves through the incidence and evaluation phases. I used a sample selected three stage sequential probit model to represent this process. In my model, I examine the role of the online rating environment (as captured by valence, volume and variance and their interactions) on each stage of this process. I also explicitly incorporated the disconfirmation in my framework to assess its relative impact on incidence decisions. Furthermore, in order to assess the proposed empirical model I developed a laboratory setting that would allow me to observe each step of the three stage decision process while controlling for other confounding factors that generally affect the secondary data.

One important result of our research is existence of substantial amount of positive correlation across all three decision stages. These cross decision correlations lead to a number of interesting observations: first, a greater propensity to choose a product is positively related to the propensity to contribute online ratings; second, a higher consumption is also positively related to a higher propensity to give a higher rating. In other words, people with greater baseline consumption propensity are more positive in
their opinion and are likelier to contribute to the rating environment. These findings, although new in the online rating dynamics literature, are consistent with well-established behavioral economics and social psychological theories such as Endowment effect (Thaler, 1980) and Cognitive Dissonance (Festinger, 1957). These theories suggest that consumers who have already acquired or consumed the product tend to have a higher evaluation of the product compared to those who have not.

Consistent with prior studies (Moe & Schweidel, 2012), I found that a positive rating environment (with high valence) tends to improve the probability of both consumption and incidence. It also tends to have a positive impact on the final rating evaluation decision. Furthermore, our results show that higher opinion unanimity increases the likelihood of product consumption and rating contribution, whereas greater amount of disagreement among the previous posters discourage such behaviors. It is also observed that opinion intensity is one of the crucial rating environmental variables that positively influence all the three decision stages. In other words, a higher level of opinion intensity improves the propensity to consume the product as well as the probability to contribute a higher evaluation rating. These results indicate that consumers tend to exhibit a strong propensity to conform to social norms when they consume a product or express their online opinions. As I have already mentioned in the literature review that earlier researchers have often reported contradictory impacts of various rating environment characteristics on product choice (sales). Some researchers (Clemons, Gao, & Hitt, 2006; Duan, Gu, & Whinston, 2008; Liu, 2006) have showed that increased volume leads to increased sales; while others (Chintagunta et al. 2010) have reported that volume does not really matter. Similarly, in some contexts it has been found that there is negative or no
relationship between valence and sales (Duan, Gu, & Whinston, 2008; Liu, 2006), while in other contexts researchers (Moe & Schweidel, 2012) have found a significant positive impact of valence on sales. In order to resolve these conflicts some recent work in this area have suggested that possibly product and consumer characteristics moderate the effects of rating environment on sales (Zhu & Zhang, 2010). The current work does not directly resolve these apparent anomalies reported by earlier researchers. However, contrary to these works that used noisy secondary data to draw these conclusions, I have assessed the impact of rating environment characteristics (in terms of valence, volume and variance) on choice decision in a controlled experimental environment that ensures the absence of any outside confounding factors. Therefore, my findings can possibly represent as a more precise description of the actual underlying dynamics.

The investigation of the functional relationship between various key measurement metrics and underlying product quality is another important aspect of current research. Additionally, I used these simulated environments to investigate the extent to which my proposed model can recreate the temporal and distributional dynamics of online rating platforms. We used a series of simulated studies (based on the estimated parameter values from our experiment) to examine these aspects of online review dynamics. Specifically, I simulated the rating environment under various conditions where I could manipulate the product quality level from a very low level to a very high level. These simulation results indicate that the average rating value of the posted ratings do not exactly reflect the underlying product quality but it gives a close enough approximation at the medium quality range. However, at the lower (upper) quality ends, the average online ratings reflect a more inflated (deflated) view of quality. An interesting pattern is also observed in
case of variance, which follows a non-linear functional relationship with product quality. In specific, the extent of opinion dispersions for moderate quality level were found to be much higher than the extent of dispersion at lower or upper quality end. It seems that for neutral product quality levels consumers are not able to assess the product quality clearly. However, as the quality level increases (or decreases) beyond this range the consumers become more and more certain about the product quality and they generally express their positive experiences much more boldly. Consequently, very high (or very low) quality levels lead to faster convergence to a highly polarized rating environment. Moreover, I find that the amount of variance is almost always higher for better product quality than it is for lower product quality. This is in agreement with the notion of positive-negative asymmetry (Peeters & Czapinski, 1990) in evaluation. It must be noted that an alternative explanation for the results depicted in Figure 21 is that the scales in question have a distinct upper and lower bound, attenuating the ratings. The equation for average rating looks like a logistic curve, which is consistent with this story. Similarly the smaller variance at both ends of the scale (which is consistent with a logistic curve), may result from attenuation. However, this effect is possible only when the underlying quality has a much broader range than that of the rating scale and hence in order to accommodate this wider notion of quality into the given narrower rating space users might need to attenuate their contributed opinions. In our simulation setup the underlying quality strictly adheres to a mean range (5 levels) that maps exactly with the rating scale range (5 levels). Consequently, I posit that there is a strong possibility that the simulation results are indeed representing an interesting underlying phenomena.
These simulation results also offer some indirect support for our proposed three stage model as it can successfully reproduce many of the key distributional and evolutionary characteristics of online rating environment. In contrast to much of the existing research (Godes & Silva, 2011; Li & Hitt, 2008) on online ratings that explained these temporal and distributional characteristics in terms of consumer characteristics, we offer an alternative explanation to these observations using product quality as a driver. However, it must be noted that despite offering alternative explanations for online review dynamics, the present research does not allow us to rule out the existing explanations.

Finally, I find that the expectation disconfirmation exerts a very strong influence on incidence decisions while quality exerts a strong influence on evaluation decision. Though, the impact of disconfirmation on conventional word-of-mouth is well documented in earlier marketing literature (Anderson 1998; Rogers 1962; Westbrook 1987), my results indicate that this conventional wisdom can very well be generalized to online platform of communication.

The main contribution of this paper is threefold. First, we propose and examine a generalized framework that explicitly incorporates the consumption as a part of the overall decision mechanism. Second, we contribute methodologically by developing a three stage (two binary stages followed by an ordinal stage) sample selected decision model with correlated errors. This three stage model (purchase, incidence and evaluation) is different from the existing two stage decision models that treated incidence and evaluation decisions as two distinct but interrelated individual level processes. The present work is also distinctive in its attribute based approach towards modeling the perceived product quality. This has been discussed in detail in our covariate specification section. Third,
using a series of simulated studies (based on the estimated parameter values from our experiment) we examine if our model can successfully reproduce some of the key distributional and evolutionary characteristics of online rating environment. In these simulations we use intrinsic product quality as a driving element and demonstrate that the innate product quality can drive the evolutionary dynamics of the online rating environment. Moreover, we show that the interrelationships between product quality and various statistical properties (e.g. average and variance) of the rating environments follow complex and non-linear pattern. In many ways the current paper is distinctly different from the earlier papers in this domain. In Table 22 we summarize the key differences in terms of its structure, focal aspects and modeling approach:

<table>
<thead>
<tr>
<th>Decision Stages Considered</th>
<th>Data Source</th>
<th>Latent Quality</th>
<th>System Credibility</th>
<th>Aggregate Measures vs. Product Quality</th>
<th>Impact of Choice Decision on Rating Behavior</th>
<th>Role of Disconfirmation</th>
<th>Decision Stage Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current Paper (2014)</td>
<td>Controlled Experiment</td>
<td>Product Attribute Based</td>
<td>Not Modeled</td>
<td>Examined Through Simulation</td>
<td>Modeled</td>
<td>Not</td>
<td>Error</td>
</tr>
<tr>
<td>Ho et al. (2013)</td>
<td>Secondary Source</td>
<td>Long Term Average Rating</td>
<td>Modeled</td>
<td>Modeled</td>
<td>Modeled</td>
<td>Not</td>
<td>Error</td>
</tr>
</tbody>
</table>

Table 22: Comparison of the modeling approach of the current research with earlier works
In the current work we have reported number of findings that are new and unique. We have also found corroborating evidence for some of the earlier findings. These findings have been presented in a concise and comparative manner in table 23.

<table>
<thead>
<tr>
<th>Impact On Choice /Sales</th>
<th>Valence</th>
<th>Volume</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Dellarocas et al. 2005</td>
<td>Clemons, Gao, &amp; Hitt, 2006; Moe and Schweidel, 2012</td>
<td>Monic Sun, 2011 (only when valence is low)</td>
</tr>
<tr>
<td>Neutral</td>
<td>Duan, Gu, &amp; Whinston, 2008; Liu, 2006</td>
<td>Chintagunta et al. 2010; Luca, 2011</td>
<td><strong>-------</strong></td>
</tr>
<tr>
<td>Negative</td>
<td>Chevalier &amp; Mayzlin, 2006</td>
<td><strong>-------</strong></td>
<td>Monic Sun, 2011 (only when valence is high)</td>
</tr>
</tbody>
</table>

Table 23: Comparison of findings of the current research with earlier works

I recognize some of the limitations in my research. First, the data is collected from a rather homogeneous sample with a limited size. This limits the generalizability of the results. Future researchers can look into this issue by using a more representative dataset that can track the pre-consumption behavior of individual consumers. The subscription based online services can be an ideal platform for collecting such dataset. Second, in my experimental setup I did not use any barrier to consumption (i.e. price). The experiment was time limited and hence there was a penalty for making an inferior or superfluous choice but still there is a possibility that the number of positive choice decisions might be slightly inflated. Third, I did not consider the prior experiences that the respondents may have as a possible source of information that can affect the pre consumption perception. This however, can be achieved easily by incorporating an appropriate set of covariates.
into the linear equations. Fourth, I treated the opinion contribution process only as a post-consumption process. This prevents my model from explaining the mechanics behind the opinions that are expressed by individuals who have no direct experience with the product. Individuals usually express their opinion about products that they have not consumed by comparing the characteristics of the new products with the products that they have already consumed before. Therefore, opinion contributed by individual on the products that individuals have not consumed can possibly be explained by looking at the extent of similarity between the focal product and the product that these individuals have experience with. Fifth, my model does not incorporate the impact of competing products on the rating environment even though I recognize that in reality every product is judged by consumers in the face of a constant stream of new competitive products. The inclusion of competition can be a significant extension of the present model. Finally, my model estimates are at population level so an extension of the model to estimate the parameters at individual level can be an interesting extension.

With technological advancement and proliferation of product related information, internet will continue to play an increasingly important role in our day to day consumption decisions. It will keep growing in terms of both reach and usage. Consequently, we shall see more and more individual level participation in the online opinion platforms. It is therefore crucial for marketers to identify the various factors that influence these processes and understand the individual level decision mechanism that underlines the online opinion creation and consumption process. The current paper is an effort in this direction.
REFERENCES


Ho, Yi-Chun “Chad”, Yong Tan, and Junjie Wu (2013), "Effect of Disconfirmation on Online Rating Behavior: A Dynamic Analysis."


APPENDIX

The estimation process can be described in terms of following iterative steps:

1. Given $y_{i1}^*, y_{i2}^*$ and other parameters, draw $y_{i1}'$ from the corresponding full-conditional normal distribution truncated below 0 for $y_{i1} = 0$ and truncated above 0 for $y_{i1} = 1$

2. Given $y_{i1}^*, y_{i3}^*$ and other parameters, draw $y_{i2}'$ from the corresponding full-conditional normal distribution truncated below 0 for $y_{i1} = 1$ and $y_{i2} = 0$ and truncated above 0 for $y_{i1} = 1$ and $y_{i2} = 1$; moreover, draw $y_{i2}'$ from the full-conditional non-truncated normal distribution when $y_{i1} = 0$

3. Given $y_{i1}^*, y_{i2}^*$ and other parameters, draw $y_{i3}'$ from the corresponding full-conditional normal distribution truncated between upper and lower cutoff points when $y_{i3} = k(>0)$; otherwise, draw $y_{i3}'$ from the full-conditional non-truncated normal distribution when $y_{i3} = 0$

4. Update $\beta_1, \beta_2$ and $\beta_3$ from the respective full conditional distributions

5. Update $\rho_{12}, \rho_{13}$ and $\rho_{23}$ simultaneously using a RWMH step with proposal density confined between (-1, 1)

6. Update $\gamma$ using a RWMH step from a set of proposal cut points corresponding to the ordinal value of observation $y_{3i}$

I proceed to implement estimation algorithm using a data sampler that is designed around a parameter vector that is augmented with three latent continuous variables $y_{i1}^*, y_{i2}^*$ and $y_{i3}^*$. The conditional distribution of each latent continuous variable is a
truncated or un-truncated normal depending on the observed individual level choice, incidence and evaluation data. The sampling strategy for the conditional mean and variance of the partitioned matrix is based on Poirier (1995). Generally, from the theory of multivariate normal distribution I know that if $Y_1, Y_2, \ldots, Y_p$ are multivariate Gaussian, then conditional on a subset $Y_1, Y_2, \ldots, Y_q$, the remaining variables $Y_{q+1}, Y_{q+2}, \ldots, Y_p$ follow a Gaussian distribution as well. If $A$ denotes the conditioning variables and $B$ denotes the conditioned variables and I define the mean vector as $\mu = (\mu_A, \mu_B)$ and the correlation matrix as

$$
\Sigma = \begin{bmatrix}
\Sigma_{AA} & \Sigma_{AB} \\
\Sigma_{BA} & \Sigma_{BB}
\end{bmatrix}
$$

then

$$
\tilde{Y}_B | \bar{Y}_A \sim MVN(\mu_B + \Sigma_{BA} \Sigma_{AA}^{-1} (Y_A - \bar{Y}_A), \Sigma_{BB} - \Sigma_{BA} \Sigma_{AA}^{-1} \Sigma_{AB})
$$

In the present context the continuous latent variables are correlated and the correlation matrix in this case is defined in equation (1.3) as follows:

(W2)

$$
\Sigma = \begin{bmatrix}
1 & \rho_{12} & \rho_{13} \\
\rho_{12} & 1 & \rho_{23} \\
\rho_{13} & \rho_{23} & 1
\end{bmatrix}
$$

Consequently, each of the latent continuous variables conditioned on the other two should also follow a multivariate Gaussian distribution as indicated in equation (1.8). However, in order to account for the sample selection process, these multivariate distributions must be truncated at appropriate points. Therefore, the posterior inference process can proceed via data augmentation by first introducing the latent variables $y_{i1}^*, y_{i2}^*$ and $y_{i3}^*$ for all $i$ in a sequential manner as follows:

1. Generate $y_{i1}^*$ from the conditional posterior given as:

82
(W3) \[ y_{i1}^* \mid \beta, y_{i1}, y_{i2}^*, y_{i3}^* \sim \begin{cases} \text{TN}(\mu_{y_{i1}}, \sigma_{y_{i1}}^2)_{(0, +\infty)} & \text{if } y_{i1} = 1 \\ \text{TN}(\mu_{y_{i1}}, \sigma_{y_{i1}}^2)_{(-\infty, 0)} & \text{if } y_{i1} = 0 \end{cases} \]

Where,

(W4) \[ \mu_{y_{i1}} = \beta_1 X_{i1} + \left( \rho_{12} \rho_{13} \right)^{\top} \begin{pmatrix} 1 & \rho_{23} \\ \rho_{23} & 1 \end{pmatrix} \begin{pmatrix} y_{i2}^* - \beta_2 X_{i2} \\ y_{i3}^* - \beta_3 X_{i3} \end{pmatrix} \]

(W5) \[ \sigma_{y_{i1}}^2 = 1 - \left( \rho_{12} \rho_{13} \right)^{\top} \begin{pmatrix} 1 & \rho_{23} \\ \rho_{23} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_{12} \\ \rho_{13} \end{pmatrix} \]

Generate \( y_{i2}^* \) from the conditional posterior given as:

(W6) \[ y_{i2}^* \mid \beta, y_{i1}, y_{i1}^*, y_{i3}^* \sim \begin{cases} \text{TN}(\mu_{y_{i2}}, \sigma_{y_{i2}}^2)_{(0, +\infty)} & \text{if } y_{i1} = 1 \text{ and } y_{i2} = 1 \\ \text{TN}(\mu_{y_{i2}}, \sigma_{y_{i2}}^2)_{(-\infty, 0)} & \text{if } y_{i1} = 1 \text{ and } y_{i2} = 0 \\ \text{N}(\mu_{y_{i2}}, \sigma_{y_{i2}}^2) & \text{if } y_{i1} = 0 \end{cases} \]

Where,

(W7) \[ \mu_{y_{i2}} = \beta_2 X_{i2} + \left( \rho_{12} \rho_{13} \right)^{\top} \begin{pmatrix} 1 & \rho_{23} \\ \rho_{23} & 1 \end{pmatrix} \begin{pmatrix} y_{i1}^* - \beta_1 X_{i1} \\ y_{i3}^* - \beta_3 X_{i3} \end{pmatrix} \]

(W8) \[ \sigma_{y_{i2}}^2 = 1 - \left( \rho_{12} \rho_{13} \right)^{\top} \begin{pmatrix} 1 & \rho_{23} \\ \rho_{23} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_{12} \\ \rho_{13} \end{pmatrix} \]

Generate \( y_{i3}^* \) from the conditional posterior given as:

(W9) \[ y_{i3}^* \mid \beta, y_{i1}, y_{i1}^*, y_{i2}^* \sim \begin{cases} \text{TN}(\mu_{y_{i3}}, \sigma_{y_{i3}}^2)_{(y_{i1}, y_{i3})} & \text{if } y_{i3} = k \\ \text{TN}(\mu_{y_{i3}}, \sigma_{y_{i3}}^2) & \text{if } y_{i3} = 0 \end{cases} \]

Where,

(W10) \[ \mu_{y_{i3}} = \beta_3 X_{i3} + \left( \rho_{13} \rho_{12} \right)^{\top} \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix}^{-1} \begin{pmatrix} y_{i1}^* - \beta_1 X_{i1} \\ y_{i2}^* - \beta_2 X_{i2} \end{pmatrix} \]
\[ \sigma^2_{\lambda} = 1 - \begin{pmatrix} \rho_{13} \\ \rho_{23} \end{pmatrix} \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_{13} \\ \rho_{23} \end{pmatrix} \]

2. Generate \( \beta \) from \( N(M, V) \), where

\[ M = V^{\frac{1}{2}} \begin{bmatrix} V^{-1}M_0 + \Sigma_i^N X_i^T \Sigma_i^{-1} y_i^* \end{bmatrix} \]

\[ V = \left[ V_0^{-1} + \sum_i^N X_i \Sigma_i^{-1} X_i^T \right]^{-1} \]

Where, \( V_0^{-1} = 0.01 \cdot I; M = 0; \) and \( I \) represents an identity matrix of dimension \( k \times k \) where \( k \) represents the total number of covariates across all three decision stages.

3. Generate \( \rho_{12}, \rho_{13}, \) and \( \rho_{23}: \) The conditional posterior of \( (\rho_{12}, \rho_{13}, \rho_{23}) \) is given by

\[ p(\rho_{12}, \rho_{13}, \rho_{23} \mid y^*, \beta) \propto |\Sigma|^{-N/2} \exp\left( -\frac{1}{2} \sum_{i=1}^N (y_i^* - X_i \beta) \Sigma^{-1} (y_i^* - X_i \beta) \right) \]

I use a Metropolis-Hastings algorithm to sample \( \rho_{12}, \rho_{13}, \) and \( \rho_{23} \) simultaneously in the following fashion:

i. Generate the errors terms from the realized \( \beta \) values and the latent data

ii. Compute the pairwise sample correlations \( (\hat{\rho}_{12}, \hat{\rho}_{13}, \) and \( \hat{\rho}_{23}) \) between errors across equations described in 1.1. Let the sample correlation matrix across the error terms be denoted by \( \hat{\Sigma} \)

iii. Generate a set of three candidate error correlations \( (\hat{\sigma}_{12}, \hat{\sigma}_{13}, \) and \( \hat{\sigma}_{23}) \) from truncated normal distribution \( TN_{[-1,1]}(\hat{\rho}_{12}, \sigma_{12}^2), TN_{[-1,1]}(\hat{\rho}_{13}, \sigma_{13}^2) \) and \( TN_{[-1,1]}(\hat{\rho}_{23}, \sigma_{23}^2) \) respectively using the computed sample correlations as mean and variance \( \sigma_{ij}^2 \) where \( \sigma_{ij}^2 = \frac{\hat{\Sigma}_{ij}^2}{N} \)
iv. Calculate the acceptance ratio $R$ as follows:

\[
R = \frac{\prod_{i=1}^{N} \varphi(y_i^* \mid \beta, \Sigma)}{\prod_{i=1}^{N} \varphi(y_i \mid \beta, \Sigma)}
\]

(W15)

Where, $\pi(\cdot) = N(\sigma_0, b_0^{-1})$ where $\sigma_0 = 0$, $b_0 = 10$; $I(\cdot)$ represents an indicator function and $\varphi(\cdot)$ is defined as follows:

\[
\varphi(y_i^* \mid \beta, \Sigma) = \Sigma^{-1/2} \exp\left\{-\frac{1}{2}(y_i^* - \mu_i)'\Sigma^{-1}(y_i^* - \mu_i)\right\}
\]

(W16)

Where, $y_i^* = (y_{i1}^*, y_{i2}^*, y_{i3}^*)'$ and $\mu_i = (\mu_{i1}, \mu_{i2}, \mu_{i3})'$

Different elements of $\mu_{ij}$ in equation are defined as follows,

(W17) \hspace{1cm} \mu_{i1} = \beta_{1} X_{i1} \\
(W18) \hspace{1cm} \mu_{i2} = \beta_{2} X_{i2} \\
(W19) \hspace{1cm} \mu_{i3} = \beta_{3} X_{i3}

v. Accept or reject proposal value:

Once computed the full ratio $R$, I compare this ratio to a $u \sim U(0,1)$ random draw, and accept $\tilde{\sigma}_j$ if $R > u$ otherwise reject $\tilde{\sigma}_{12}$.

4. Generate $\gamma$ using a random walk Metropolis–Hastings algorithm:

i. Draw a set of proposal outpoints: Sample $\tilde{\gamma}_k \sim N(\gamma_k, \sigma_{\text{MH}}^2, \tilde{\gamma}_{k-1}, \gamma_{k+1})$ for $k = 3...K$. Where, $N(a,b,c,d)$ is a truncated normal distribution with mean a, variance b, and lower and upper truncation points of c and d, respectively.

ii. Compute the acceptance ratio $r$:

\[
r = \frac{f_3(\tilde{\gamma} \mid 3, X_3) f_2(\gamma \mid \tilde{\gamma})}{f_3(\gamma \mid 3, X_3) f_2(\tilde{\gamma} \mid \gamma)}
\]

(W20)
Where $f_a(b)$ is the posterior density evaluated at $\gamma$ (or $\tilde{\gamma}$) and $f_c(d)$ is the value of the proposal density at $c$ when the proposal is centered over $d$. The $r$ value can be expressed as:

$$r = \prod_{i=1}^{N} \frac{\Phi_1(\hat{\lambda}_i, y_{y_{y_i}}, y_{y_{y_i+1}})}{\Phi_1(\hat{\lambda}_i, y_{y_{y_i}}, y_{y_{y_i+1}})} \times \prod_{k=1}^{K} \frac{\Phi_2(\gamma_k, \sigma_{MH}, \bar{y}_k, \gamma_{k+1})}{\Phi_2(\gamma_k, \sigma_{MH}, \bar{y}_k, \gamma_{k+1})}$$

(W21)

Where, $\Phi_1(a,b,c)$ is the integral of the normal density function with mean $a$ and variance 1 between thresholds $b$ and $c$; On the other hand, $\Phi_2(a,b,c,d)$ is the integral of the normal distribution with mean $a$ and standard deviation $b$ between the thresholds $c$ and $d$. The standard deviation $b$ is chosen to produce an acceptance rate around 50%. Moreover, $\tilde{y}_{y_i}$ denotes the cut point proposal corresponding to the ordinal value of observation $y_{y_i}$, similarly $y_{y_i}$ is the actual value of the cut point corresponding to the observed ordinal category. The term $\hat{\lambda}$ denotes the value of the linear predictor in the measurement model for individual $i$ resulting in

(W22)

$$\hat{\lambda} = \beta_{j} X_{i_3}$$

Here, $X_{i_3}$ is the stacked vector of respective covariates for decision stage 3 across all individuals and products and $\beta_{j}$ is the corresponding fixed-effect regression parameter including intercept term.

iii. Accept or reject proposal value:

The product of these components constitutes the full ratio $R$. Once computed, I compare this ratio to a $u \sim U(0,1)$ random draw, and accept $\tilde{\gamma}$ if $R > u$. 

86