INFERRING INDIVIDUAL PREFERENCES AND LATENT BEHAVIORAL FACTORS WITH INCOMPLETE INFORMATION

Liangbin Yang

A DISSERTATION

in

Marketing

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

2017

Supervisor of Dissertation

____

Eric T. Bradlow
K.P. Chao Professor of Marketing,
Statistics, Education and Economics

Co-Supervisor of Dissertation

____

Peter S. Fader
Frances and Pei-Yuan Chia Professor
of Marketing

Graduate Group Chairperson

____

Catherine Schrand, Celia Z. Moh Professor, Professor of Accounting

Dissertation Committee:

David R. Bell, Xinmei Zhang and Yongge Dai Professor of Marketing
Robert J. Meyer, Frederick H. Ecker/MetLife Insurance Professor of Marketing

INFERRING INDIVIDUAL PREFERENCES AND LATENT BEHAVIORAL FACTORS WITH INCOMPLETE INFORMATION

ACKNOWLEDGMENTS

First of all, I would like to express the deepest appreciation to my advisors Profs. Eric Bradlow and Pete Fader for being excellent mentors during my Ph.D. journey. I feel very fortunate to be blessed with my advisors who are extremely considerate, supportive and above all, accessible when I needed their help. They have provided me extensive personal and professional guidance and supports, as well as taught me a great deal about both scientific research and life in general. I honestly could not have hoped for better advisors and role models for my career and life. I also thank them for always believing in me and encouraging me to pursue my goals. Without their guidance, encouragement, and help this dissertation would not have been possible. Words are inadequate to express my heartfelt gratitude to them.

I would like to thank my committee members, Profs. Bob Meyer and David Bell. Bob and David have provided me not only valuable and insightful feedback and comments but also constant encouragement and support. I have benefited tremendously by having them as my committee members. I also would like to express my sincere gratitude to Prof. Jagmohan Raju who have helped me get through the challenging time during my Ph.D. journey. I am also extremely grateful to Profs. Jehoshua Eliashberg and Raghuram Iyengar who I closely work with. They have provided me valuable guidance, scholarly inputs and consistent encouragement throughout my Ph.D. journey. I would like to thank Profs. Christophe Van den Bulte, John Zhang, Barbara Kahn, Qiaowei Shen, Elea Feit, David Reibstein, Wes Hutchinson, Deb Small, Patti Williams, Barbara Mellers, Rom Schrift, Pinar Yildirim, Ron Berman, Maria Ana Vitorino, for many insightful discussions and helpful suggestions. I would like to thank all my teachers during my Ph.D. journey at Penn. The training and the vision they provided are undoubtedly valuable for my entire career.

I would like to thank the Baker Retailing Center and the Wharton Customer Analytics Initiative for generous financial and data support that was instrumental in the completion of this dissertation. I am so grateful for Wharton Research Computing team. The supports Hugh MacMullan and Gavin Burris provided are above and far beyond their jobs. This dissertation and my job market process would be more difficult without their kind help. Many thanks go to Denise Dahlhoff, Anthony Kutovoy, Angela Di Santo, Liz Page, Karen B. Ressler and Shawn Zamechek, SaraRose Christodoulou, for all the administrative kind help and excellent support. I also thank my dear fellow students and friends for their help and company. Many thanks go out to Eric Schwartz, Arun Gopalakrishnan, Jeff Cai, Alix Barasch, Kathy Li, Joy Lu, Ludo Stourm, Valeria Stourm, Jae Young Lee, Amit Bhattacharjee, Cindy Chan, Jonathan Berman, and Evan Weingarten.

Finally, I would like to thank my family for their constant encouragement and unconditional love.

ABSTRACT

INFERRING INDIVIDUAL PREFERENCES AND LATENT BEHAVIORAL FACTORS WITH INCOMPLETE INFORMATION

Liangbin Yang

Eric T. Bradlow

Peter S. Fader

This dissertation extends prior research on inferring individual preferences from the following two aspects: one is to examine important latent behavioral factors affecting consumers' consumption decisions; the other is to overcome the challenges arising from incomplete information. Regarding latent behavioral factors, this dissertation considers the following two aspects: (1) three types of intragroup dynamics behavior, and (2) variety-seeking behavior. Regarding incomplete information, this dissertation focuses on two types of incomplete information: individual's behavior and identity, and order of consumption. Specifically, Chapter 2 presents a method to infer heterogeneous individual preferences and three components of intragroup dynamics using just aggregate and de-identified data. Chapter 3 emphasizes the effect of consumption outcomes on an individual's propensity for variety-seeking when the order of consumption is unobserved. To overcome the challenges arising from incomplete information, this dissertation develops advanced individual-level Bayesian models and uses two-step iterative algorithms to estimate the proposed models in an MCMC framework. In-depth simulation studies show that the parameters are well recovered, suggesting that the proposed models are identified. Furthermore, this dissertation shows that ignoring latent behavioral factors may lead to biased estimation of individual preferences, which could result in many consequences. This dissertation applies the proposed methods to two empirical settings: an individual-level TV viewing and targeted TV advertising setting using Nielsen People Meter (NPM) data, and an online video game environment. In the TV viewing setting, it is shown that the proposed method could sig-

nificantly improve the efficiency of TV ad targeting through counterfactual analysis. In the video-game environment, results show that although there is extensive heterogeneity, on average, positive consumption outcomes lead to inertial preferences, while negative consumption outcomes lead to variety-seeking. In sum, this dissertation shows the importance to incorporate important latent behavioral factors in inferring heterogeneous individual preferences especially when data are incomplete, and proposes innovative methods to overcome the challenges emerging from incomplete information.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Today, rapid advancement of technology enables marketers to reach individual customers with customized marketing strategies in many ways such as online ads, emails and so on. It is widely accepted that a customized ad for individual customers can be much more effective than a universal "one-size-fits-all" one for all customers. However, the advantage of individually customized marketing strategies does not "come for free." Firms need to correctly infer individual's preferences before they send out effective ads or recommendations to individual customers. Correspondingly, a great body of research in marketing has been done to reveal a customer's preferences based on his/her historical consumption behavior (Allenby & Rossi 1998). For example, based on a customer's historical purchase data, choice models infer the customer's preferences on each of several brands in a product category. A retailer then may be able to send the customer customized ads based on his/her preferences.

However, most existing models that examine individual preferences come with an implicit assumption that customers make their decisions individually, without any interaction with other customers or environment when they make decisions. Such assumptions are unlikely to hold in some cases and may cause bias in the estimation of individual preferences. For example, individuals in a household may make consumption decisions together instead of alone. In this case, it is important to take into account the interaction among household members. Therefore, it is more challenging than it looks to correctly infer customers' individual preferences in joint decision-making environments.

In this dissertation, we intend to overcome some challenges associated with inferring individual preferences and latent behavioral factors in marketing practice; more importantly, we develop novel methods and models, as well as innovative estimation algorithms, and apply them to empirical marketing contexts. We hope that this dissertation will help marketers overcome the challenges emerging from incomplete information and avoid possible biases in inferring individual preferences using existing methods, and shed some light on how to incorporate important latent behavioral factors which may significantly affect customers' consumption decisions.

Specifically, we look into the following two challenges in this dissertation. One is to take into account the important latent behavioral factors that affect customers' consumption decisions (in addition to the product or brand itself). In this dissertation, we focus on factors coming from two sources: one is the possible social influence and agent-to-agent interactions; the other is possible interactions with past consumption behavior of the customer himself/herself. Specifically, we focus on the cases that customers may make decisions in a group setting (for example, household members make a decision together about where to travel for their vacation). Over the last few decades, many researchers define a group from various perspectives (e.g., interdependence, shared identification, shared goals, shared tasks, structure, categorization, and so on). For instance, Shaw (1981) defines a group as "two or more persons who are interacting with one another in such a manner that each person influences and is influenced by each other person" (Shaw 1981, p454). To better understand individuals' preferences, we will have to learn about the complex social actions, processes and changes that happen within a group, i.e., within-group or intragroup dynamics (Lewin 1951). Thus, we would like to take into account possible intragroup dynamics when we model individuals' preferences.

Second, we take into account possible effects from individual's prior consumption experience, especially for experiential products such as playing video games or watching movies. In particular, if a customer consumes a brand or an option intensively, he/she may become satiated to the brand/option and exhibit variety-seeking behavior by switching to different brands/options, or may become "locked-in" by the brand/option and exhibit inertial behavior by consistently choosing the same brand/option. In this case, the customer's prior consumption affects his/her choices in the future, in addition to his/her intrinsic preference of each brand/option. Furthermore, such variety-seeking or inertia behavior may be affected by the valence of the experience gained in the consumption (especially for experiential products). For example, a video game player may be more likely to continue to play if he/she has a good experience in a previous round than a bad one. Thus, it is important to account for the consumption experience of outcomes and related variety-seeking/inertia behavior when we infer individuals' preferences.

Another challenge in inferring individuals' preferences is that sometimes data are not "complete". Many reasons can cause incomplete data, such as technical issues (e.g., it is difficult to track customers' offline behavior), or privacy regulation, and so on. In this dissertation, we investigate two possible sources of incomplete data. One is that we may have missing information about individuals' behavior and identity, so only aggregate and de-identified data are available. This refers to the cases that we observe the total consumption for a group of customers but we do not know how many and which of them consume which product. For instance, we observe members in a household (that consists of a husband and a wife) watch sports three times in a week, but we do not know who watches each time (either the husband, or the wife, or both of them). This is a very general problem in marketing. As long as multiple consumers may share the same account or device, we may have the same issue. Such kind of missing information brings challenges to estimating individual preferences since the data available are at the aggregate level, and we would have to first infer individual-level consumption behavior before we estimate the individual preferences.

Another source of incomplete data we examine is the missing information about the order of consumption (or the timing of the consumption). This refers to the case that we observe the consumptions in a period, but we do not know the consumption order: which item was consumed first, etc. For example, we observe that a video game player plays three rounds of a video game in a day, but we do not know which round comes first, and which one comes last. Or a Netflix user rents several DVDs together, but we do not observe in which order this user watches the DVDs. Or a customer purchases multiple items in one shopping occasion, and his or her consumption order is not recorded or trackable. This type of missing information brings challenges to examine the consumption sequences across time, since we would first have to infer the order of consumption taking place to investigate the relationship among multiple consumption events across time.

Inferring behavioral factors has been investigated by many researchers using different types of data. The majority of previous studies have extensively focused on understanding observed or unobserved behavioral factors based on complete information (i.e., the important information is not missing in the data). Relatively few studies examine how to infer observed or unobserved behavioral factors when information is incomplete. Although with technology advances, the data we collect may be much richer than before, unfortunately, it is often that we are "data rich but information poor." It is not rare that critical information is missing even with big data. It would seem, therefore, that further investigation on how to infer latent behavioral factors with incomplete information are needed. Table 1.1 summarizes that this dissertation bridges the task of better understanding important latent behavioral factors with overcoming the essential challenges in information availability, aiming to fill this important gap in the literature.

Table 1.1: Inferring Behavioral Factors with Different Types of Data

		Information	
		Complete	Incomplete
Behavioral Factors	Observable		
	Unobservable		This dissertation

In this dissertation, demonstrating ways to overcome the two challenges discussed above, we present two essays in Chapter 2 and Chapter 3 on inferring individual preferences and latent behavioral factors with incomplete information, as shown in Table 1.2.

Table 1.2: Comparison of Chapter 2 and Chapter 3

		Incomplete Information	
		Individual's Behavior and Identity	Order of Consumption
Unobserved Behavioral Factors	Intragroup Dynamics	Chapter 2	
	Variety-Seeking/Inertia		Chapter 3

In Chapter 2, "Inferring Individual Preferences and Intragroup Dynamics with Aggregate and De-identified Data: An Application to Targeted TV Advertising," we present a method that can be used to model and infer customers' consumption, preferences and intragroup dynamics at the individual level when there is a lack of rich individual-level data. The model aims to enable customized marketing strategies (e.g., personalization) in areas where individual behavior and consumption is difficult to track and observe, such as offline shopping and TV ad targeting. Specifically, this chapter proposes a joint consumption theory taking into account intragroup dynamics, state dependence, time-varying factors, and observed/unobserved heterogeneity, to estimate the model with just aggregate and de-identified data. The proposed method can disentangle three potentially confounded components of intragroup dynamics: 1) preference revision, i.e. when an individual's preferences depend on the preferences of others, 2) behavioral interaction, i.e. when an individual's consumption utility depends on the choices of other group members, and 3) decision power, i.e. the influence an individual exerts when his or her group makes a decision. We propose a new identification strategy which relies on two important identification sources and an innovative estimation algorithm, followed by a series of in-depth simulation studies where we validate the proposed method, identification strategy, and estimation algorithm. Finally, we conduct a series of calibrated counterfactual simulations demonstrating that our proposed model will enable advertisers to significantly improve the efficiency of targeting intragroup individuals.

In Chapter 3, "Inferring Individual Preferences and Variety Seeking with Non-Ordered Data: An Application to Video Games," we propose a method that examines the variety-seeking behavior at the individual and attribute level, and estimate the proposed model using incomplete data where the order of consumption occasions is missing. At the heart of the model, we propose that consumers may become variety-seeking in response to the valence of consumption outcomes which indicate the quality of the consumer's experience for a specific consumption occasion. We predict that positive consumption outcomes will generally lead to inertial preferences, while negative consumption outcomes lead to variety-seeking. Accordingly, in our model, we allow variety-seeking to be a function of consumption outcomes. Furthermore, we allow the variety-seeking to vary across attributes of the options, which provides managerial insights to product design. We test our hypotheses within a context of an online video game, in which players choose between different map environments for each round of play, and consumption outcomes can be measured objectively by a player's performance during the round. We observe consumption outcomes that are consistent with our hypotheses, suggesting that in our context, in general, firms should place players in a familiar environment in the next round of play if they are performing well, and introduce variety if they are performing poorly. Since heterogeneity in variety-seeking across players is observed, this suggests the importance to incorporate customized strategies to individual players.

In sum, in Chapter 2 and Chapter 3, we propose novel methods to overcome challenges arising from incomplete information and incorporate important latent behavioral factors mentioned above to help marketers better understand heterogeneous individual preferences and design corresponding marketing strategies more efficiently. Specifically, regarding modeling important latent behavioral factors that affect customers' consumption decisions, Chapter 2 takes into account possible intragroup influence and interactions, while Chapter 3 considers variety seeking and inertial behavior, and other possible interactions and correlations with past consumption behavior of each customer himself/herself. In terms of overcoming the challenge of incomplete information, Chapter 2 tackles the issue of missing information on individual's behavior and identity while Chapter 3 addresses the problem of missing information on the order of the consumption. In both chapters, we use in-depth simulation studies to validate the proposed methods and estimation algorithms, and apply the proposed model to empirical contexts. This dissertation shows the importance of incorporating latent behavioral factors in inferring individual preferences especially when the data are incomplete, and proposes innovative methods to overcome the challenges emerging from incomplete information.

Chapter 2

Inferring Individual Preferences and Intragroup Dynamics with Aggregate and De-identified Data: An Application to Targeted TV Advertising

2.1 Introduction

With rapid and tremendous technological advancements and breakthroughs, there has been a growing interest in personalization. Research conducted by EPiServer indicated that there were already about 33% of U.S. marketers in 2011 who believed that personalized campaigns were "highly effective and measurable"[1]. In September 2016, a marketing research company, Ascend2, found that 50% of marketers believed that sending individualized messages was the most efficient strategy for various marketing campaigns[2]. In 2017, according

[1] "Personalized marketing brings rewards and challenges -difficulties in managing data hinder more personalized efforts." eMarketer, June 2, 2011, June 3, 2017 accessed.

[2] "Email marketing strategies: survey summary report." Ascend2, September 2016, June 3, 2017 accessed.

to a recent study conducted by Evergage, 88% of marketers in their study stated that efforts in personalization generated high payoffs and provided substantial and measurable improvements[3].

However, the majority of U.S. marketers haven't implemented personalization[4]. It is surprising that even for email which is the most common and successful application for personalization, there are still about 61% of marketers who never reach their customers through personalized emails[5]. If most marketers believe in the effectiveness of individualized messages and products, why hasn't the industry widely adopted personalization? One possible explanation is due to several barriers that exist towards implementing personalization. The biggest challenges are the inability of gaining insights from data (40%) and unavailability of individual-level data (39%)[6]. It goes without saying that the ability to send customized messages to targeted individuals is primarily based on data availability, as well as data analysis methods and techniques. Among the challenges, it seems that data availability is a basic premise of implementing personalization. Without rich information about individuals, it is difficult for marketers to predict an individual customer's behavior, needs and wants. As a result, online and digital media have attracted the most attention in personalization so far, where it is relatively easy to track and observe individual's information.

Nonetheless, it is worth noting that tailoring messages and products for offline customers is equally important[7]. Studies showed that 95% of marketers who adopted personalization in their offline channels had noticed a measurable improvement in conversion rates, outranking any online and digital channels studied[8]. Unfortunately, in most offline cases, as well as

[3]"2017 trends in personalization survey report." Evergage, 2017, June 3, 2017 accessed.

[4]"Personalized marketing brings rewards and challenges -difficulties in managing data hinder more personalized efforts." eMarketer, June 2, 2011, June 3, 2017 accessed.

[5]"Personalization research: how retailers personalize across five channels." Certona, 2017, June 3, 2017 accessed.

[6]"What are personalization's biggest challenges and opportunities?" Marketingcharts, July 24, 2014, June 3, 2017 accessed.

[7]"Offline personalization matters just as much: marketers who personalize offline most likely to see lift in conversions." eMarketer, January 14, 2015, June 3, 2017 accessed.

[8]"Offline personalization matters just as much: marketers who personalize offline most likely to see lift

many online cases, it is often difficult to monitor and collect information at the individual level. Correspondingly, it is very common that only group (e.g., household) data, instead of individual-level data, are available. For example, retail scanner data often only include transaction information for each household but not for individuals in the same household. With retail scanner data, we know when a household purchases what product (e.g., yogurt), but we do not know which individual within the household purchases/consumes the products. This becomes a big roadblock which prevents many retailers from adopting and implementing personalization.

Furthermore, in some cases where marketers can collect individual's information from their customers, the data collected are often "de-identified": the individual-level information observed isn't connected with each person's identity. For instance, for a two-member household which has two TVs, if we observe that two different programs are being watched at the same time in this household, we know that two members are watching different programs but still don't know who is watching which one. De-identified data is a common issue not only for offline but also online environments (Novak et. al, 2015). In many online scenarios, individuals' identity can also be unobservable because of various reasons, such as privacy regulations, technical limitations, or non-technical reasons (e.g., multiple people may share an online account or a computer which makes it difficult to identify who is logging in the online account or using the computer).

It is worth noting that this problem doesn't arise purely because the data collected lacks individual-level information or are de-identified. Today, everyone lives in a social world. Many people prefer making decisions and enjoying daily activities together with others rather than being alone. Such kinds of choices and consumption processes are often complicated, not only driven by intragroup heterogeneous individual preferences, but also by intragroup dynamics (e.g., individuals' preferences and choices are influenced by others within the group, etc.). For example, a wife may watch the Super Bowl with her husband,

in conversions." eMarketer, January 14, 2015, June 3, 2017 accessed.

not because she likes watching football, but just because she enjoys watching TV with her husband. Another obstacle which impedes the advance of personalization is that intragroup dynamics data are normally unavailable. The importance of intragroup dynamics defines this problem to be different from a simple assignment problem (who is watching). That is, we cannot resolve this fundamental problem by simply collecting better individual-level data and directly using independent single-agent consumer models to assign the consumption information to an individual consumer. In many cases, we will need to infer the latent behavioral factors (e.g., intragroup dynamics in this chapter) with missing information on intragroup interaction.

Given that most data we have today are aggregate and de-identified, many marketers had no choices for decades but to use the group (e.g., household) as the unit of analysis, which typically assumes homogeneous preferences within a group and ignores intragroup interactions and influence. In some cases, it may be reasonable to assume homogeneous preferences within a group and use the group as the unit of analysis, or it may be possible that heterogeneous individual preferences and complex intragroup dynamics can cancel out with aggregation (Adamowicz et al. 2005); however, making individual-level inferences without incorporating the role of heterogeneity and intragroup dynamics could lead to biased estimates and erroneous predictions (Yang, Narayan and Assael 2006).

Undoubtedly, modeling heterogeneous individual preferences and incorporating possible intragroup interactions and influence have great importance to infer individuals' preferences and exercise personalization appropriately and efficiently. Take the application in this chapter, household TV viewing and targeted TV advertising, as an example. We demonstrate that when constrained by aggregate and de-identified set-top box (STB) data, without properly inferring individual preferences and intragroup dynamics, targeted TV advertising campaigns will be inefficient. Consider a beauty product company that wants to broadcast its female beauty product ad. With STB data, this company observes only the viewing behavior from each TV, without knowing who is (are) watching and which program (if more than one program is being watched simultaneously) is watched by whom. Airing its ad while a male but not a female is watching can be much less effective. The inability of targeting heterogeneous individuals within the same household becomes a great barrier for multichannel video programming distributors (MVPD) and networks to compete with online/digital advertising platforms. This may be one of the reasons why marketers are gradually moving away from traditional offline media (e.g., TV), and allocating more and more of their ad budgets to online/digital media, which has more advanced ad targeting technology to reach relevant individuals, not just households (Vranica and Perlberg 2015).

Not only does TV advertisers face this pressing issue but also many other marketing contexts. For instance, more than one individual may use the same computer (e.g., public computer) to browse online content, multiple group members may play games together on a video-game console, or household members often share the same store loyalty account (e.g., grocery store loyalty account) or club membership (e.g., Amazon Prime, Costco membership, museum membership). As long as two or more group members share the same device/account, this issue exists. In these cases, due to the aggregate and de-identified nature of information, it is questionable to simply deliver targeted ads/promotions based on the historical behavior observed from the same device/account without recognizing the important roles of unobserved heterogeneous individual preferences and intragroup dynamics.

When two or more group members share the same device or account and only aggregate and de-identified data are available, personalization seems unrealizable. Inability to infer heterogeneous individual's preferences and intragroup dynamics becomes a great barrier for marketers to underpin new waves of marketing innovations. This motivates us to look closely into how to jointly infer heterogeneous individual preferences and intragroup dynamics with commonly available aggregate and de-identified data.

Although considerable research has been devoted to inferring individual behavior using aggregate-level data (e.g., Chen & Yang 2007; Feit et al. 2013; Musalem, Bradlow & Raju 2008, etc.), rather less attention has been paid to jointly infer heterogeneous individual preferences and intragroup dynamics. This chapter is different from previous research that has studied individual behavior using only aggregate-level data in several ways. First, an interesting feature of TV viewing is that household members can share the "viewership". In other words, when two or more household members watch a TV program together, we only observe one view. This is different from previous studies which focus on consumption of regular goods or products. For example, in the orange juice setting examined by Chen and Yang (2007), if each of two individuals consumes one bottle of orange juice, then they observe consumption of two bottles, but not one. That is, we don't observe the variation of quantity, which brings us another level of uncertainty which needs to be resolved: when a view is observed, it not only can be watched by one of household members but also can be watched by any possible combinations of them. Second, prior literature which focuses on inferring individual behavior using aggregate-level data usually assumes that individuals' preferences are independent. We extend prior work by inferring not only intragroup heterogeneous individuals' preferences (e.g., whether the wife in the example above likes watching sports) but also allowing dependent individuals' preferences (e.g., the wife in the example above enjoys watching TV with her husband).

Recently, there has been a growing interest in inferring intragroup dynamics using individual-level data (e.g., Arora and Allenby 1999; Kato and Matsumoto 2009; Yang, Narayan and Assael 2006; Yang et al. 2010; Zhang et al. 2009, etc.). First, each prior study in this stream tends to focus on one component of intragroup dynamics[9]. We investigate three important components of intragroup dynamics simultaneously to help marketers better understand how group members affect and shift each other's preferences persistently, how group members interact with each other across time, and how group members make group decisions with and without conflict. To our knowledge, this chapter is the first to simul-

[9] Kato and Matsumoto (2009) focus on two components of intragroup dynamics.

taneously examine these three intragroup dynamics components. Second, we show in this chapter that marketers can disentangle these three potentially confounded components of intragroup dynamics using just aggregate and de-identified data.

We first propose a theory of joint consumption by developing a novel individual-level joint consumption model and incorporating two major types of group-decision making mechanisms, which takes heterogeneous individual-level preferences, intragroup dynamics (e.g., interactions and influence across individuals) and state dependence (interactions across time) into account. We directly model intragroup dynamics to examine how group members influence and shift each other's preferences in a long-term stable way, how group members interact with each other across time, and how group members exercise their influence when they make group decisions with and without conflict.

Specifically, we look into the following three aspects: 1) preference revision; 2) behavioral interaction and 3) decision power. In the context of household TV viewing, these three aspects refers to the following three examples respectively: 1) if member A in a household likes comedy, member B may be influenced by member A and become a person who likes comedy; 2) if member A is watching TV, member B may be more (less) likely to watch TV since he/she enjoys (hates) the time watching TV with member A; 3) when member A and member B watch TV together, the one with higher decision power may have the authority to decide or largely influence what program the group will watch.

We then further investigate our model and method using real data on TV viewing and targeted TV advertising setting. Specifically, we use Nielsen People Meter (NPM) data to examine TV viewing choices. In particular, we estimate our model through two Bayesian steps iteratively, following and extending previous literature (e.g. Tanner & Wong 1987): first, we use data augmentation to impute missing data about individual choices, and then we estimate the model using the individual-level choice data generated in the data augmentation step. In the imputation step, we provide an algorithm which simultaneously incorporates the dependence of time-correlated missing information and the uncertainty of estimated values to estimate the unobserved state dependence across time. In the modeling step using the augmented individual level data, we implement a Gibbs sampling framework to estimate the model. Our estimation results show the existence of intragroup heterogeneity and intragroup dynamics as well as strong state dependence across periods. We also find that part of the heterogeneity in base preferences and intragroup dynamics can be explained by household-level covariates and individual's demographics including income, age, gender and working hours.

The identification of the model is achieved by leveraging cross-sectional and longitudinal variation, with the existence of single-member groups, changes of available choice sets and time-varying factors. We demonstrate the identification and validity of the proposed model in several ways: a discussion, an empirical validation and a series of simulations. Specifically, we first discuss the intuition of our identification strategy with some simulation studies to show how variation at the aggregate level help us identify the preferences and group dynamics at the individual level. Second, we use simulation studies to create data sets (at the aggregate level) and then estimate the proposed model using the simulated data sets; we are able to recover the simulated parameters and also show that ignoring intragroup heterogeneous individual preferences and/or intragroup dynamics causes biased estimates and yields several consequences for marketing managers. Finally, the proposed model and its identification are also validated by the empirical application to the TV targeting setting using NPM data. In the empirical settings, we show that our proposed model outperforms benchmark models regarding model fit, the efficiency of targeting, and economic outputs.

To summarize, methodologically, we propose a new individual-level model and a novel algorithm to infer individual preferences and intragroup dynamics using just commonly available aggregate and de-identified data. With our method, marketers can better understand who actually consumes the products, and how individuals within a group interact with and influence each other. We propose an identification strategy by leveraging cross-sectional and longitudinal variation, with the existence of single-member groups, changes of available choice sets and time-varying factors to identify heterogeneous individual preferences and different components of intragroup dynamics.

Theoretically, this chapter addresses several important gaps in the literature. First, we develop a joint consumption theory and emphasize three important components of intragroup dynamics (i.e., preference revision, behavioral interaction, and decision power). To date, there has been little systematic investigation that has considered multiple components of intragroup dynamics in the same study. Second, unlike most previous literature on estimating individual-level models using aggregate data which usually looks into the cases at the market level, we provide a model at a more granular level where the influence and interactions among individuals are usually strong, such as intra-household. Third, different from prior studies on intragroup consumption which either focus on private consumption (i.e., consumer consumes products alone) or public consumption (i.e., consumers consume products together), our study extends the prior work and allows products (e.g., TV viewing in our setting) to be consumed both privately (e.g., watch TV alone) and publicly (e.g., watch TV together with others). Moreover, the information we have in our setting is even more incomplete: we don't observe the variation of quantity to infer whether a consumption choice is private or public. Last but not least, our study advances the understanding of intragroup consumption and decision making in a household TV viewing setting.

Managerially, we applied our method to household TV viewing and targeted TV advertising. We show that our proposed method will enable advertisers to better target within-household individuals and significantly improve the effectiveness of ad targeting. The proposed model can potentially be used in many marketing contexts where only group-level data are observed. For example, in the TV viewing context examined in this chapter, marketers usually observe household data only and thus target customers at the household-level. With the proposed method, marketers may now be able to infer preferences and choices of individuals within a household and thus effectively target intragroup individuals accordingly. Retailing is another area where the proposed method can contribute. Retailers usually observe transactions made by households while they do not observe how individuals in a household consume the purchased products. The proposed method can provide important insights for retailers about how to design efficient customized marketing strategies, such as coupon design and promotion strategies.

In sum, this chapter is a supplement to several streams of literature, including the literature on estimating individual-level models with aggregate data, the literature on inferring intragroup dynamics with individual-level data, and the literature on intragroup consumption and decision making. We argue that ignoring intragroup heterogeneous individual preferences and intragroup dynamics results in biased estimates and yields several consequences. We develop a theory of joint consumption which accounts for intragroup dynamics, state dependence, and observed and unobserved heterogeneity, propose a novel identification strategy and an innovative algorithm to overcome the challenges arising from identification and estimation, apply the proposed method to a household TV viewing and targeted TV advertising setting using Nielsen People Meter (NPM) data, and demonstrate that the proposed method significantly improves the efficiency of targeting intragroup individuals.

The remainder of this chapter is as follows. In Section 2.2, we introduce our theory of joint consumption and a model incorporating three potentially confounded components of intragroup dynamics and state-dependence. This is followed by an estimation and identification discussion in Section 2.3. In Section 2.4, we apply our model to TV viewing and TV targeted advertising context using a novel data set obtained from Nielsen, and conduct a series of model comparisons and a counterfactual analysis. Discussion of potential future extensions of the proposed method is provided in Section 2.5. Finally, in Section 2.6, we provide a summary of conclusions, managerial findings, and implications.

2.2 Theory of Joint Consumption

In this section, we introduce our theory of joint consumption by defining latent preference (including two components: base preference and preference revision), individual-level utility for joint consumption, and group decision-making processes respectively. As discussed above, unlike prior studies which either focus on private consumption (i.e., consumer consumes products alone) or public consumption (i.e., consumers consume products together), we extend the prior work and allow products to be consumed both privately and/or publicly.

For the sake of simplicity, in what follows we focus on a group (e.g., a household) h that consists of two members ($i = A, B$; member A and member B). We discuss how to generalize our model to more than two group members in Section 2.5. Our joint consumption theory consists of three parts: 1) latent preference; 2) individual-level utility for joint consumption; 3) group decision-making process.

Specifically, we assume and model the following: first, each household member (member A or member B) has his or her base preference before forming a group (i.e., h) in period t_0. This base preference drives his or her original choices. When household h is formed and two members cohabitate, there is a preference revision process in which household members influence and shift each other's preferences. This preference revision process can be continued until preferences of both members reach steady states such that their preferences do not change any longer. We call these new steady states of preferences as "revised preferences". After household h is formed, the preference revision process completes and is irreversible. From $t = 1$, individual household member's utility for joint consumption in each period depends on his/her revised latent preference, as well as temporal factors including behavioral interaction, state-dependence, observed time-varying factors and unobserved random shocks. Each household member may have different influence (i.e., decision power) for joint consumption. To make a decision, each household member follows a predefined decision-making mechanism[10]to maximize either individual-level or household-level utility.

2.2.1 Latent Preference: Base Preference and Preference Revision

2.2.1.1 Base Preference

In this chapter, we refer to an individual's original preferences before forming a group as base preference (z_{hg}^{A0}, z_{hg}^{B0} for member A and B in household h over a choice g in period t_0). When two members (A and B) form a group and cohabitate, they influence each other's preferences; as a result, their preferences are permanently changed or shifted to their revised preferences (z_{hg}^{A} and z_{hg}^{B}) respectively.

Similar to previous literature (Allenby and Rossi 1998; Yang, Narayan and Assael 2006), we first describe the base preferences of member A and member B in a household h by exogenous covariates (e.g., explanatory variables: age, income, education, etc.). Specifically, we have

$$z_{hg}^{A0} = X_{hg}^{A} \cdot \beta_g + v_{hg}^{zA0} \tag{2.1}$$

$$z_{hg}^{B0} = X_{hg}^{B} \cdot \beta_g + v_{hg}^{zB0} \tag{2.2}$$

where $\left(v_{hg}^{zA0}, v_{hg}^{zB0}\right) \sim MVN\left[0, \Sigma_z\right]$. X_{hg}^{A} (X_{hg}^{B}) is a vector containing an intercept (the term "1") and explanatory variables that could be either specific to member A (B), such as age, or common to both of them, such as household income. $X_{hg}^{A} \cdot \beta_g$ ($X_{hg}^{B} \cdot \beta_g$) captures the observed heterogeneity in base preference of member A (member B); whereas v_{hg}^{zA0} (v_{hg}^{zB0}) captures the unobserved heterogeneity in base preference of member A (B), in which v stands for unobserved heterogeneity. Σ_z describes the unobserved covariation between member A and

[10]Two group decision-making mechanisms have been tested in this chapter.

member B before forming a household. There may be some unobserved factors which are correlated with the decision of group formation and cohibitation for member A and member B. For instance, member A and member B cohibitate because they have several common interests, such as travelling. If this is the case, the elements in Σ_z tend to be positive for household h.

2.2.1.2 Preference Revision

When household h is formed and two members cohabitate, there is a preference revision process in which household members influence and shift each other's preferences. There is a growing body of literature recognizing the notion of interdependent preferences among members (e.g., Aribarg, Arora, and Bodur 2002; Case 1991; Yang and Allenby 2003; Yang, Narayan and Assael 2006). However, researchers model the preference revision in different ways. In particular, Aribarg, Arora, and Bodur (2002; henceforth AAB) models preference revision by assuming that how much an individual affects another one depends on the difference in their base preferences (i.e., $z_{hg}^{B0} - z_{hg}^{A0}$), while Yang, Narayan and Assael (2006; henceforth YNA) describes a revision process where the influence a member (e.g., member A) has on another member (e.g., member B) depends on the other member's (e.g., member B's) revised preference only (i.e., the preference interdependence of A only depends on revised preference of B, z_{hg}^{B}).

However, these previous models of preference revision may not be applicable to our setting. For example, per AAB (2002), the preference revision should depend on the difference between two member's base preferences. AAB is more appropriate for a revision process which takes place once (which is consistent with their setting), but not a revision process taking place again and again. In this chapter, we follow but extend AAB (2002)'s idea in two aspects. First of all, built on AAB (2002)'s idea of preference revision, we assume that when household h is formed and two members cohabitate, there is a preference revision process in which household members influence and shift each other's preferences until preferences of both members reach an equilibirum/steady state. We argue that this equilibirum state of revised preference is more suitable for our setting where household members, who live with each other, have lots of opportunities to influence and shift other members' preferences. More importantly, households in our empirical application may have formed a long time before we start to observe their behavior. The preference revision process should have been completed and reach the equilibirum states.

Specifically, consider a process that member A and B affect each other depending on the difference of their latest preferences. Note that, what AAB (2002) models is the first step of the revision process where the latest preference is base preference. When the household h is formed, household members influence each other again and again until an equilibrium state is reached and the revised preferences are steady. In this case, the revision will depend on the difference of their final revised preferences, i.e., $z_{hg}^{B} - z_{hg}^{A}$.

Second, unlike previous literature (e.g., AAB (2002)) which assumes that the outside option does not change during the preference revision process, we incorporate a shift, δ, on the outside option. By doing so, previous literature nests within our general model (the case which previous literature models is equivalent to a case of δ=0 in our model). For example, when member A and member B marry each other and form a household, they may have more outside options than they were single. They may enjoy cooking together after they formed a household, but cooking together was not a feasible outside option before they cohabitated. Since we normalize the outside option as zero, this is equivalent to having a shift in the preferences. In sum, per the two aspects discussed above, we describe the revised preferences in the following way:

$$z_{hg}^{A} = z_{hg}^{A0} + \delta_h^{A} + \alpha_h^{BA} \cdot \left(z_{hg}^{B} - z_{hg}^{A} \right) \tag{2.3}$$

$$z_{hg}^{B} = z_{hg}^{B0} + \delta_h^{B} + \alpha_h^{AB} \cdot \left(z_{hg}^{A} - z_{hg}^{B} \right). \tag{2.4}$$

where the final states on the left-hand-side of the model equal to the revised states on the right-hand-side. In this way, the preference revision will reach an equilibrium state[11].

In the model above, we assume that preference revision consists of two parts: a shift and influence from other household members. Specifically, for each individual, there is a permanent shift δ_h^A (δ_h^B), which, as mentioned above, takes into account possible changes of outside options. So, this shift is homogeneous across all possible choices. In our setting, for example, a negative shift δ_h^A means that, after forming a group, member A has a lower overall preference on watching TV, or equivalently, he or she becomes a person who likes outside options more. For example, A may prefer having dinner instead of watching TV with other household members.

The total influence from B to A depends on the degree of influence member B has on member A (denoted as α_h^{BA}, and similarly, α_h^{AB}) and the difference of the revised preference between B and A (i.e., $\left(z_{hg}^B - z_{hg}^A\right)$ and $\left(z_{hg}^A - z_{hg}^B\right)$). The rationale behind is that, each member, e.g., A, has two propensities, one is to stick to his or her original preference, which is related to $z_{hg}^A - z_{hg}^{A0}$, the other is to be influenced by the other member B, which is related to $z_{hg}^B - z_{hg}^A$. The revised preference will reach a steady state if the effect from these two propensities equal to each other and reach an equilibrium. α_h^{BA} can be considered as a ratio to balance these two propensities and it indicates how strong A is influenced by B, compared with sticking to its original preferences.

In addition, we allow asymmetric coefficients of preference revision within household (i.e., α_h^{BA} can be different from α_h^{AB}). For example, a large positive α_h^{AB} indicates that member A influences B greatly, and member B's preference become more similar to member A after preference revision. While when α_h^{AB} is close to 0 then member A's latent preference has

[11]Note that this dynamic preference revision process is not modeled in this chapter because for our empirical application, the group formation often took place a long time ago (outside the observation time window) and preference revision has already reached a steady state. Even for the preference revision process which takes place within the observation time window, household members can learn about each other's revised preference almost instantaneously. As a result, we assume the evolution process completes fairly quickly.

a negligible influence on member B's latent preference, leading to a result that member B doesn't revise his or her latent preference after the group is formed. Note that, we allow α_h^{AB} and α_h^{BA} to be less than 0 to capture the possibility of negative influence.

In summary, for a two-member household, member A's and member B's latent preferences are represented as follows including base preference and preference revision (substituting equations (2.1) and (2.2) to (2.3) and (2.4)):

$$
\begin{aligned}
z_{hg}^{A} &= \delta_h^A + \left(X_{hg}^{A} \cdot \beta_g + v_{hg}^{zA0}\right) + \alpha_h^{BA} \cdot \left(z_{hg}^{B} - z_{hg}^{A}\right) && (2.5)\\
z_{hg}^{B} &= \delta_h^B + \left(X_{hg}^{B} \cdot \beta_g + v_{hg}^{zB0}\right) + \alpha_h^{AB} \cdot \left(z_{hg}^{A} - z_{hg}^{B}\right). && (2.6)
\end{aligned}
$$

where $\left(v_{hg}^{zA0}, v_{hg}^{zB0}\right) \sim MVN\left[0, \Sigma_z\right]$.

In this model, we allow individual-specific preference revision coefficients and shifts. We further assume that the heterogeneity in preference revision coefficients and shifts can be decomposed as follows:

$$
\begin{aligned}
\delta_h^A &= X_h^A \cdot \zeta^\delta + v_h^{\delta A} && (2.7)\\
\delta_h^B &= X_h^B \cdot \zeta^\delta + v_h^{\delta B} && (2.8)\\
\alpha_h^{BA} &= X_h^A \cdot \zeta^\alpha + v_h^{\alpha BA} && (2.9)\\
\alpha_h^{AB} &= X_h^B \cdot \zeta^\alpha + v_h^{\alpha AB} && (2.10)
\end{aligned}
$$

where $v_h^{\delta A}\left(v_h^{\delta B}\right) \sim N\left(0, \sigma_\delta^2\right)$, $v_h^{\alpha BA}\left(v_h^{\alpha AB}\right) \sim N\left(0, \sigma_\alpha^2\right)$.

Note that X_h^A (X_h^B) is a vector containing an intercept (the term "1") and explanatory variables that could be either specific to member A (B), e.g., age, or common to both of them, e.g., household income.

2.2.2 Individual-level Utility for Joint Consumption

After household h is formed, a preference revision process completes and is irreversible. From $t = 1$, individual household member's utility for joint consumption in each period depends on his/her revised latent preference, as well as temporal factors including behavioral interaction, state-dependence, observed time-varying factors and unobserved random shocks.

Consider the individual-level utilities for joint consumption for two members (i.e., member A and member B). Member A consumes g^A and member B consumes g^B. The utility of member A to consume g^A consists of several parts. First of all, A's intrinsic utility from g^A can be described by his or her revised preferences on it. Second, A's utility for joint consumption may also depend on other household member's choices since there may be a behavioral interaction which can play a significant role in determining member $A's$ utility. For example, A may gain an extra utility to consume the same choice if A consumes with B together, compared with the case when A consume alone. When g^A=g^B, member A and member B consume the product together. Third, there may be some unobserved time-specific factors, such as state-dependence, time-specific quality or price. Finally, there may also be an unobserved random shock which captures the unknown random factors affecting A's utility on g^A.

Specifically, the joint consumption utility includes five components, including 1) latent preference after preference revision, 2) utility of behavioral interaction, 3) utility of state dependence, 4) time-specific factors, and 5) random shocks. Denoting utility of member A and member B in household h at time t as $U_{ht}^A\left(g^A, g^B\right)$ and $U_{ht}^B\left(g^A, g^B\right)$ respectively, we have the following

$$U_{ht}^A\left(g^A, g^B\right) = z_{hg^A}^A + I_{ht} \cdot \theta_{hg^A}^A + S_{hg^A,t-1}^A \cdot \tau_{hg^A}^A + \lambda_h^A \cdot Q_{g^A t} + \varepsilon_{hg^A t}^A \tag{2.11}$$

$$U_{ht}^B\left(g^A, g^B\right) = z_{hg^B}^B + I_{ht} \cdot \theta_{hg^B}^B + S_{hg^B,t-1}^B \cdot \tau_{hg^B}^B + \lambda_h^B \cdot Q_{g^B t} + \varepsilon_{hg^B t}^B \tag{2.12}$$

where

$$I_{ht} = \begin{cases} 1 & if\ g^A = g^B \\ 0 & otherwise \end{cases}$$

and $\left(\varepsilon^A_{hgt}, \varepsilon^B_{hgt}\right) \sim MVN[0, \Sigma_\varepsilon]$ for any g; $S^A_{hg^A,t-1} = 1$ if member A in household h consumes g^A also at time $t-1$; otherwise, $S^A_{hg^A,t-1} = 0$. $\varepsilon^A_{hg^At}$ is a random shock to member A in household h on g^A at time t. $\theta^A_{hg^A}$ ($\theta^B_{hg^B}$) and $\tau^A_{hg^A}$ ($\tau^B_{hg^B}$) describe behavioral interaction and state-dependence for member A (B) respectively where I_{ht} and $S^A_{hg^A,t-1}$ are indicators for these two components respectively. Q includes observable temporal factors, and λ^A_h describes A's sensitivity to Q. Σ_ε captures unobserved covariation of member A and member B in unobserved random shocks. Note that, for identification, following most probit models, we assume that the diagonal elements in Σ_ε are 1. We further explain behavioral interaction, state-dependence and time specific factors as follows.

The behavioral interaction we model here is fundamentally different from preference revision as preference revision describes how a member's utility is shifted for all occasions and time by the difference of revised preferences between two members, while behavioral interaction captures how a member's utility depends on the other group member's behavior (e.g., choice g) on a given occasion (Yang et al. 2010). In particular, following previous literature (Hartmann 2010; Yang et al. 2010), we model behavioral interaction as $I_{ht} \cdot \theta^A_{hg^A}$ and $I_{ht} \cdot \theta^B_{hg^B}$ for member A and B respectively (as shown in the equations above). $\theta^A_{hg^A}$ ($\theta^B_{hg^B}$) is the utility member A (B) gains or loses by consuming together with the other member B (A), capturing member A's (B's) tendency to consume the same choice with others. In case both members choose the same choice g together, the behavioral interaction effect kicks in, in this way, the utility of member A could depend on member B 's choice g^B. For example, a husband (member A) and his wife (member B) may both enjoy watching Drama together (i.e., $\theta^A_{h,drama} > 0$ and $\theta^B_{h,drama} > 0$), or the wife enjoys watching Sports with her husband

($\theta^B_{h,sports} > 0$) but the husband does not ($\theta^A_{h,sports} < 0$). Note that this utility from behavioral interaction is in addition to members' latent utility/preference, and is gained or lost only when both members consume together (i.e., $I_{ht} = 1$). Similar to preference revision, we further examine the heterogeneity of θ^A_{hg} and θ^B_{hg} by decomposing them as follows:

$$\theta^A_{hg} = X^A_h \cdot \zeta^\theta_g + v^{\theta A}_{hg} \tag{2.13}$$

$$\theta^B_{hg} = X^B_h \cdot \zeta^\theta_g + v^{\theta B}_{hg} \tag{2.14}$$

where $v^{\theta m}_h \sim N\left(0, \sigma^2_\theta\right)$ for $m \in \{A, B\}$.

With preference revision and behavioral interaction, we capture the intragroup interdependence and interaction across members. Here we use state dependence to account for dependence across time (Dube, Hitsch and Rossi 2010; Heckman 1991; Gupta, Chintagunta and Wittink 1997). $S^A_{hg^A,t-1}$ ($S^B_{hg^B,t-1}$) is an indicator variable to show whether A (B) consumes the same choice in last and current periods. $\tau^A_{hg^A}$ and $\tau^B_{hg^B}$ measure how the consumption choices in a previous period affect household members' current utilities. For example, a household member could be more likely or less likely to continue watching a drama if he/she already watched the drama in the last period. Again, similar to preference revision and behavioral interaction, we further examine the heterogeneity of τ^A_{hg} and τ^B_{hg} by decomposing them as follows:

$$\tau^A_{hg} = X^A_h \cdot \zeta^\tau_g + v^{\tau A}_{hg} \tag{2.15}$$

$$\tau^B_{hg} = X^B_h \cdot \zeta^\tau_g + v^{\tau B}_{hg} \tag{2.16}$$

where $v^{\tau m}_{hg} \sim N\left(0, \sigma^2_\tau\right)$ for $m \in \{A, B\}$.

We further assume that there are time-varying factors, Q, which affect individuals' choices. Thus, Q is time-specific and could also be choice-specific. For example, in the TV viewing setting, an individual A may prefer the choice "drama" better when the number of available drama programs is higher. In this case, the number of available drama programs (similarly, the number of available comedy programs and so on) is one of the time-varying factors. The coefficients of the time-varying factors, λ_h^A and λ_h^B, shows how sensitive individuals' choices are to the time-varying factors. Once more, similar to preference revision, behavioral interaction and state dependence, we further examine the heterogeneity of λ_h^A and λ_h^B by decomposing them as follows:

$$\lambda_h^A = X_h^A \cdot \zeta^\lambda + v_h^{\lambda A} \tag{2.17}$$

$$\lambda_h^B = X_h^B \cdot \zeta^\lambda + v_h^{\lambda B} \tag{2.18}$$

where $v_h^{\lambda m} \sim N\left(0, \sigma_\lambda^2\right)$ for $m \in \{A, B\}$.

2.2.3 Group Decision-Making Process

After setting up the individual-level utility for joint consumption, we examine how group members make decisions together. A variety of different group decision-making mechanisms have been studied by previous literature, which can be classified into two main categories: 1) cooperative or collective mechanisms, which predict that under certain sharing rules members of a group will cooperate and reach Pareto-efficient intragroup allocations, and 2) non-cooperative or strategic mechanisms, which predict that each member of a group will act to maximize his or her own utility and the group will reach Nash equilibrium in aggregate. In this chapter, we assume that each household member may have different influence (i.e., decision power) for joint consumption. To make a decision, each household member follows a predefined decision-making mechanism to maximize either individual-level or household-level utility. Here, we demonstrate how group decision-making mechanism can be incorporated into our model and how decision power may play a role in the consumption process. With a probit model framework, one of the advantages of the proposed model is its flexibility to incorporate different forms of decision-making mechanisms. Our framework can be easily extended to incorporate other decision mechanisms in the future. In Appendix D, we discuss another decision-making mechanism, which we call "Joint Optimization"; we also estimate the model with the joint-optimization mechanism to show the ability of the proposed model to extend to other decision mechanisms.

In this chapter, we focus on a decision mechanism which we call "Winner Maximization." We assume there is a "winner" in each period where all members use the "Winner Optimization" decision mechanism. The winner has the priority to make the consumption decision first over his or her partner. The rationale behind this mechanism is that individuals try to maximize his or her own utility. The one who has higher weighted utility will "win" and will be able to move first. Specifically, we have

$$Winner \quad = \quad \begin{cases} A & if\ \gamma_h \cdot U_{ht}^{A*} > (1-\gamma_h) \cdot U_{ht}^{B*} \\ B & if\ \gamma_h \cdot U_{ht}^{A*} \leq (1-\gamma_h) \cdot U_{ht}^{B*} \end{cases} \tag{2.19}$$

where $U_{ht}^{A*} = max\left(U_{ht}^{A}\right)$ and $U_{ht}^{B*} = max\left(U_{ht}^{B}\right)$ are A and B's maximum utilities at time t over all feasible choices respectively; γ_h refers to the weight of member A's utility, which refers to decision power (of A) in this chapter; whereas $1-\gamma_h$ is the decision power of B.

Then the decision process follows a two-step sequential game where the winner moves first followed by the other household members. In this case, there is one and only one subgame perfect equilibrium in this sequential game. Therefore, one and only one consumption decision can be reached following this decision mechanism. Specifically, in the case that A is the winner, using backward induction, we have the following conditions for the subgame perfect equilibrium $\left(g^A, g^B\right)$:

$$Condition\, A \quad \equiv \quad \begin{cases} U_{ht}^{A}\left(g^A, g^B\right) \geq U_{ht}^{A}\left(\hat{g}^A, \widetilde{g}^B\right) \\ U_{ht}^{B}\left(\hat{g}^A, \hat{g}^B\right) \geq U_{ht}^{B}\left(\hat{g}^A, \widetilde{g}^B\right) \end{cases}$$

where $\hat{g}^A$ refers to the feasible consumption choices for A; whereas $\widetilde{g}^B$ refers to any feasible consumption choices of B given $\hat{g}^A$. The second condition means that B chooses his or her best option $\hat{g}^B$ given A's move $\hat{g}^A$; while the first condition means that, considering B's response, A makes his or her decision to maximize his/her utility. Similarly, following the process described above, we can have the conditions for the subgame perfect equilibrium when B is the winner:

$$Condition\ B \equiv \begin{cases} U_{ht}^B\left(g^A, g^B\right) \geq U_{ht}^B\left(\hat{g}^A, \widehat{g}^B\right) \\ U_{ht}^A\left(\hat{g}^A, \hat{g}^B\right) \geq U_{ht}^A\left(\widetilde{g}^A, \hat{g}^B\right) \end{cases}$$

In summary, the probability of household members to choose a choice of $\left\{y_{ht}^A = g^A, y_{ht}^B = g^B\right\}$ is

$$\begin{aligned} & p\left\{\left(y_{ht}^A = g^A, y_{ht}^B = g^B\right)\right\} \\ = & \, p\left\{\gamma_h \cdot U_{ht}^{A*} > (1-\gamma_h)\cdot U_{ht}^{B*} \bigcap U_{ht}^A\left(g^A, g^B\right) \geq U_{ht}^A\left(\hat{g}^A, \widehat{g}^B\right) \bigcap U_{ht}^B\left(\hat{g}^A, \hat{g}^B\right) \geq U_{ht}^B\left(\hat{g}^A, \widetilde{g}^B\right)\right\} \\ & + p\left\{\gamma_h \cdot U_{ht}^{A*} \leq (1-\gamma_h)\cdot U_{ht}^{B*} \bigcap U_{ht}^B\left(g^A, g^B\right) \geq U_{ht}^B\left(\hat{g}^A, \widehat{g}^B\right) \bigcap U_{ht}^A\left(\hat{g}^A, \hat{g}^B\right) \geq U_{ht}^A\left(\widetilde{g}^A, \hat{g}^B\right)\right\} \end{aligned} \tag{2.20}$$

where the first part on the right hand side is the probability when A is the winner whereas the second part is that for B is the winner. Similar to Browning, Chiappori and Lewbel (2013) and Cherchye, Rock and Vermeulen (2012), we further assume decision power of member A in a two-member household h as follows

$$\gamma_h = \frac{exp\left(\left[X_h^A - X_h^B\right]\cdot\psi + v_h^\gamma\right)}{1 + exp\left(\left[X_h^A - X_h^B\right]\cdot\psi + v_h^\gamma\right)} \tag{2.21}$$

where $v_h^\gamma \sim N\left(0, \sigma_\gamma^2\right)$. Note that, ψ is not identified for X variables which are common to both A and B, for example, household income, as in standard choice models.

In all, in this section, we propose a theory of joint consumption consisting of three parts. First is the latent preference which includes base preference and a preference revision component to examine how individuals revise their preferences after forming a group. After that, we model individual-level utility for joint consumption based on latent preferences and several temporal factors including behavioral interaction, state dependence, observable time-specific factors and unobservable random shock. Finally, we illustrate how a group decision-making mechanism can be incorporated into our theory and investigate the effect of decision power on group decisions.

2.2.4 A Single-Member Household

In Section 2.2.1 to Section 2.2.3, we describe a model for a two-member household. When there is only one member in the household (i.e., a single-member household), we do not have the three types of intragroup dynamics. Instead, we model its preference on choices, state dependence and sensitivity to the time-varying factors. Specifically, since there is no preference revision for single-member households, preference of a single-member household $\hat{h}$ on a choice g can be considered as the same as his base preference, which can be described following equations (2.1) and (2.2). Similarly, its state dependence can be described following equations (2.15) and (2.16), whereas its sensitivity to the time-varying factors can be described following equations (2.17) and (2.18). Particularly,

$$z_{\hat{h}g}^{1P} = X_{\hat{h}g}^{1P} \cdot \beta_g + v_{\hat{h}g}^{z1P} \tag{2.22}$$

$$\tau_{\hat{h}g}^{1P} = X_{\hat{h}}^{1P} \cdot \zeta_g^{\tau} + v_{\hat{h}g}^{\tau 1P} \tag{2.23}$$

$$\lambda_{\hat{h}g}^{1P} = X_{\hat{h}}^{1P} \cdot \zeta_g^{\lambda} + v_{\hat{h}g}^{\lambda 1P} \tag{2.24}$$

where $v_{\hat{h}g}^{z1P} \sim N\left(0, \sigma_z^2\right)$, $v_{\hat{h}g}^{\tau 1P} \sim N\left(0, \sigma_\tau^2\right)$, $v_{\hat{h}g}^{\lambda 1P} \sim N\left(0, \sigma_\lambda^2\right)$; superscript "1P" stands for single-member household.

2.3 Model Estimation, Identification and Simulation

2.3.1 Model Estimation

We now discuss the estimation of our model using aggregate and de-identified data. Specifically, we consider that only household-level data is available and the data may also be de-identified, but we would like to estimate the model at the individual level to infer heterogeneous individual preference and intragroup dynamics. To achieve our estimation goal, we generate individual-level data by imputing the missing information of who consumes which choice. Our estimation process as stated involves two iterative Bayesian steps: first, we use data augmentation to impute missing data about individual choice, and then we estimate the model parameters using the individual-level choice data generated in the data augmentation step.

Specifically, assume that at time t, members in household h consume $\left\{Y_{ht}^{A} = g^{A}, Y_{ht}^{B} = g^{B}\right\}$. In summary, there are three possible cases of observed data. In the first case, we observe that none of the choices are consumed, in this case, we can infer that $\left\{Y_{ht}^{A} = 0, Y_{ht}^{B} = 0\right\}$ without any further imputation. Second, we may observe that one choice g is consumed. This corresponds to three situations in individual-level consumption: either A consumes it alone $\left\{Y_{ht}^{A} = g, Y_{ht}^{B} = 0\right\}$, or B consumes it alone $\left\{Y_{ht}^{A} = 0, Y_{ht}^{B} = g\right\}$, or A and B consume it together $\left\{Y_{ht}^{A} = g, Y_{ht}^{B} = g\right\}$. This refers to the case of aggregate data where consumption information of A and B are aggregate. Third, we may observe that two choices $\{g_1, g_2\}$ are consumed. This corresponds to two possible situations in individual-level consumption, either A consumes g_1 and B consumes g_2 $\left\{Y_{ht}^{A} = g_1, Y_{ht}^{B} = g_2\right\}$ or the other way. This refers to the de-identified data where we know that both A and B consume, but we cannot identify who consumes which one. We summarize the observed information, Y_{ht}^{obs} as follows

$$Y_{ht}^{obs} = \begin{cases} g \; if \; \{Y_{ht}^{A} = g, Y_{ht}^{B} = g\} \bigcup \{Y_{ht}^{A} = g, Y_{ht}^{B} = 0\} \bigcup \{Y_{ht}^{A} = 0, Y_{ht}^{B} = g\} \\ \{g_1, g_2\} \; if \; \{Y_{ht}^{A} = g_1, Y_{ht}^{B} = g_2\} \bigcup \{Y_{ht}^{A} = g_2, Y_{ht}^{B} = g_1\} \end{cases} \tag{2.25}$$

where $g_1 \cdot g_2 \cdot (g_1 - g_2) > 0$. A challenge to implement data augmentation for this model is that there may be state dependence across periods. So, individual members' choice at time t (which is missing) depends on another missing information, the individuals' consumption choice at time $t-1$; and it could further rely on the missing information at time $t-2$ and so on. Therefore, it is hard to compute or adequately sample the posterior distribution of missing data. In this case, traditional data augmentation (e.g., Tanner and Wong 1987) is not applicable.

To impute time-correlated missing individual-level information, we propose a forward-backward augmentation algorithm as detailed in Appendix A, which incorporates the dependence of time-correlated missing information and the uncertainty of estimated values simultaneously. After the imputation step, we then have obtained an individual-level data set which can be used for parameter estimation. In this chapter, we use a Bayesian MCMC Gibbs sampling framework. Specifically, we draw unknown parameters one-by-one using Metropolis-Hastings steps conditional on all other parameters. Then we estimate the heterogeneity structure on intragroup dynamics, state-dependence, and sensitivity across households using regressions from the corresponding equations. A detailed description of the parameter estimation process is provided in Appendix A.

2.3.2 Identification

In this chapter, two identification sources are used: one is the variation across time; the other is the existence of single-member households. Specifically, there are two types of variation across time. First, available choice set changes over time: some choices may not be available in certain periods (for example, in our TV viewing setting, some genres may not be available in some periods); second, there could be time-varying factors which affect individual's choice decision (for instance, in our TV viewing setting again, the available number of programs may change over time).

To illustrate our identification strategy, consider an example: TV viewing in a household with two members, A and B. Although we don't necessarily directly observe how member A and member B behaved when they were single[12], the information we learn from similar single-member households[13] (A' and B') can serve as the "reference points". This information helps us to have an idea about how member A and member B will respond to variation across time in expectation. For example, imaging A is expected to be sensitive to time-varying factors while B is not. We then observe how consumption at the aggregate level responds to the variation across time to infer who is (are) watching. If we see that for a given household, the household-level watching of a genre g is sensitive to the time-varying factors, then it indicates that most genre g may be watched by A since A is sensitive to time-varying factors while B is not. Similarly, across time, the available choice sets can be different. Changes on available choice sets are important sources of variation which help us identify who is (are) watching. For example, if A is more likely to switch to drama when sports is not available, while B is more likely to switch to comedy. Then, observing at the household level about what the household switches to (e.g., drama or comedy) when sports is not available will provide us useful information about who is more likely to watch sports in this household.

The identification sources discussed above provide us information about "who consumes which" when only household-level data are available, enabling us to augment individual

[12]Note that our identification strategy does not rely on observing each member's behavior before cohabitation.

[13]Here, similar single-member households can be defined through various attributes, such as demographics, geographics, consumption habits, etc.

member's choices from observed household-level data. Once individual-level data of all choice decisions are augmented, we are able to identify individual's revised preferences and intragroup dynamics. For example, what genres member A usually watches when he or she watches alone will help us pin down A's revised preferences. The revised preferences, together with the information learned from similar single-member households, further enable us to infer the component of preference revision, the shift and heterogeneity in preferences (through regression in equations (2.5) and (2.6)). Specifically, the difference between the revised preferences and the expected preferences based on single-member households will provide us information about how the individual revises his or her preference and how he or she is affected by the other member in the same household. Behavior interaction of a genre g is determined by how often the two household members watch genre g together. For example, if A and B watch a little bit of drama when they watch TV alone, but they watch a lot of drama when they watch TV together, it is likely that they have a high positive behavioral interaction on watching drama together. Decision power is identified by what genres are watched when A and B watch together. Specifically, it is identified by how the consumption pattern is in line with A's preference when A and B consume together. For example, if A likes watching sports but dislikes drama, then a high decision power for A will make it more likely for A and B to consume sports instead of drama when they watch TV together. Please note that the identification of preference revision and decision power rely on the viewing distribution across genres when people watch alone and watch together respectively, so, for identification purposes, we assume that, for individuals, preference revision and decision power are homogeneous across genres. If a member A has high decision power on watching a genre g, we assume he/she will have the same decision power across all genres. Table 2.1 below summarizes the main sources of identification and whether the parameters are homogeneous across genres or genre-specific.

Our identification strategy can be seen from another perspective. Consider a household with member A and B again. Based on their demographics and information learned from

Table 2.1: Identification of Preference Heterogeneity and Group Dynamics

	Watching alone	Watching together
Homogeneous across genres	Preference Revision and Shift	Decision Power
Genre-specific	Preference Heterogeneity	Behavioral Interaction

similar single-member household members, we will be able to expect, from aggregate level (i.e., household level) data, how the viewing behavior of this household will be if there is no intragroup dynamics and preference heterogeneity. Then the difference between observed aggregate level viewing behavior and the expected viewing behavior may come from any one or more of the three types of group dynamics or preference heterogeneity. For example, in a household, we see that the observed aggregate-level data has a higher viewing of drama than expected. This may be because (assuming that we expect A likes drama more according to similar single member households): (1) preference revision that B is influenced by A and B becomes a person who likes drama more; (2) behavioral interaction that A and B like watching drama together; (3) decision power that A has more power to decide what to watch; or (4) A has a preference heterogeneity in drama that he or she likes drama more than other similar single-member households. So the question is how we can disentangle these three group dynamics as well as preference heterogeneity. Variation across time provides us the information to separate these potentially confounded effects. Specifically, these four factors, although they all can cause higher viewing of drama at the household level, the effects of these components actually have different extents, shapes, or different directions as long as time-varying factors change over time. For example, assume that we may expect the effect to become stronger (i.e. bigger differences between observed and expected viewing of drama) over time if it is due to decision power of A; while the effect may remain relatively constant if it is due to preference heterogeneity. In this case, by observing how the household-level viewing behavior changes over time along with the variation from time-varying factors, we will be able to infer where (which factors) the effect comes from. In Appendix B, with several simulated examples, we show how variation across time helps us disentangle these three confounded intragroup dynamics as well as preference heterogeneity.

Our identification idea is similar to some identification strategies in collective models in family economics, where a central issue is how to infer individual preferences and intra-household decision processes using individual's labor supply and household-level consumption data. Various methods have been used and discussed for identification of such kind of models. For instance, Blundell et al. (2000) test and discuss how price variations help the identification of collective models, by estimating price responses under different price conditions (using aggregate consumption data). Another approach for identification is to include single-member households where individual's consumption behavior is observed and ready to estimate individual preferences, and then use these estimates to further identify the decision process in households with two or more members (Chiappori and Donni 2009). In this chapter, we observe how consumption at the aggregate level responds to changes in the choice sets (e.g., one or more choices are not available) and the time-varying factors to infer individual preferences and intragroup dyamics. This is similar to the idea of utilizing price variations as a source of identification for collective models, in the sense that both approaches observe changes of aggregated consumption behavior due to variation in variables which affect individuals' consumption decisions.

2.3.3 Parameter Recovery and Model Comparisons Using Simulation Studies

We test and show the empirical identification of our individual-level model using household-level data in simulation studies. In the simulation, we generate a data set consisting of 120 households (80 two-member households and 40 single-member households), with three covariate variables, four available choices and 500 periods. All parameters and explanatory variables are randomly drawn from normal distributions or multivariate normal distributions. Specifically, the first covariate variable is for intercept estimation, so the first covariate variable is 1. The second and third covariates are randomly drawn from a normal distribution with a mean of 0 and standard deviation of 0.5. We then simulate parameters so that we can have a setting close to the empirical data set. For example, the estimation results of empirical data set have an estimated z mostly between -2 and -3; correspondingly, in the simulation, we draw βs from a normal distribution with mean of -2 and standard deviation of 0.2 to generate the similar scale of z in the simulation. Following the same logic, we draw ζ_α, ζ_γ, ζ_θ, ζ_τ, ζ_λ, ζ_δ from normal distributions with mean of 0.3, 0.2, 1.0, 1.0, 0.5, -0.5 respectively (and standard deviation of 0.1, 0.1, 0.2, 0.2, 0.1, 0.1 respectively). We then aggregate the viewing data into the household level and estimate the model using aggregated viewing data only. We first show that the full model (Model 1) is able to recover simulated parameters. To show the biases caused by ignoring group dynamics, we further compare the full model (Model 1) which includes all three components of intragroup dynamics with the following four models: a model without preference revision (Model 2), a model where all individuals have the same decision power (Model 3), a model without behavioral interaction (Model 4), and a model without any intragroup dynamics (Model 5).

Table 2.2 summarizes the parameter estimation results, where the second column is the true values that the simulation is based on, followed by five columns for estimation results of the five models respectively. For each model, we run two chains of MCMC, each for 20,000 iterations and the first 10,000 iterations are discarded by trace plots and Gelman-Rubin diagnostics (Gelman and Rubin 1992). The MCMC chains are thinned to remove autocorrelation between draws and every 50th draw is used for the subsequent analysis. For each parameter, we present the mean and standard deviation (in parenthesis) of its kept MCMC iterations. For parameters which are genre-specific, each of them is a 3*4 matrix (three covariates and four choices); for parameters which are not genre-specific, each of them is a 3*1 matrix. For example, $\beta[1,2]$ is the intercept term for choice 2. From

Table 2.2: Estimation Results of Five Models in Simulation

Parameter	TRUE	Model 1	Model 2	Model 3	Model 4	Model 5
$\beta[1,1]$	-2.19	-2.27(0.07)	-2.67(0.04)	-2.32(0.03)	-2.3(0.03)	-2.37(0.04)
$\beta[1,2]$	-2.05	-2.00(0.09)	-2.40(0.02)	-2.11(0.05)	-1.91(0.03)	-1.89(0.04)
$\beta[1,3]$	-2.35	-2.41(0.09)	-2.71(0.03)	-2.47(0.03)	-2.3(0.04)	-2.32(0.03)
$\beta[1,4]$	-2.34	-2.37(0.08)	-2.82(0.03)	-2.4(0.04)	-2.35(0.04)	-2.38(0.03)
$\beta[2,1]$	-1.79	-1.61(0.14)	-2.19(0.07)	-1.60(0.09)	-1.94(0.09)	-1.76(0.19)
$\beta[2,2]$	-2.12	-2.03(0.09)	-2.50(0.04)	-1.98(0.06)	-2.41(0.08)	-1.92(0.08)
$\beta[2,3]$	-2.36	-2.27(0.08)	-2.68(0.09)	-2.19(0.06)	-2.56(0.08)	-1.98(0.08)
$\beta[2,4]$	-2.03	-1.82(0.11)	-2.34(0.06)	-1.66(0.09)	-2.12(0.07)	-1.85(0.06)
$\beta[3,1]$	-1.77	-1.70(0.13)	-2.20(0.08)	-1.84(0.09)	-2.19(0.10)	-1.9(0.07)
$\beta[3,2]$	-1.81	-1.71(0.11)	-2.09(0.07)	-1.77(0.09)	-2.11(0.06)	-1.67(0.05)
$\beta[3,3]$	-1.88	-1.92(0.08)	-2.14(0.1)	-2.00(0.1)	-2.27(0.1)	-1.89(0.05)
$\beta[3,4]$	-1.77	-1.77(0.09)	-2.30(0.05)	-1.78(0.08)	-2.23(0.09)	-1.76(0.07)
$\zeta_\alpha[1,1]$	0.27	0.22(0.04)		0.32(0.03)	0.45(0.03)	
$\zeta_\alpha[2,1]$	0.32	0.25(0.06)		0.38(0.06)	0.41(0.06)	
$\zeta_\alpha[3,1]$	0.26	0.17(0.06)		0.23(0.04)	0.26(0.03)	
$\zeta_\gamma[2,1]$	0.16	0.26(0.21)	-0.03(0.22)		-0.58(0.18)	
$\zeta_\gamma[3,1]$	0.17	0.11(0.16)	0.03(0.15)		-0.44(0.17)	
$\zeta_\theta[1,1]$	0.57	0.74(0.06)	0.77(0.04)	0.53(0.05)		
$\zeta_\theta[1,2]$	1.14	1.06(0.07)	1.01(0.04)	1.04(0.06)		
$\zeta_\theta[1,3]$	0.84	0.86(0.09)	0.78(0.05)	0.83(0.05)		
$\zeta_\theta[1,4]$	0.62	0.93(0.09)	0.92(0.05)	0.87(0.06)		
$\zeta_\theta[2,1]$	0.97	1.06(0.1)	0.81(0.1)	0.64(0.16)		
$\zeta_\theta[2,2]$	0.82	0.90(0.12)	0.68(0.08)	0.54(0.09)		
$\zeta_\theta[2,3]$	0.99	0.87(0.13)	0.75(0.06)	0.91(0.09)		
$\zeta_\theta[2,4]$	1.22	1.31(0.13)	0.42(0.14)	0.65(0.09)		
$\zeta_\theta[3,1]$	1.26	1.2(0.12)	1.15(0.12)	0.75(0.11)		
$\zeta_\theta[3,2]$	0.69	0.86(0.10)	0.73(0.08)	0.4(0.11)		
$\zeta_\theta[3,3]$	1.18	1.07(0.11)	0.91(0.18)	1.25(0.11)		
$\zeta_\theta[3,4]$	1.02	0.77(0.15)	1.13(0.11)	1.29(0.1)		
$\zeta_\lambda[1,1]$	0.36	0.39(0.04)	0.4(0.04)	0.38(0.03)	0.41(0.04)	0.33(0.02)
$\zeta_\lambda[2,1]$	0.48	0.50(0.07)	0.53(0.07)	0.53(0.05)	0.50(0.06)	0.41(0.06)
$\zeta_\lambda[3,1]$	0.60	0.68(0.06)	0.76(0.07)	0.61(0.05)	0.70(0.06)	0.67(0.06)
$\zeta_\tau[1,1]$	1.04	0.99(0.06)	1.04(0.08)	1.13(0.07)	0.95(0.06)	1.01(0.07)
$\zeta_\tau[1,2]$	0.86	0.96(0.05)	0.90(0.03)	0.94(0.04)	0.98(0.03)	0.95(0.05)
$\zeta_\tau[1,3]$	1.25	1.25(0.06)	1.14(0.04)	1.28(0.05)	1.32(0.03)	1.36(0.03)
$\zeta_\tau[1,4]$	1.07	0.98(0.05)	0.87(0.05)	0.85(0.04)	0.97(0.04)	1.03(0.04)
$\zeta_\tau[2,1]$	1.10	1.15(0.1)	1.15(0.09)	1.4(0.1)	1.16(0.1)	1.3(0.08)
$\zeta_\tau[2,2]$	0.91	0.94(0.12)	0.83(0.1)	0.86(0.14)	0.9(0.08)	1.05(0.07)
$\zeta_\tau[2,3]$	0.82	0.7(0.1)	0.67(0.06)	0.78(0.09)	0.78(0.12)	0.82(0.08)
$\zeta_\tau[2,4]$	0.89	0.71(0.09)	0.56(0.07)	0.58(0.08)	0.78(0.07)	1.03(0.11)
$\zeta_\tau[3,1]$	0.96	1.07(0.12)	1.2(0.07)	1.01(0.18)	0.91(0.11)	0.99(0.06)
$\zeta_\tau[3,2]$	0.97	1.08(0.21)	0.92(0.09)	0.97(0.09)	0.96(0.17)	0.92(0.06)
$\zeta_\tau[3,3]$	0.83	0.94(0.09)	0.64(0.08)	0.94(0.1)	1.01(0.11)	1.09(0.07)
$\zeta_\tau[3,4]$	1.04	1.05(0.1)	1.01(0.11)	0.94(0.09)	1.24(0.09)	1.18(0.09)
$\zeta_\delta[1,1]$	-0.52	-0.62(0.05)		-0.58(0.03)	-0.44(0.03)	
$\zeta_\delta[2,1]$	-0.59	-0.89(0.16)		-0.89(0.06)	-0.15(0.07)	
$\zeta_\delta[3,1]$	-0.44	-0.60(0.12)		-0.28(0.07)	0.00(0.07)	

Table 2.2, we show that, first of all, the full model can recover the simulated parameters in the sense that almost all (45 out of 47) estimated parameter values are not significantly different from their corresponding simulated true values at a significance level of 0.05, with 2 parameters slightly deviating from true values. Second, comparing results in Model 2 to Model 5 with the true parameters, we show that ignoring any one of the group dynamics may cause biases on the estimation results. In Model 2 to Model 5, there are 24 out of 41 parameters, 19 out of 45 parameters, 18 out of 35 parameters, 10 out of 27 parameters respectively having true values out of the 95% posterior interval.

For example, in Model 2 where we ignore the preference revision, there are biases on estimating individual preferences and other group dynamics. In this simulation study, we assume an overall negative preference shift (i.e. δ) in the preference revision process. Ignoring preference revision process will ignore such a negative shift from single-member households to two-member households and thus downwardly bias the estimation of intercept terms in β (as shown in Table 2.2). In Model 3, where we assume that individuals have the same decision power, there are also biases on estimating preferences. In households where member A is simulated to have decision power higher than 0.5 (that is A has more power than B to decide what to watch), ignoring decision power in the estimation process would pull upward the estimated preferences for all genres of member A. In Model 4, where we ignore the behavioral interaction, we have bias on preference estimation as well as preference revision and decision power estimation. Note that ignoring behavioral interaction will distort the augmentation directly about whether household members watch TV together. In this simulation study, we simulate the overall positive behavioral interaction between A and B. Ignoring such kind of behavioral interaction will downwardly bias the number of times that A and B watch together and upward bias the number of times they watch alone. Since some cases where A and B watching together are incorrectly inferred as A or B watching alone, A and B would be estimated to be similar to each other (more than it should be), which upward biases the estimation of preference revision. Similar to

Model 2 to Model 4, Model 5 which ignores all three group dynamics also has estimation biases on preferences.

2.4 Empirical Application: Household TV Viewing and Targeted TV Advertising

In this section, we apply the proposed model to the setting of television viewing and targeted advertising on Nielsen People Meter (NPM) data. Each subject in NPM panel presses his or her unique meter whenever he or she watches TV. The NPM data contains individual-level viewing data which reports when and what each household member watched on TV, as well as household and person classification data. But before we estimate our model, we first aggregate the TV viewing behavior to the household level to generate another data set which is similar to the set-top box data. We call it as the aggregate and "de-identified" NPM data. Then, we take no notice of intra-household individuals' TV viewing behavior and estimate our model using the aggregate and "de-identified" NPM data. In short, although we use an individual-level data set, the individual-level viewing data are used only for model validation and discussion. We illustrate how to estimate the proposed model using aggregated and de-identified data, and further demonstrate the importance to incorporate group dynamics.

2.4.1 Data Description

The NPM data set consists of two parts: one is TV viewing data, and the other is household and person classification data. Each TV viewing record contains information about who (which household member) watches what TV program at what time. Household and person classification data describe characteristics of households and individuals, including household size, household income, whether there is cable in the household, whether the household has internet, geographic location, as well as individual-level demographics including age, gender, education and working hours each week.

In this chapter, following previous studies, we focus on TV viewing behavior in prime time of weekdays (i.e., Monday to Friday 8:00 PM to 11:00 PM) for 1,204 households in 24 weeks. Our sample includes 411 single-member households and 793 two-member households. Specifically, we focus on households which have one or two members, no child and no cable in the household. We do not include households with cable because, in the NPM data set, we do not have information about these households' cable subscription and thus do not know available choices they have, an important source of identification.

Nielsen assigns each program to a genre. There are 22 genres defined by Nielsen in the NPM data set. Table 2.3 shows the top six genres by household viewing frequencies. Household viewing frequency of a genre is the percentage of household views on this genre. As displayed in Table 3, first, the top six genres contains over 95% of household viewing. Second, as shown in Table 3, there is large heterogeneity among genres. The top genre, General Drama, has a viewing percentage of over 35%, while Sports has less than 6%. In summary, households focus and spend most of their watching time on a few genres instead of evenly spending it across genres.

Table 2.3: Viewing of Genres

Genre	Percentage of Viewing	Availability
Comedy	9.4%	41.7%
News Documentary	8.4%	21.1%
General Drama	35.4%	100.0%
General Variety	6.2%	34.7%
Participation Variety	31.0%	69.3%
Sports	5.7%	12.6%
Others	3.9%	21.0%

In this chapter, we focus our analysis on the top six genres, including general drama (GD), participation variety (PV), situation comedy (CS), general variety (GV), sports event (SE), and news documentary (DN), with all other genres coded as "others" (Os). We show in Table 2.3 (column "Availability") the availability of each genre over time. For example, SE is available about 12% of the time. Most genres (except Drama) are not available in some periods respectively, which creates 37 different choice sets over time. In addition, we use the number of available programs for each genre as a time-varying factor, in the sense that, for example, a higher number of available programs for drama may lead more people watch it.

Some exploratory analyses and basic statistics indicate the possibility of intragroup dynamics in this data set. For example, there is a significant difference between the total time spent on TV for individuals in single-member households and those in two-member households. Table 2.4 summarizes the average time each of the following groups spent on each genre as well as their total time on TV: females and males in two-member households, and females and males in single-member households. For example, females in two-member households spent about 22% of their time on TV on average, while females in single-member households spent about 33% of their time on TV. The less time on TV for individuals in two-member households, compared with those in single-member households, may be due to the more attractive outside options for two-member households, and corresponds to the preference shift, δ, in our model.

Table 2.4: TV Viewing Frequency

	Two-member HH		Single-member HH	
Genre	Females	Males	Females	Males
Comedy	2.1%	1.8%	2.8%	2.5%
News Documentary	1.9%	1.3%	3.3%	2.0%
General Drama	7.7%	6.5%	11.0%	10.0%
General Variety	1.4%	1.1%	1.9%	1.5%
Participation Variety	7.1%	5.4%	11.3%	6.6%
Sports	1.0%	1.5%	1.0%	1.8%
Others	0.8%	0.7%	1.3%	1.0%
total	22.0%	18.2%	32.7%	25.3%

To check whether there is a possible effect of preference revision in an exploratory way, we take females as an example. First, we calculate the ratio of viewing frequency of females in two-member households over that of females in single-member households for each genre, which is presented by the solid line in Figure 2.1. This gives us a proxy of preference revision (from single-member females to two-member females). Per our model, we expect this kind of revision to be related to the differences in revised preferences of two members in the two-member households. Since 90% of the two-member households consist of one male and one female, we look into the ratio of viewing frequency of males in two-member households over that for females in two-member households (as presented in the dashed line in Figure 2.1), using it as a way to explore the differences in revised preferences. As shown in Figure 2.1, these two lines are highly corrected with each other, as our model predicts that preference revision depends on the differences in revised preferences.

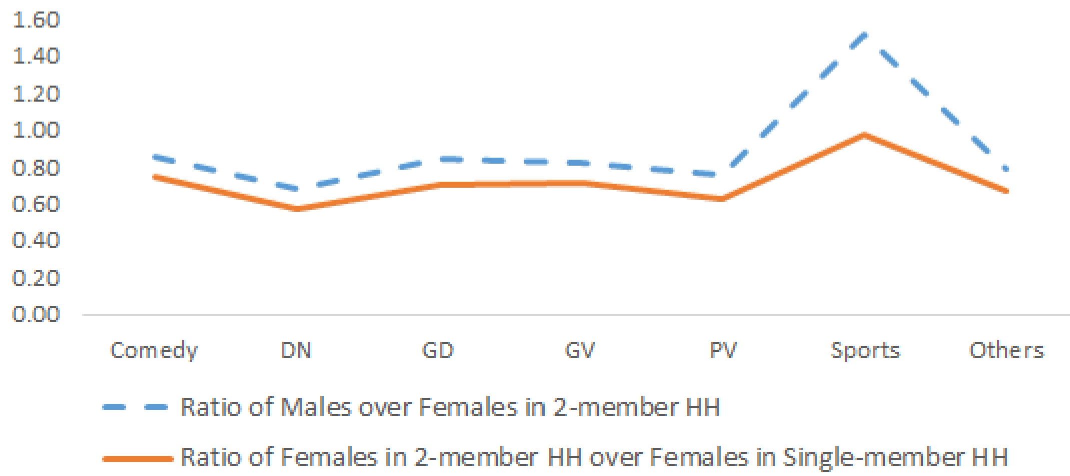

Figure 2.1: Possible Preference Revision for Females vs. Differences Between Males and Females

We also see possible effects of behavioral interaction in two-member households. In two-member households, members spend a lot of time watching TV together. Table 2.5 summarizes the percentage of times watching together versus watching alone in two-member households on each genre. For example, for every 100 times of viewing Comedy (from the household level), about 53 of them are watched by a household member alone while the other 47 are watched by two members together. As shown in Table 2.5, over 40% of the viewing time members in two-member households watch TV together. Without behavioral interaction, it is difficult to explain such a high percentage of viewing time that household members watch TV together.

Table 2.5: Watching Together vs. Watching Alone in Two-Member Households

	CS	DN	GD	GV	PV	SE	Os
Watching alone	53.4%	58.5%	55.3%	57.2%	56.3%	61.0%	60.9%
Watching together	46.6%	41.5%	44.7%	42.8%	43.7%	39.0%	39.1%

2.4.2 Estimation Results

We now report the estimation results based on the Nielsen People Meter (NPM) data. In the estimation, we incorporate the following demographics information as explanatory variables: three household-level demographics (including household income, size of geographic location (whether it is in a big city), whether the household consist of a married couple) and six individual-level demographics (including gender, age, education level, time spent on the internet, working hours per week, whether the individual is a household head). We further use the number of available programs for each genre at each period as a time-varying factor. We first aggregate the TV viewing data into household level and pretend that only household-level data are available. We run two chains of MCMC for the estimation, each for 20,000 iterations and the first 10,000 iterations are discarded on the basis of trace plots and Gelman-Rubin diagnostics. We further thin the MCMC chains and kept every 50th draw to remove autocorrelation between draws. Tables 2.6 and 2.7 summarize the heterogeneity in preferences, intragroup dynamics and state dependence, where we present the mean and standard deviation of central prior distributed parameters across all individuals. For example, the preference revision across all individuals has a mean of 0.33 and standard deviation of 0.06. We also show the distribution of the heterogeneity in Appendix C using Figures C.1 to C.7.

First, overall, there is evidence of positive preference revision and behavioral interaction among household members. That is, on average, household members positively affect each other's preferences on TV genres (thus revise their preferences toward each other) and enjoy watching TV together. Furthermore, there is a negative shift from single-member household to two-member households overall. Household members in two-member households may enjoy more from outside options such as shopping together and thus have a negative shift of watching TV at home.

Table 2.6: Heterogeneity Across Members in Two-Member Households

	Preference Revision	Preference Shift	Decision Power	Sensitivity
mean	0.33	-0.62	0.50	0.80
St.dev.	0.06	0.18	0.04	0.29

Table 2.7: Genre-Specific Heterogeneity Across Members in Two-Member Households

		CS	DN	GD	GV	PV	SE	Others
Behavioral Interaction	mean	0.73	0.77	0.85	0.63	0.88	0.57	0.73
	St.dev.	0.77	0.77	0.66	0.76	0.71	0.82	0.82
State Dependence	mean	1.71	2.51	1.69	2.08	1.84	2.44	2.20
	St.dev.	0.82	0.72	0.50	0.68	0.55	0.73	0.69
Revised Preferences	mean	-2.83	-2.81	-2.28	-2.91	-2.39	-2.71	-2.92
	St.dev.	0.48	0.42	0.47	0.42	0.40	0.35	0.34

Second, the heterogeneity across households emphasizes the importance to incorporate and examine intragroup dynamics. Moreover, it has some important managerial implications. Take behavioral interaction as an example. It could be more effective to air ads of family products (e.g., furniture, traveling, etc.) for households with strong and positive behavioral interaction since their household members are more likely to watch TV together. That is, the model allows us to examine intragroup dynamics with heterogeneity across households, which provides us the opportunity to make better-customized marketing strategies, demonstrated later in this chapter.

Third, we find strong state dependence for all genres. Results in Table 2.7 suggest that there is heterogeneity in state-dependence across household members for each of the seven genres, and Figures C.3 and C.4 in Appendix C shows that almost all household members have strong and positive state dependence for each genre. Therefore, household members are inclined to continue watching the same genres across time. First, this provides insights into the design of TV program schedule. For example, offering TV programs of the same genre consecutively could be desirable to take advantage of the effect of state dependence, thus keeping audiences staying with TV. The strong state dependence may explain why more and more TV channels focus on one genre (to lock in and keep audiences watching). We also find strong sensitivity to the number of available programs. This may also provide some implications into the design of TV program schedule.

Table C.1 to Table C.4 in Appendix C summarize the estimation results of the effect of covariates on preferences, intragroup dynamics, state dependence, and sensitivity. From these, we can infer some relationships between modeled parameters and demographics. For example, estimation results show that, in general, higher income, higher education level and lower working hours are associated with less time on TV. Males watch more sports while females watch more PV. Higher income is associated with less behavioral interaction in GV and higher behavioral interaction in Sports. People in large cities are associated with high interaction in News. Older people are associated with high behavioral interaction in most genres. Household heads are associated with less interaction in News. Higher income and younger people are associated with higher state dependence.

Based on the estimation results, we examine how well the proposed model is able to identify who is (are) watching. Each time we observe a TV viewing from a two-member household, for each of the two members, we predict whether he/she watches the TV. So, each time, there are two conditions: a true condition which is the truth about whether a household member watches the TV from the NPM data, and a predicted condition which refers to the prediction based on the modeling results. Table 2.8 below illustrates a contingency table to examine how well the model can identify who is (are) watching. Specifically, there are four areas in the table, which stands for four different cases. For example, true positive (TP) refers to the case that people watch the TV, and our prediction correctly says they watch TV. We then look into two metrics to examine the performance of the proposed model. One is positive predictive value (PPV), which is the proportion of true watch out of all predicted watch (i.e., PPV= TP / (TP+FP)); the other is true positive rate (TPR) which is the proportion of true watch that are that correctly predicted (i.e., TPR=TP/(TP+FN)). We

Table 2.8: An Illustration Example of a Contingency Table

		True Condition	
		Watch	Not-Watch
Predicted Condition	Watch	True Positive (TP)	False Positive (FP)
	Not-Watch	False Negative (FN)	True Negative (TN)

focus on these two metrics because these two are more relevant for marketing practice (such as ad targeting) than other metrics such as false positive rate. In particular, both of these fit metrics are important for inferring TV watching. For example, in TV ad targeting, a high positive predictive value means that most ad impressions sent out correctly reach targets (i.e. correctly infer targets are watching most time), which reduces the direct ad cost to reach targets; while a high true positive rate means that most people who are watching TV are correctly found/detected, which means a low opportunity cost. If the true positive rate is low, then we will miss some targets who are watching TV, which loses the opportunity to reach them, resulting in a high opportunity cost.

In our estimation, we randomly divided all households into two groups; one is an in-sample group to run estimation, the other is an external-sample group to test and validate our estimation results. Figure 2.2 summarizes the distribution of positive predictive value and true positive rate across all two-member households for both in-sample and external sample. The overall positive predictive value and true positive rate are around 75% and 68% respectively for the in-sample group. The external sample has comparable results, with 72% and 65% for the overall positive predictive value and true positive rate respectively. As shown in Figure 2.2, the proposed model has a very high positive predictive value for most households (e.g., in some households, the positive predictive value is as high as more than 90%). This is particularly important for TV ad targeting, in the sense that it enables us to focus on those households with high positive predictive value (i.e., so we can correctly infer who is watching) to target the desired audience efficiently. Note that some households have lower positive predictive value and true positive rate; these households usually have sparse data, and it is difficult to make inference about individuals' watching behavior.

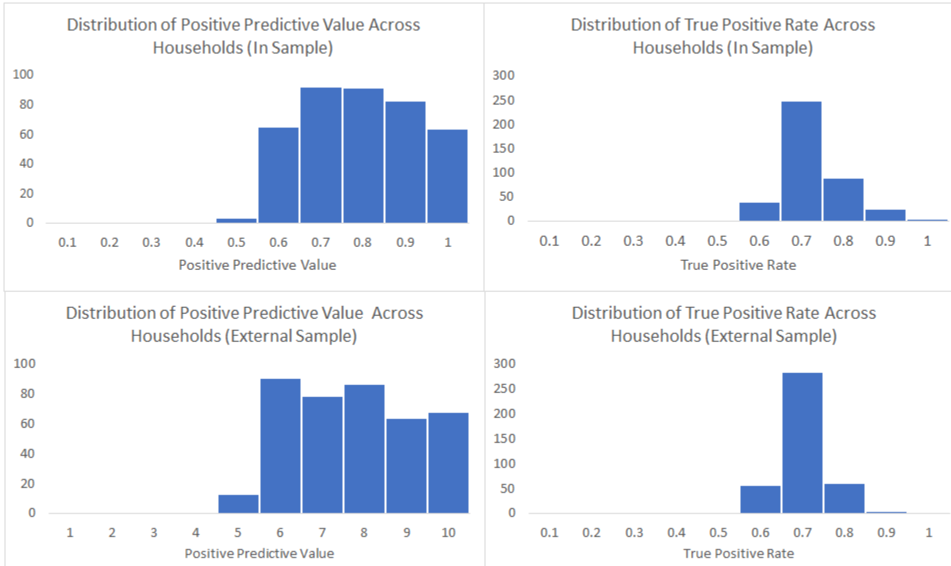

Figure 2.2: Distribution of Positive Predictive Value and True Positive Rate across Two-Member Households

2.4.3 Model Comparisons and Targeted TV Advertising Campaign

We compare the proposed model with the four benchmark models: each of first three models turns off one type of group dynamics (preference revision, behavioral interaction and decision power) respectively to see the consequence of ignoring each of these three group dynamics; the fourth benchmark model is a model without any of the three intragroup dynamics. We first compare the full model with the first four benchmark models in terms of DIC which has been widely used in Bayesian models as a model selection criteria. Table 2.9 summarizes the DIC for these five models. It is clear that the full model has lower DIC than benchmark models, indicating that the full model outperforms the other models in term of model fit.

Table 2.9: DIC of Five Models

	DIC
Full Model	571438
w/o Preference Revision	571941
w/o Decision Power	572646
w/o Behavioral Interaction	575351
w/o any group dynamics	580590

We then conduct a counterfactual analysis on TV targeting using the NPM data set based on the estimated results. Specifically, imaging that a TV advertiser wants to target particular types of audiences, for example, males with ages ranging from 31 to 50. Further assume that household members' demographics information is available and observed, so that the TV advertiser is aware of the households that contain 31 to 50-year-old male members. We call these households as "targeted households" and 31 to 50-year-old male members as "targeted audiences." Specifically, we pretend that we do not have individual-level data and only observe household-level viewing data. In this case, once a TV viewing from a targeted household is observed, the TV advertiser decides whether to air its ads or not. Note that, these households may contain household members who are not targeted (e.g., the wives in these households).

To compare the proposed model with the benchmark models, we run a counterfactual experiment on targeting 31 to 50 years old males. Specifically, assume that an advertiser would like to have a certain amount of exposures from targeted audiences (i.e., 31 to 50 years old males), where an exposure (EX) is counted each time the ad is exposed to a target audience. Note that, some targeted households may consist of two targeted audiences. In these households, if both members are watching TV while the ad is aired, then it is counted as two exposures. To save advertising cost, the advertiser tries to reach its goal of exposures from targeted audiences by airing its ad as few times as possible. To do so, the advertiser needs to infer how likely that the targeted audience(s) in the household is (are) watching TV once a TV viewing from a targeted household is observed. It then sends its ad (an impression) to households where the ad would be most likely watched by a target. Then the cost of the advertiser can be measured by the ratio of impressions over exposures, which describes the number of impressions needed to get an exposure to the targeted audiences. The lower the ratio, the lower the cost, and thus the more efficient it is to target.

We look into two different scenarios. In one scenario, the advertiser would like to reach 2,000 exposures whereas in the second scenarios the advertiser would like to reach 3,000 exposures. We examine the expected number of impressions needed by each model in each of the two scenarios, comparing with randomly sending out the ad to targeted households. First, we find that the proposed full model is able to improve the efficiency to reach targeted audiences. For example, for the in-sample data set, the average impression-to-exposure ratio is around 1.82 if we randomly send out the ad to targeted households. That is, for 2000 exposures and 3000 exposures, we expect to need about 3640 and 5460 impressions respectively. The proposed full model can reach the same exposures with much fewer impressions (i.e., 2027 and 3547 impressions respectively). Table 2.10 displays the expected reduction in impression with each model comparing with sending out ads randomly to targeted households. Second, the proposed full model also outperforms reduced models where one or more group dynamics are ignored. For example, to reach 2,000 exposures to the targeted audiences, the proposed model needs only 2027 impressions while the model without group dynamics needs 2528 impressions which is 35% higher. Moreover, the impression-exposure ratio is lower when the number of exposures is smaller. Specifically, for the proposed full model, to reach 2,000 exposures, the impression-exposure ratio is about 1.01 while the ratio is about 1.18 to reach 3,000 exposures. This indicates that the proposed model is able to correctly find out households which are more likely to watch. So with fewer exposures, the model will focus more on targeted audiences who are more likely to watch.

Table 2.10: Expected Reduction in Impressions Needed

	In-Sample		Out-of-Sample	
	2,000 Exposures	3,000 Exposures	2,000 Exposures	3,000 Exposures
Full Model	44.3%	35.0%	33.0%	30.3%
w/o Preference Revision	43.7%	34.9%	27.6%	24.2%
w/o Decision Power	41.4%	32.6%	32.5%	29.5%
w/o Behavioral Interaction	35.6%	21.1%	29.9%	27.4%
w/o any group dynamics	30.5%	16.8%	27.4%	26.8%

2.5 Discussion and Potential Future Extensions

In this section, we discuss future extensions to the model and research. First, there are other strategies which may help us identify the proposed model. We briefly discuss some of them below, including (1) time-varying covariates for consumers, (2) survey data, (3) auxiliary aggregate data, and (4) exclusion consumption. Incorporating these identification strategies may help us identify the proposed model better. Second, the current study looks into two-member households. In the future, we may extend the proposed model to groups/households with three or more than three members.

2.5.1 Other Identification Strategies

First, there are two types of time-varying covariates: one is covariates for the consumption goods; the other is covariates for household members. The first category is discussed and incorporated in this chapter. The second category includes variables describing the changes of household members across time. For example, suppose we have a variable describing the available leisure time for each household member each period. Undoubtedly, such time-varying covariates could help model identification by providing variation over time and examining how consumption choices change correspondingly. Another example is a variable indicating whether household members are available for watching TV in each period. Then, say in some periods, if only one household member has time for watching TV, we will be able to observe the consumption of this member in these periods, which enables us to identify the preferences of this household member well akin to the variation currently provided by single-member households.

For instance, marketers constantly collect survey data to understand the genre viewing habits of households. For example, a survey on, if a genre g is watched, what is the probability that it is watched by member A or B, or A and B together. Suppose such data for each genre is available, we then have a prior distribution $\mathscr{F}_g$ of, when a genre g is watched, what is the probability that it is watched by A, B and A and B together respectively. We then can implement a method which is similar to the Information Reweighted Prior (IRP) method proposed by Wang (2012), except that we incorporate it with the data augmentation method.

Another possible external data source is aggregate TV viewing data which contains information about what percent of individuals watch genre g at time t. Existing TV rating data may be feasible if it contains information for each genre (instead of specific programs). Assuming that we have a data set which includes the following aggregate information: the percentage that member As (Bs) watch genre g at time t, denoted as π_{gt}^A (π_{gt}^B). We can then use this information as a constraint to facilitate the data augmentation process (as shown in Chen and Yang 2007; Musalem, Bradlow and Raju 2008):

$$\sum_h y_{hgt}^A = H \cdot \pi_{gt}^A \quad (2.26)$$

$$\sum_h y_{hgt}^B = H \cdot \pi_{gt}^B \quad (2.27)$$

where $y_{hgt}^i = 1$ if individual i in household h watching genre g at time t according to augmented data.

Finally, an approach that has been used to identify the collective models is to have an exclusion assumption that each household member does not consume at least one good. For example, in a two-member household, member A never watches drama, and member B never watches Sports. Chiappori and Ekeland (2009) show that such an exclusion assumption can guarantee generic identifiability of the collective models. Our proposed model is surely different from collective models, but this identification strategy can shed some light on improving identification of the proposed model. If member A never watches drama, we would know that all drama viewing is from member B alone. Similarly, if member B never watches Sports, we would know all Sports viewing is from A alone. This information provides us individual-level data on some choices, which reduces our uncertainty in the data augmentation and improves the identification of the proposed model.

2.5.2 More than Two Group Members

Another possible extension is to have more than two members in the households. We develop our model assuming that there are two group members. The model proposed can easily be extended to incorporate more than two group members. When there are more than two group members, the model development process is similar to what we present, except that it will need to incorporate more parameters due to the model complexity. We start our discussion with the three intragroup dynamics. First, the preference revision can be generalized as

$$z^i_{hg} = X^i_h \cdot \beta^i_g + \delta^i_h + \sum_{i' \neq i} \left(z^{i'}_{hg} - z^i_{hg} \right) \cdot \alpha^{i,i'}_h + v^i_{hg} \tag{2.28}$$

where $\alpha^{i,i'}_h$ measures the influence of group member i' to member i; whereas δ^i_h is the preference shift for member i; If we allow asymmetric preference interdependence then $\alpha^{i,i'}_h \neq \alpha^{i',i}_h$. Second, the behavioral interaction can be generalized as $\sum_{i' \neq i} \theta^{i,i'}_{hg} \cdot I^{i,i'}$ where

$\theta_{hg}^{i,i'}$ measures the behavioral interaction between group member i' and member i. Third, decision power of member i is γ_h^i where $\sum_i \gamma_h^i = 1$. Model components other than group dynamics can then be easily extended to the case of more than two group members. For example, each household member may have its own state dependence term for each choice.

2.6 Conclusion

When there is a lack of rich individual-level data, we only observe choices at the group level and therefore lack important information about individuals' behavior. For decades, it presented a great barrier and a central issue which hinders marketers to underpin new waves of marketing innovations (e.g., personalization). To solve this issue, we propose a new model and a novel algorithm to infer individual preferences and intragroup dynamics using just commonly available aggregate and de-identified data. Even if only aggregate and de-identified data are available, marketers can apply our method to estimate an individual-level model with aggregate and de-identified data. With our method, marketers can better understand who actually consumes the products, and how individuals within a group interact and influence each other.

In particular, we propose a joint consumption theory which focuses on three important components of intragroup dynamics (i.e., preference revision, behavioral interaction, and decision power). We show that we can disentangle these three components apart in our setting without observing individual-level data. The model is estimated with just aggregate and de-identified data in two iterative Bayesian steps: first, we use data augmentation to impute missing data about individual choices, and then we estimate the model parameters using the individual-level choice data generated. To impute time-correlated missing individual-level information based on the aggregate and de-identified data, we propose a forward-backward augmentation algorithm, which incorporates the dependence of time-correlated missing information and the uncertainty of estimated values simultaneously.

In this chapter, we offer a new identification strategy by leveraging cross-sectional and longitudinal variation to identify heterogeneous individual preferences and different components of intragroup dynamics. We identify the proposed model with aggregate and de-identified data using two identification sources. One is the variation across time due to time-varying factors and changes of choice set over time. The other is the existence of single-member households (where we can observe individual-level behavior directly). To illustrate how we achieve identification, a series of in-depth simulations and empirical validations are provided.

To further calibrate our proposed model and algorithm, we applied our model to an empirical setting of household TV viewing and targeted TV advertising on a data set obtained from Nielsen People Meter (NPM) data where we knew the truth, but pretended that we didn't. This empirical application provides us several valuable insights into intragroup dynamics. Overall, we find that preference revision and behavioral interaction among household members tend to be positive. That is, household members positively affect each other's preferences on TV genres (thus revise their preferences toward each other) and enjoy watching TV together. Next, we found a negative shift from single-member household to two-member households overall. Household members in two-member households may enjoy more from outside options such as shopping together and thus have a negative shift of watching TV at home.

Finally, we use our estimates to conduct a series of calibrated counterfactual simulations demonstrating that our proposed model will enable advertisers to better target intragroup individuals and significantly improve the efficiency of ad targeting. In addition, our proposed model will make it possible for marketers to design better marketing strategies for a variety of marketing contexts, such as customizing promotions catering to the "powerful" member's preference, conducting effective product recommendation based on inference of behavioral interaction, and enhancing the customer experience by better understanding customer joint-consumption processes.

This chapter could be extended in several aspects in the future. First, we focus on intragroup dynamics in this chapter and assume that the decision mechanism is known. In the future, we may want to identify the decision mechanism. However, this may require much more rich data and variation in the data for the identification purpose. Second, we assume that all households have the same decision mechanism. In the future, we may extend this to allow heterogeneity in decision mechanism across households. Third, for identification purpose, we assume that preference revision and decision power in a two-member household are homogeneous across genres. In the future, we may release these two assumptions and allow genre-specific preference revision and decision power.

In summary, learning about intragroup dynamics and individual preferences within a group is important, but has been understudied. In this chapter, we propose a joint consumption model which can be used to jointly infer heterogeneous individual preferences, intragroup dynamics, and state dependence when only aggregate and de-identified data are available. We hope this model is useful for marketers to examine heterogeneous intragroup individual preferences through observed group consumption choices in a variety of marketing contexts, which may further enable them to design customized marketing strategies such as targeting specific types of individuals.

Chapter 3

Inferring Individual Preferences and Variety Seeking with Non-ordered Data: An Application to Video Games

3.1 Introduction

As consumers spend time engaging in experiential products, they may become satiated on certain product attributes and exhibit preferences for new experiences (variety-seeking), or they may become "hooked" on certain familiar features and prefer consistency (inertia). It is useful for firms to identify when consumers are variety-seeking or inertial when predicting purchase patterns and offering recommendations for future consumption occasions. For example, to a variety-seeking video-game player, it could be more effective to recommend him a different gaming environment from his previous consumption instead of a similar one. In the present research, we focus on how game players choose between the various options across multiple consumption occasions within the same product category. For example, in a video game setting where players play an online video game round by round (each round they choose a map to play, and there are multiple maps available), each round is a consumption occasion, and maps are the available options for each round.

Considering the importance of understanding variety-seeking behavior, extensive research has been done to examine the reasons of variety-seeking (e.g., Kahn 1995, etc.), and to develop models for variety-seeking behavior (e.g., Seetharaman and Chintagunta 1998, etc.). In this chapter, we intend to identify and fill some gaps between research on variety-seeking and marketing practice. First, to model and quantify variety-seeking behavior across available options, we would first need to know the order of the options consumed by each customer. However, in some cases, we only observe non-ordered data where consumption order information is missing. This can be commonly seen in some industries. For example, in a grocery shopping setting, customers may do grocery shopping once a week or even once a month. So retailers usually observe customers' consumptions in a certain period (i.e. a week or a month) but they do not observe customers' consumption order within the period. Another example is in the DVD rental industry. Customers usually rent multiple DVDs at the same time. But their consumption order (i.e., watching order in DVD rental example) is often not recorded or untrackable.

In this chapter, we propose an identification strategy to handle non-order data, and conduct empirical analysis in a video game setting where players play a video game round-by-round. Particularly, we observe the rounds played by each player each day, as well as the map used in each round, but we do not have information about the order of the rounds/maps played within a day. To understand individual's variety-seeking behavior, the consumption order within a day is critial. To overcome the challenges emerging from non-ordered data, we propose an augmentation algorithm building on Musalem, Bradlow, and Raju (2008, 2009) in the model estimation process to obtain an augmented data set with full inferred consumption order within each day. Then, how a customer switches from a map to another provides us information to infer the propensities of variety-seeking and inertia for individual customers.

Second, we hypothsize that consumption outcomes may affect consumers' variety-seeking behavior. Researchers in psychology and consumer behavior have been paying extensive attention to the effect of emotions and moods on variety-seeking and inertia (e.g., Kahn and Isen 1993; Poor, Duhachek, and Krishnan 2012). Although emotions and moods can be potentially associated with consumption outcomes (e.g., a good consumption outcome could potentially be associated with a positive mood), little systematic investigation has been paid directly on how consumption outcomes affect a consumer's variety-seeking behavior. From a marketing practice perspective, learning about how consumption outcomes affect consumers' variety-seeking would enable the firm to improve its marketing strategies corresponding to the observed consumption outcomes.

We predict that positive consumption outcomes lead to inertial preferences, while negative consumption outcomes lead to variety-seeking. For example, imagine that a Netflix user just started watching the Netflix Original Series, Jessica Jones. After binge-watching 6 episodes, she gives the show 5 out of 5 stars, which indicates a positive consumption outcome and will likely lead to her continuing to watching the show or choosing to watch similar shows, for example, Daredevil which shares several attributes with Jessica Jones, including being a gritty action series based on Marvel superhero characters. By the same logic, if the user hated Jessica Jones and gave it 1 star, then it's likely that she would prefer to watching something completely different the next time she logs onto Netflix.

Third, prior research has shown that variety-seeking behavior could be related to product attributes. For example, satiation on product attributes over time is one of the proposed reasons for variety-seeking behavior (Kahn 1995). Moreover, consumers could have different extent of variety-seeking on different attributes. For example, Inman (2001) found that consumers have more intensive variety-seeking on sensory attributes such as flavor than on nonsensory attributes such as brand. In this chapter, we develop a model to quantify the extent of variety-seeking at the attribute-level. From a perspective of marketing practice, the attribute-level variety-seeking could potentially shed light on product design (e.g., what attributes to be incorporated into the design of a product) or recommendations.

In sum, we develop a model to quantify the variety-seeking behavior at individual and attribute level, taking into account the possible effect on consumption outcomes. Moreover, we estimate the model with non-ordered data where consumption order within a period is not observed. We use simulation studies to illustrate the identification ideas and show that our estimation process can well recover the simulated parameters. We then apply the proposed model to a data set in an online video-game environment. Across 30 to 40-minute rounds of play, individual players choose which map they want to play on and experience consumption outcomes that can be measured by their performance or points earned during the round. We find that overall better performance results in players choosing similar maps in subsequent rounds, while poorer performance results in players choosing maps with different attributes. Furthermore, there is wide heterogeneity across players in terms of the effect on consumption outcomes on variety-seeking behavior. The heterogeneity across players suggests the importance to make customized marketing strategies for an individual player instead of the same strategy for all players.

This chapter contributes to both the academic literature and managerial practice. Regarding academic literature, to our best knowledge, we are among the first to explore Bayesian data augmentation as a solution for non-ordered data. Traditional augmentation practices in marketing mostly focus on the missing information about whether an individual consumes the product or not in a period. Instead, we emphasize the order of the consumption occasions within a period. This consumption order information is important to examine the dynamics of consumer behavior over time. Second, we look into variety-seeking behavior and propose an important factor, consumption outcomes, which may drive the variety-seeking and inertial behavior. This is a new factor examined, in addition to some factors proposed in existing literature.

In addition to the contribution to the literature, our research also provides interesting insights to managerial practice. First, this chapter may help firms better design or provide recommendations to their customers. In many industries, such as video gaming, restaurants, travel, vacation, and so on, recommendations have been proven to be effective. Our method provides a framework to examine individual's variety-seeking behavior even when the order of consumption is missing, which enables the firms to design customized recommendations to an individual customer, based on consumers' past consumption options and outcomes. Second, we model the variety-seeking behavior at the attribute level, which hasn't been studied and may provide new insights on option design. For example, if customers are variety-seeking on one attribute but inertial on another, the firm may design options accordingly by offering more variations on the former attribute. In sum, in the video game context examined, our findings would inform the company on how to match players based on performance level and map preferences, as well as on future map designs and releases. Although we focus specifically on the context of online video games, our findings can be applied to a broader set of experiential products, including watching movies and dining at restaurants, and so on.

The remainder of this chapter is as follows. In Section 3.2, we briefly introduce some relevant literature, followed by a model and identification discussion in Section 3.3 and data description in Section 3.4. In Section 3.5, we apply our model to a context of an online video game, in which players choose between different map environments for each round of play and consumption outcomes can be measured "objectively" by a player's performance during the round. In this section, we also show the ability of the estimation process to recover parameter values via a simulation study. Finally, in Section 3.6, we provide a summary of conclusions, managerial findings, and implications.

3.2 Related Literature

In this chapter, we focus on decisions that consumers make when choosing between different options across consumption occasions within the same product category. Variety-seeking is defined to be when consumers frequently switch between options, while inertia (or reinforcement) is defined to be when consumers repeatedly choose the same option across multiple consumption occasions. We review the main reasons that researchers have identified for why consumers exhibit variety-seeking or inertial behavior, the most recent models that have been developed to capture variety-seeking, and the behavioral literature that supports our hypothesis that positive (negative) consumption outcomes lead to inertial (variety seeking) behaviors.

3.2.1 Why Consumers are Variety-Seeking or Inertial

The concepts of inertia or reinforcement behaviors and variety-seeking were developed separately before researchers began to think of them as two ends of the same continuum, so much of the research that provides explanations for these behaviors focuses on one extreme. Early models used time-lagged variables to capture inertial choices and attributed them to "brand loyalty" (Jacoby and Kyner 1973; Guadagni and Little 1983). Researchers explored alternative explanations for inertial behaviors such as state dependence and habit persistence, which can be disentangled using more sophisticated utility models (Erdem and Keane 1996; Seetharaman 2004).

On the other hand, Kahn (1995) summarizes the three main reasons for why consumers may be variety-seeking: (1) external situations, (2) satiation, and (3) future preference uncertainty. External situations include marketing decisions made by firms. For example, different firms may promote in alternating weeks (Kahn and Raju 1991) or engage in price discrimination (Shaffer and Zhang 2000), driving consumers to switch between brands.

Satiation is a well-studied phenomenon in both behavioral and quantitative research. Satiation may occur on brands or attributes and lead consumers to seek out products with new features (McAlister 1982; Inman 2001). Finally, forward-looking consumers may use variety-seeking as a way to resolve future preference uncertainty and learn about unknown choices (Walsh 1995; Erdem 1996).

3.2.2 Dynamic Discrete Choice Models

In the classic models of variety-seeking, the underlying assumption is that the consumer is making choices between options following a first-order Markov process (Jeuland 1979; Givon 1984; Kahn, Kalwani, and Morrison 1986). The key feature is that there is an explicit variety-seeking parameter that can be estimated for an individual consumer. Brand choices are formulated as a standard logit model, but the first-order Markov property allows the probability of choice to depend on the brand that was chosen previously. The individual specific variety-seeking parameter determines whether repeat choices or brand switching is more likely between subsequent consumption occasions.

There have been various extensions to this classic model to take into account variation across brands, consumers, and time. The brand choice probabilities can be revised to include brand-specific marketing variables (Seetharaman and Chintagunta 1998). The variety-seeking parameter can also vary within shoppers by assuming they come from a flexible distribution. For example, the Beta distribution allows for a bimodal pattern that can account for shoppers switching between inertial and variety-seeking states (Trivedi et al. 1994). Heterogeneity across individuals can be modeled as individuals receiving information that arrives according to a Poisson timing function (Roy et al. 1996). In our current model, we will demonstrate the advantages of attribute-based variety-seeking using a distance between options on each attribute. We also allow variety-seeking to change over time based on previously experienced consumption outcomes, and to vary across individuals and attributes.

3.2.3 Effects of Consumption Outcomes on Variety-Seeking

Consumption outcomes or their proxies are observable for a variety of experiential products. These include star ratings for movies or TV shows on Netflix, star ratings for restaurants on Yelp, thumbs up or down for videos on YouTube, and a player's score on a video game. We are going to focus on the context of video games, which has been mostly unexplored within the marketing literature. Player scores allow for a clean, relatively objective, and continuous measure of each player's consumption outcomes.

We hypothesize that positive consumption outcomes will lead to inertia, while negative consumption outcomes will lead to variety-seeking. Our model is also able to account for the magnitude of consumption outcomes in either direction, so the degree of variety-seeking or inertia also depends on how positive or negative the experienced outcomes are, relative to some reference point. Although we model the effects of consumption outcomes across a continuum, when we examine the behavioral work in support of our hypothesis, we focus on the valence of the outcomes (positive or negative) and how they might map to emotional responses. For example, for the Netflix TV show Jessica Jones, 5 stars would indicate a positive consumption outcome. For a military-based shooting game, positive net kills would be a positive consumption outcome. Likewise, a 1-star rating for Jessica Jones would indicate a negative consumption outcome, while a net of 15 deaths would indicate a negative consumption outcome for the shooter game.

Researchers in psychology and consumer behavior have long been interested in the effects of emotions on people's choice behaviors, but there is some disagreement on how positive and negative aspects influences variety-seeking. A positive valence has been shown to increase variety-seeking behavior among enjoyable products, as long as they don't have any negative features (Kahn and Isen 1993). Positive moods seem to drive people to seek out more stimulation, but this pattern might break down at very extreme positive moods (Roehm and Roehm 2005). Other research shows that differentiation of positive and negative emotions of the experience slows the satiation process due to cognitive appraisal, and so focusing on negative emotions may result in more enjoyment of repeated experiences (Poor, Duhachek, and Krishnan 2012).

In contrast to this prior research, we focus on the aspect generated by the same source as the choices being made, rather than an external manipulation of mood. We hypothesize that positive consumption outcomes should lead people to have more inertial preferences. This is consistent with literature that suggests that encountering high-value rewards will intensify motivational states towards the same reward source (Berridge 2012) and positive rewards may "whet" the reward appetite (Wadhwa, Shiv, and Nowlis 2008). In our context, a video game player may experience a hot streak and expect continued positive outcomes from playing within the same or similar map environments. On the other hand, negative consumption outcomes lead to variety-seeking, which is consistent with the notion that helplessness and sadness result in people wanting to change their current state (Keltner, Ellsworth, and Ed wards 1993; Lazarus 1991), and they may choose to do this through consumption choices (Lerner et al. 2004). A player may feel sad or frustrated after a tough loss, but a change of scene in the next round may boost their engagement in the game.

3.3 Model

The degree of variety-seeking may vary across product categories (Kahn et al. 1986) or individuals (Givon 1984), but the model we develop is more appropriate for capturing how the degree of variety-seeking varies within individual consumers. Our model captures individual choices across multiple consumption occasions. All variables are specified at the individual level, for individual n. In the base model, the probability of a consumer choosing option j depends on the individual n's intrinsic preferences for option j, z_j^n. Without any time varying effects, the probability of each choice j is formulated by a standard logit:

$$p^n(j) = \frac{exp\left(z_j^n\right)}{\sum_{j'} exp\left(z_{j'}^n\right)} \tag{3.1}$$

Here, to examine the variety-seeking/inertial behavior, we focus on the switching behavior of choices among several options instead of whether the individual consumes an option. The above equation (32) can be considered as, given individual n consumes an option, what is the probability that she consumes option j. To capture how consumers respond to consumption outcomes over time, we specify choice preferences to be first-order Markov across rounds. So the probability of selecting map j depends on the option i that was selected in the previous round. This choice depends on the individual's intrinsic preferences for option j, z_j^n, and the distance d_{ijk} between option j and the previous option i on each attribute k (which represents the extent of difference on attribute k between the two options; the more the difference, the higher the distance):

$$p_t^n(j) = \frac{exp\left(z_j^n + \sum_k \left(\alpha_k^n + \beta_k^n \cdot X_{t-1}^n\right) d_{ijk}\right)}{\sum_{j'} exp\left(z_{j'}^n + \sum_k \left(\alpha_k^n + \beta_k^n \cdot X_{t-1}^n\right) d_{ij'k}\right)} \tag{3.2}$$

where $\alpha_k^n + \beta_k^n \cdot X_t^n$ describes the individual n's variety-seeking factor on attribute k, with α_k^n as an intercept and β_k^n describes how the variety-seeking on attribute k depends on previous consumption outcome variable X_{t-1}^n. If $\alpha_k^n + \beta_k^n \cdot X_{t-1}^n \geq 0$, then the player n is variety-seeking on attribute k at occasion t; on the other hand, if $\alpha_k^n + \beta_k^n \cdot X_{t-1}^n \leq 0$, then the individual is inertial on attribute k. β_k^n describes how outcome variable affects the variety-seeking factors. A positive β_k^n means higher outcome leads to stronger variety-seeking on attribute k. The probability also depends on the distance between the options on each attribute. So a variety-seeking player is also more likely to switch to options that are farther away, while an inertial player is more likely to choose options that are closer to the previously chosen option.

In the model above in equation (3.2), all parameters are specified at individual level (for individual n), including preferences on choice z_j^n (for $j = 1, ..., J$ where J is the total number of choices), and α_k^n, β_k^n (for $k = 1, .., K$ where K is the total number of attributes). We further assume that these parameters follow independent normal distribution across individuals.

$$z_j^n \sim N\left(\bar{z}_j, \sigma_{zj}\right) \; for \; j \in \{1, ..., J\} \tag{3.3}$$

$$\alpha_k^n \sim N\left(\bar{\alpha}_k, \sigma_{\alpha k}\right) \; for \; k \in \{1, ..., K\} \tag{3.4}$$

$$\beta_k^n \sim N\left(\bar{\beta}_k, \sigma_{\beta k}\right) \; for \; k \in \{1, ..., K\} \tag{3.5}$$

3.3.1 Identification with Non-ordered Data

First, we discuss the identification of the proposed model when we have complete data information including the map consumption order. In this case, equation (3.2) above can be written as

$$p_t(j) = \frac{1}{\sum_{j'} exp\left(z_{j'} - z_j + \sum_k \left(\alpha_k + \beta_k \cdot X_{t-1}\right)\left(d_{ij'k} - d_{ijk}\right)\right)} \tag{3.6}$$

where the probability for option j depends on the difference in intrinsic preferences $z_{j'} - z_j$ and distance on each attribute, $\left(d_{ij'k} - d_{ijk}\right)$. To identify the model, we assume that $z_1 = 0$ since we can only identify the difference among preferences instead of preferences itself.

Then we consider a case when part of the consumption order information is missing. Specifically, assuming there are W periods, in a period w, the consumer consumes the product in multiple occasions. Although we have information about what options the consumer choose on each occasion, we do not have information about the order of the occasions in this period. For example, we observed that a video game player plays map1 twice and map2 three times in a day, but we do not have information the order of the five occasions in this day. In this case, first of all, even with aggregate and non-ordered data, we have information about the total number of times a consumer chooses each option j, which provides us information about preference z_j (compared with z_1). Second, the variation across periods help us identify the variety-seeking factor.

In Appendix E, we provide a more detailed numerical example to show the identification source of the proposed model. Here we illustrate the identification idea of variety-seeking using a small example vignette. Consider a player who equally likes two maps: map1 and map2. In each period, the player plays five times. We then compare the difference in the following two cases observed from aggregate and non-ordered data (where we observe only the aggregate consumption information in each period instead of the order of the five consumption occasions). First, in case that the player is high in variety-seeking, we expect to observe the following pattern: map1 twice and map2 three times in a period, followed by map2 twice and map1 three times in another period, and then repeat. This is because if the player has strong variety-seeking, he will switch to different maps every time. Second, in case that the player has strong inertia, we expect the player to consume the same map for a long time and occasionally switch to the other map only. So we expect the following pattern from aggregate and non-ordered data: we may see, for example, in first few periods, the player always plays map1 (so we observe five map1 in each of the first few periods), and he switches to map2 occasionally, but once switches, he sticks to map2 for a while. Thus, different patterns will be observed with different levels of variety-seeking.

In short, we will be able to recover preference on a choice j based on how frequently it has been selected; and, based on the variation across periods observed with non-ordered data.

In Section 3.5.2, we use a simulation study to further show that we are able to fully recover the parameters simulated with aggregate non-ordered data.

3.4 Data Description

Our data set was awarded through the Wharton Customer Analytics Initiative (WCAI) from a large video game developer. We have data on the activity of 1,309 frequent players of an online multiplayer first-person shooter video game. Players engage in campaigns averaging 20 minutes in length in two competing teams. We focus on the rounds played on the firm's public servers and exclude the rounds played on player's private servers. There are on average 20 players involved in each round, and people rarely play with the same player twice. In this chapter, we focus on each player's rounds played across two months, starting from the game's release. Each round is considered to be a particular consumption occasion t. During each round, players are allowed to choose what map they want to play on. The map is basically a game environment with a set of attributes and features. We are interested in the player's map choice. Players are presented with a set of $M = 9$ maps. The player then chooses a map and is dropped into a server by the firm's matching algorithm with other players to play a round. After the round, the player is shown his/her individual round outcomes, which include the total score, the number of kills, the number of deaths, individual points earned for completing certain tasks in the round, etc. In the next round, the player will again have the opportunity to select a map.

Whether a player switches frequently among maps is an indicator to assess the variety-seeking of the player. However, simply looking at the switching rate could be misleading sometimes. For example, if a player switches frequently between two favorite maps which are very similar to each other, then we may see high switching rate but in fact, this player is inertia to these two maps. To address this issue, we enable attribute-specific variety-seeking in the proposed model and incorporate the distance between any two maps in each attribute. Specifically, there are two attributes for each map: Combat Type and Terrain. For example, there are four possible combat types, including CombinedArms (CA), Vehicle, Infantry, and UrbanWarfare (UW). Similarly, each map may have one or more of the eight possible terrains: Countryside (CS), Woodland (WL), Mountainous, Urban, Rocky, Docks, Desert, and Underground (U.G.). Table 3.1 summarizes the attributes of the 9 maps. For example, Map 1 has a combat type of Infantry and a terrain of Urban.

Table 3.1: Attributes of Maps

Map	Attribute 1: Combat Type			
	Combined Arms	Vehicle	Infantry	UrbanWarfare
MAP 1	0	0	1	0
MAP 2	0	0	1	0
MAP 3	1	1	0	0
MAP 4	0	0	1	0
MAP 5	1	1	0	0
MAP 6	0	1	0	0
MAP 7	1	0	0	0
MAP 8	1	1	0	0
MAP 9	0	0	1	1

Map	Attribute 2: Terrain							
	Country	Woodland	Mount.	Urban	Rocky	Docks	Desert	U. ground
MAP 1	0	0	0	1	0	0	0	0
MAP 2	0	0	0	1	0	0	0	0
MAP 3	1	1	0	0	0	0	0	0
MAP 4	0	0	0	1	0	0	0	0
MAP 5	0	0	0	0	0	0	1	0
MAP 6	0	0	1	0	0	0	0	0
MAP 7	0	0	0	1	1	1	0	0
MAP 8	0	0	0	1	1	0	0	0
MAP 9	0	0	0	1	0	0	0	1

We then describe the distance between two maps on each of the two attributes. Specifically, for example, for Combat Type attribute, we calculate the sum of the difference between two maps in four possible combat types. For example, map2 has combat type of Infantry, while map3 has combat type of CombinedArms (CA) and Vehicle. So out of the four possible types, map2 and map3 are different in three of them (the only common is that they both do not have UrbanWarfare), so their distance on Combat Type attribute is 3. Following the same way to calculate the distance between two maps on Terrain attribute, we summarize the distance on the two attributes in Table 3.2 and Table 3.3 respectively.

Table 3.2: Distance on Attribute: Combat Type

	Map 1	Map 2	Map 3	Map 4	Map 5	Map 6	Map 7	Map 8	Map 9
Map 1	0	0	3	0	3	2	2	3	1
Map 2	0	0	3	0	3	2	2	3	1
Map 3	3	3	0	3	0	1	1	0	4
Map 4	0	0	3	0	3	2	2	3	1
Map 5	3	3	0	3	0	1	1	0	4
Map 6	2	2	1	2	1	0	2	1	3
Map 7	2	2	1	2	1	2	0	1	3
Map 8	3	3	0	3	0	1	1	0	4
Map 9	1	1	4	1	4	3	3	4	0

Table 3.3: Distance on Attribute: Terrain

	Map 1	Map 2	Map 3	Map 4	Map 5	Map 6	Map 7	Map 8	Map 9
Map 1	0	0	3	0	2	2	2	1	1
Map 2	0	0	3	0	2	2	2	1	1
Map 3	3	3	0	3	3	3	5	4	4
Map 4	0	0	3	0	2	2	2	1	1
Map 5	2	2	3	2	0	2	4	3	3
Map 6	2	2	3	2	2	0	4	3	3
Map 7	2	2	5	2	4	4	0	1	3
Map 8	1	1	4	1	3	3	1	0	2
Map 9	1	1	4	1	3	3	3	2	0

In addition to the distance between two maps, an important variable in the proposed model is the consumption outcome. After each round, the player experiences a set of individual consumption outcomes. These include Total Points, Combat Points, Kills, Deaths, and Net Kills. These variables are all pretty highly correlated (see Table 3.4), so we use one of them, Total Points, to measure the overall performance. Total Points is generally the primary objective of the game. Besides, total points represent the overall performance of the player in a round, and can be considered as a measurement which takes into account all the other outcomes.

Table 3.4: Correlation of Outcomes

	Total Points	Combat Points	Kills	Deaths	Net Kills
Total Points	1				
Combat Points	0.79	1			
Kills	0.73	0.84	1		
Deaths	0.47	0.59	0.66	1	
Net Kills	0.44	0.44	0.57	-0.24	1

3.5 Model Estimation

In this section, we first describe the estimation process we use for the proposed model. Then we demonstrate the identification of the model parameters in a simulation study by showing that we are able to recover the simulated parameters with non-ordered data. Finally, we present the estimation results using the video game data set and discuss relevant findings.

3.5.1 Estimation Process

In this chapter, due to the missing information about the consumption order in a period, we incorporate a data augmentation process in the estimation process. Assuming we have W periods, in period w, the consumer consumes the products in l_w occasions (each occasion consumes an option). Since the consumption order information is missing, we do not know the order of the l_w occasions in period w. We do the estimation in a two-step iterative process as follows.

In the first step, we augment the missing information about the order of consumption occasions in each period w. With l_w occasions, there are $l_w!$ possible consumption orders in period w. So the number of possible consumption orders could increase sharply with an increase of l_w. For example, $l_w = 5$ and $l_w = 10$ correspond to 120 and 3,628,800 possible consumption orders respectively. To overcome the curse of dimensionality, we propose a switching algorithm in the augmentation process. Note that we observe the option consumed and the outcome for each consumption c, but we do not observe the order of the consumption occasions. In sum, there are l_w consumption observed in period w; the task is to "match" each of these l_w consumption (based on the option consumed and outcome) to an occasion t.

Similar to Musalem, Bradlow, and Raju (2008, 2009), we use a switching algorithm in the augmentation step. The idea of the proposed switching algorithm is to have an initial consumption order within each day first, and then gradually improve the consumption order by switching consumption occasions within the same day. The main difference between the proposed algorithm and the algorithm by Musalem, Bradlow, and Raju (2008, 2009) is that, instead of randomly generating pairs in each iteration, we follow certain orders to generate pairs to compare and decide whether to switch. Specifically, in period w, we follow the following process to augment consumed option of occasion t for $t \in \{1, ..., l_w\}$.

Algorithm 3.1 A Switching Algorithm in Augmentation Step

(a1) If there is only one occasion in period w ($l_w = 1$), we only have one consumption c and one occasion t; so we will be able to match them without augmentation;

(a2) If there is more than one occasion, from occasion 1 to l_w, we do the following:

(i) If $t = 1$, the available consumption set includes all l_w consumption in period w; otherwise, update available consumption set by removing the consumption matched to $t - 1$ (from the latest consumption set);

(ii) Switch current consumption in occasion t with each consumption in the available consumption set respectively (including a case in which we do not switch but keep current consumption order); calculate the corresponding probability of each switch according to equation (2);

(iii) Randomly choose one of the switches based on the calculated probabilities, and match it to occasion t accordingly;

(a3) Repeat the above process until finished for all periods and all occasions.

In the second step of our estimation process, with augmented consumption order information, we estimate the proposed model assuming all consumption order information is known. We implement an MCMC Gibbs-sampling framework to draw parameters one by one, including the preference on each option j, z_j^n, as well as the variety-seeking parameters for each attribute α_k^n and β_k^n. For identification purposes, preference for the first option, z_1^n, is fixed to be 0. After drawing these parameters for individuals, we draw partial pooling means and variances using conjugate priors.

3.5.2 A Simulation Study

To test the proposed estimation process and check whether we are able to recover model parameters, we simulate a data set where there are 50 players playing 9 maps in a video game. These 9 maps have the same features/attributes as the 9 maps we have in the video game data set. For each player, we simulate playing behavior in 60 days, each day has 5 occasions. Playing outcomes for each occasion, as well as parameters for individuals are randomly drawn from a normal distribution with a mean of 0 and standard deviation of 0.5. We assume that we do not observe the consumption order within a day. In sum, there are 12 parameters to estimate for each player n, including preferences for 8 maps z_j^n, and two parameters α_k^n and β_k^n for each of the two attributes. Note that preference for the first map is fixed to be 0 for identification.

For model estimation, we run two chains of MCMC, each for 10,000 iterations and the first 5,000 iterations are discarded on the basis of trace plots and Gelman-Rubin diagnostics (Gelman-Rubin 1992). The MCMC chains are thinned to remove autocorrelation between draws and every 20th draw is used for the subsequent analysis. Figure 3.1 compares the estimation results (mean of posterior MCMC iterations) with simulated values for each of the 12 parameters. Each graph is for one parameter, and each point on a graph stand for one of the 50 players. The dashed line stands for a line where estimation perfectly matches the simulated value. As shown in Figure 3.1, for each player and each parameter, the estimated value is close to the corresponding simulated value, which indicates the ability to recover simulated parameter and the identification of the proposed model.

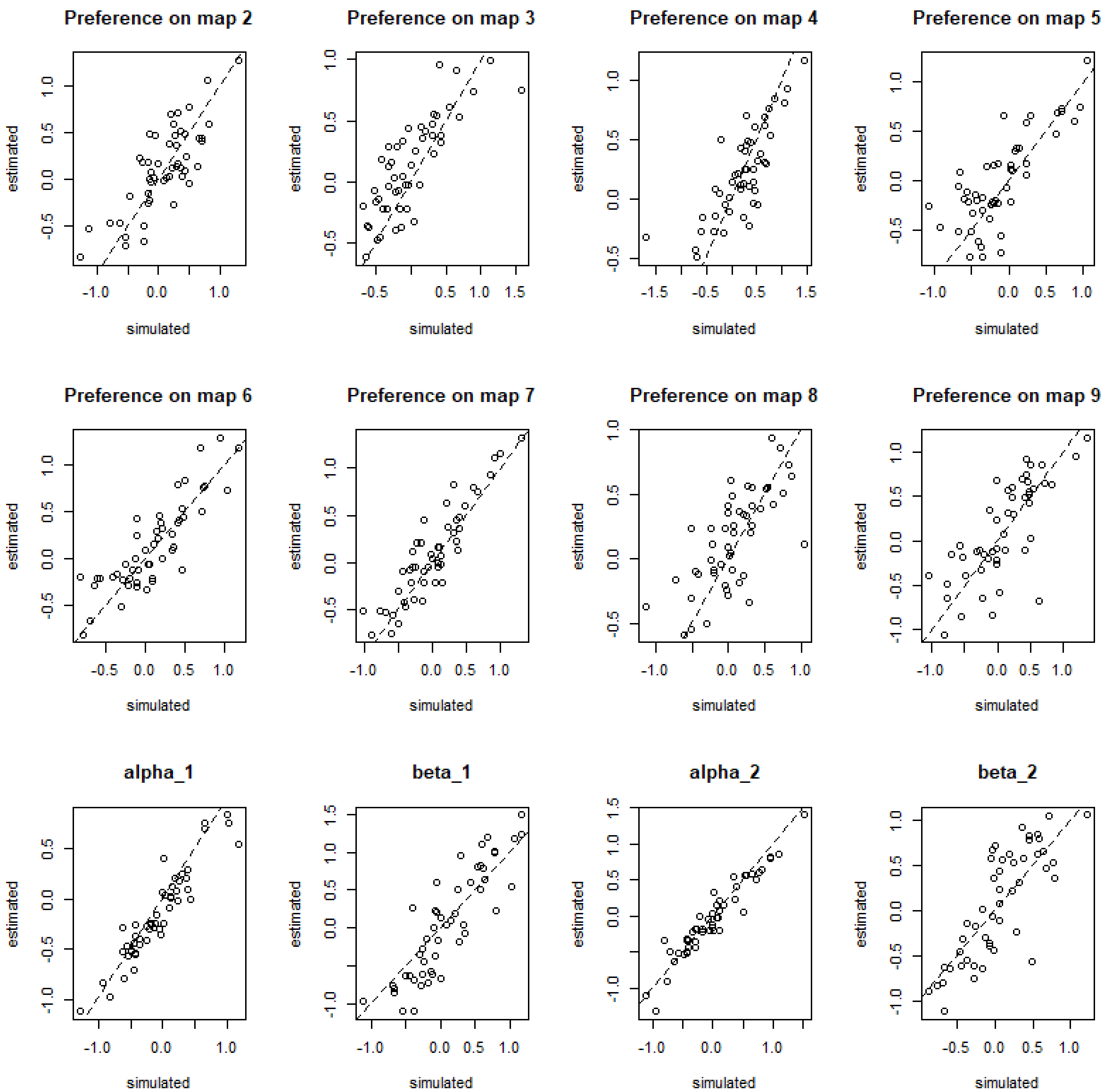

Figure 3.1: Estimation Results vs. Simulated Values

We further run a benchmark model where variety-seeking behavior is ignored. Specifically, we assume all variety-seeking behavior parameters, α and β, to be 0, and then estimate the proposed model. As shown in Figure 3.2 which compares the estimation on individual preferences with simulated parameter values, estimation results ignoring variety-seeking are biased. For example, ignoring the variety-seeking behavior, for 19 out of 50 simulated players, estimated preferences on map 2 are significantly lower than corresponding simulated preferences at a significance level of 0.01 (according to their MCMC estimation mean and standard deviation). Thus, ignoring the variety-seeking behavior may cause biases on the estimation of individual preferences.

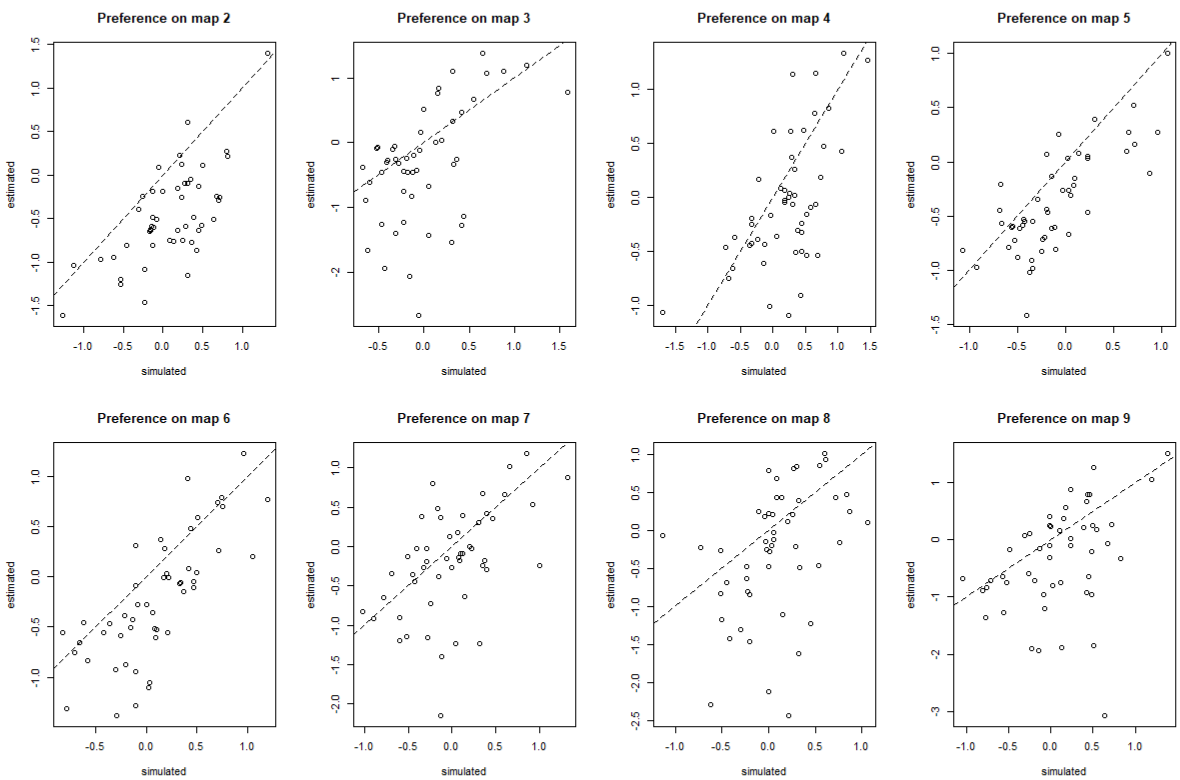

Figure 3.2: Estimation Results Ignoring Variety-Seeking vs. Simulated Values

3.5.3 Estimation Results and Discussion

We apply our model to the data set of an online multiplayer first-person shooter video game described in Section 3.4. In this video game, players play the game round-by-round, choosing a map to play each round. When the game was released, there were nine maps available for all players. Five additional map packages were released afterward; each contained three to five new maps. However, these additional packages were available only for players with premium membership or for players who pay an additional charge for the packages. In this chapter, we focus on the map choice on the original nine maps before the release of additional packages, which is a two-month time window between the release of the game and the release of the first additional package. There are 514 frequent players who have ever played the game in this two-month window.

3.5.3.1 Estimation Results

As described in Section 3.4, each map has two attributes: combat type and terrain. We look into the preference of each player on the 9 maps respectively, as well as their variety-seeking factor on each of these two attributes. We use the logarithm of total points of each round as the playing outcome. The average total points is about 3,000 across all rounds and all players. We normalize the total points each round with 3,000. So the outcome variable, logarithm of the normalized total point, indicates how much the outcome is better/worse than the overall average of 3,000. There are 12 parameters to estimate for each player n, including preferences for 8 maps z_j^n, and two parameters α_k^n and β_k^n for each of the two attributes. Note that preference for the first map is fixed to be 0 for identification. For model estimation, we run two chains of MCMC, each for 10,000 iterations and the first 5,000 iterations are discarded on the basis of trace plots and Gelman-Rubin diagnostics. We further thin the MCMC chains to remove autocorrelation between draws and keep every 20th draw. Table 3.5 illustrates the estimation results for one randomly selected player. The first row shows the estimation means of the MCMC iterations whereas the second row shows corresponding standard deviation in the MCMC iterations.

Table 3.5: An Example Estimation Result for a Player

	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	α_1	β_1	α_2	β_2
Mean	-0.49	0.95	-0.17	0.29	0.56	0.52	0.13	-0.48	-0.17	-1.00	-0.40	-0.53
Std.	0.19	0.24	0.18	0.25	0.24	0.23	0.26	0.25	0.08	0.13	0.07	0.12

In Table 3.5, z_2 to z_9 describe the preference on map 2 to map 9 respectively, with a reference of 0 on map 1. For example, the player prefers map 3 most since he has the highest preference of 0.95 on map 3. α_1 and β_1 describe the variety-seeking of this player on attribute "Combat Type" whereas α_2 and β_2 are for another attribute "Terrain". Note that, $VS_1 \equiv \alpha_1 + \beta_1 \cdot x$ stands for the overall variety seeking factor for this player on attribute "Combat Type" when playing outcome is x; a negative value on VS_1 means the player is inertial on this attribute while a positive value means the player is variety-seeking. Similarly, $VS_2 \equiv \alpha_2 + \beta_2 \cdot x$ can be used to describe the overall variety-seeking for this player on the attribute "Terrain".

Figure 3.3 illustrates the amount of variety-seeking on these two attributes and how they change with the playing outcome (which is the total points received for a playing round). The dashed line shows the variety-seeking for the Combat Type attribute while the solid line is for the Terrain attribute. As shown, both variety-seeking factors decrease when the playing outcome gets better, which means the player becomes more inertial. Second, the player switches from variety-seeking to inertia (from a positive variety-seeking to negative) at some point. For example, when the playing outcome is higher than 1600, the player is inertia on Terrain, otherwise, he is variety-seeking on this attribute.

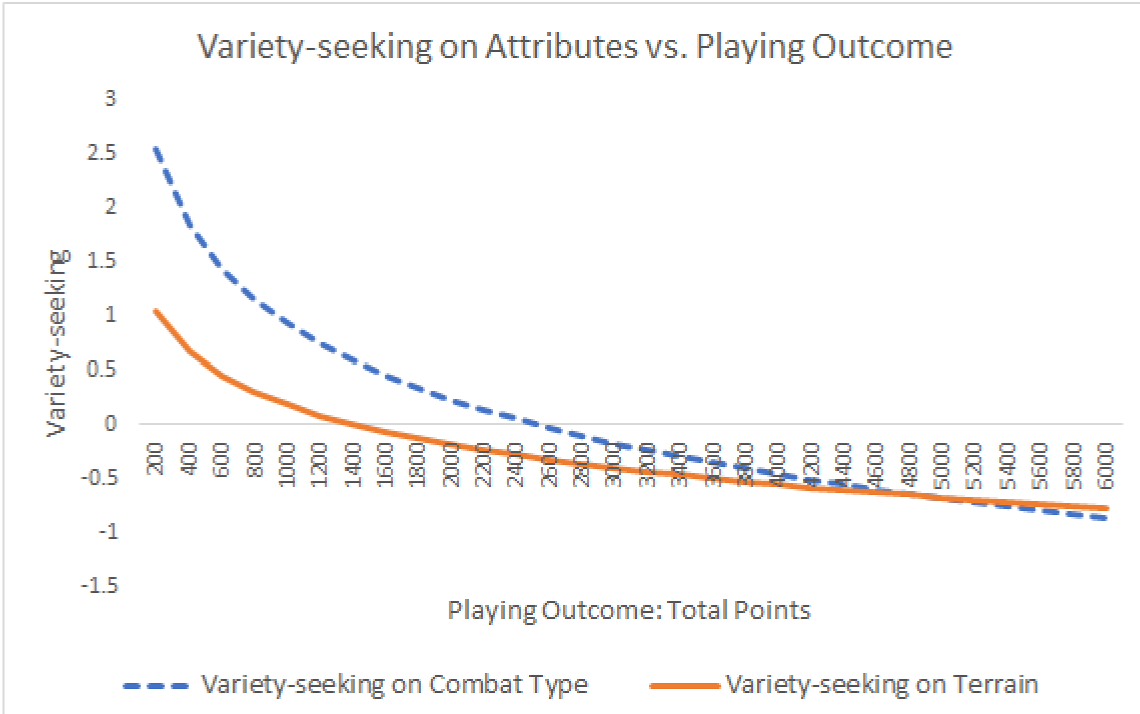

Figure 3.3: Variety-Seeking on Attributes versus Playing Outcomes

Table 3.5 provides estimation results of an example player who is randomly selected; to examine the overall variety-seeking factor on each attribute across all players, we look into the partial-pooling mean for α and β respectively. Table 3.6 provides the estimation results (mean and standard deviation of kept MCMC iterations) for four partial pooling mean parameters, including α and β for each of the two attributes. As shown, $\bar{\beta}_1$ and $\bar{\beta}_2$ are negative, which indicates that overall a better playing outcome will reduce the variety-seeking behavior and strengthen inertial behavior. Furthermore, both $\bar{\alpha}_1$ and $\bar{\alpha}_2$ are also negative. These two parameters indicate variety-seeking on the two attributes respectively when the normalized playing outcome is zero, i.e. if the playing outcome is equal to the overall mean of 3000. A negative α means that the player would be inertia (with negative variety-seeking) when he has a playing outcome equal to the overall mean. So overall, the players have the similar variety-seeking pattern on each attribute as the example player shown in Figure 3.3: they are variety-seeking when the playing outcome is low, but become more and more inertia when the playing outcome improves.

Table 3.6: Estimation of Partial Pooling Mean on Variety-Seeking Parameters

	$\bar{\alpha}_1$	$\bar{\beta}_1$	$\bar{\alpha}_2$	$\bar{\beta}_2$
Est. Mean	-0.55	-0.50	-0.17	-0.21
Est. Std.	0.07	0.07	0.05	0.05

3.5.3.2 Heterogeneity

Now we look into the heterogeneity structure of the variety-seeking parameters in Figure 3.4. Specifically, each graph in Figure 3.4 is for one of the four parameters. Each graph presents the histogram of the corresponding parameter over 514 players, where negative values are presented in orange while positive values are in blue. As shown in Figure 3.4, there is heterogeneity in each of the four parameters: most players have negative values while some are positive. For example, about 24% of players are estimated to have a positive β_1 (for attribute Combat Type). For these players, we expect them to be more variety-seeking on Combat Type when the playing outcome becomes higher.

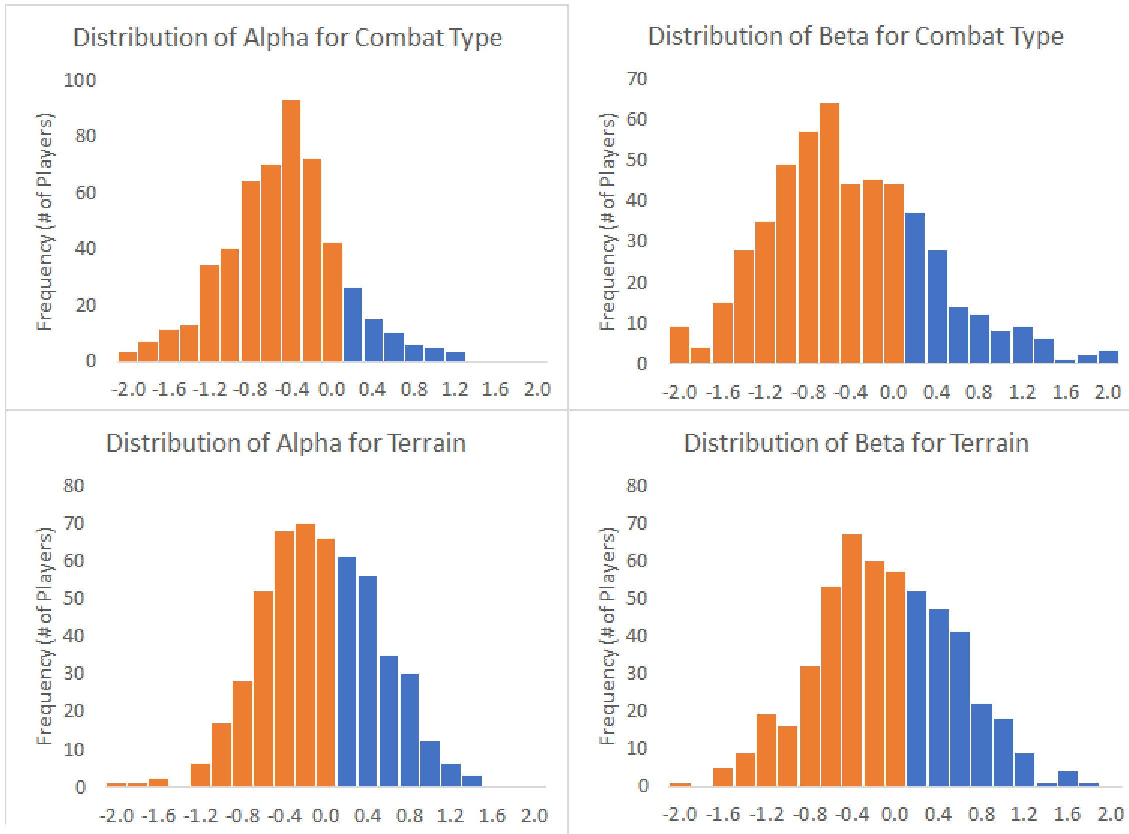

Figure 3.4: Distribution of Variety-Seeking Parameters across Players

In sum, we find that players become more inertia/less variety-seeking on both attributes when playing outcome becomes better. Furthermore, there is heterogeneity across players. To further look into how players' variety-seeking on two attributes change with playing outcome, we summarize the joint distribution of β_1 and β_2 in Table 3.7. Using a significance level of 0.05, we examine the estimation result of each player to check whether their estimated β_1 and β_2 are significantly different from 0 or not. For example, Table 3.7 shows that, only 2 players have significantly positive β_1 and β_2, so only two players become more variety-seeking on both attributes when the playing outcome improves.

Table 3.7: Distribution of β_1 and β_2

	$\beta_2 > 0$ significantly	β_2 not significant	$\beta_2 < 0$ significantly
$\beta_1 > 0$ significantly	2	17	23
β_1 not significant	19	95	80
$\beta_1 < 0$ significantly	46	160	72

One of the reasons why some players may be more variety-seeking when better playing outcomes occur may lie in the discrepancy between good playing outcome and good experience from the game. Such kind of discrepancies can come from multiple sources. First, we use total points gained from the round as the measure of the playing outcome; these outcome variables may not be appropriate for some players. For example, some players may focus on killing (the number of enemies they kill in the game) instead of total points (which is a comprehensive measure, not only killing, but also other factors are taken into account, such as the objectives destroyed, death, the weapon used and so on). Second, a good playing outcome is not always the objective. For example, for some players, their major objective could be challenging themselves, and the enjoy to overcome difficult combat or task. In this case, a good playing outcome on a map, indicating that they are skilled on the map, may make them more likely to switch to a map with different features/attributes to have more challenges in the future. The heterogeneity across players emphasizes the importance to make customized strategies to an individual player.

3.6 Conclusion

We built a model that allows for an attribute-specific parameter to describe how the variety-seeking of an individual on product attributes depends explicitly on consumption outcomes. We estimate our model with non-ordered data where we do not observe consumption order information within a period. Specifically, our estimation involves a two-step iterative

MCMC process. In the first step, we augment the consumption order using a switching algorithm; in the second step, we estimate the model with augmented data where we assume consumption order is known using an MCMC Gibbs-sampling framework.

We verify the model identification and validate the estimation process using a simulation study, where we successfully recover simulated parameters using non-ordered data. After that, we apply our model to an empirical analysis of an online video game context. The results from our analysis lend support to our hypothesis that, in general, positive consumption outcomes lead to inertia, while negative consumption outcomes lead to variety seeking. Although our analyses were conducted within the context of player map choices within an online video game, our model and findings can be extended to other domains of experiential products, including TV shows and restaurants.

There are several extensions that could be made to our current model. First, to focus on variety-seeking behavior, we look into choices among available options in this chapter but do not examine the process of whether to consume. In this future, we may look further into, for example, how consumption outcomes affect whether the consumer continues to successively consume the product or not. Second, there might be a learning process. Consumers may switch to other options to learn more about the available options. In our empirical example, players may choose different maps to learn about the features of the maps. In this chapter, we assume that consumers are familiar with available options. This assumption may be relaxed in the future to incorporate a learning process. Third, variety-seeking could be dynamic over time due to some other factors, in addition to consumption outcomes. For example, players may explore/play widely with many maps first, but with playing maps more and more, they may become gradually focus on one or two favorite maps. In this case, players' variety-seeking gradually reduces over time. Or, from another perspective, the players may become more and more variety seeking over time because they satiate on some attributes of the maps.

For managers, this research will provide a method for determining individual customer preferences and how these preferences change over time based on consumption outcomes, as well as when customers might be more susceptible to the release of new products and whether they should be novel or similar to existing products. Within the specific context of online campaign-based video games, consumption outcomes in each round of game play may be defined by the performance of the player. The firm may use these performance metrics to determine whether players are becoming bored or frustrated with the current playing experience and prefer a change of scene, or are on an exciting winning streak and want to continue with the same experience. This provides an opportunity to enhance the firm's current matching algorithm by suggesting that the player's consumption outcomes aspect their preferences over time.

In the general context of experiential products, understanding whether consumers are variety-seeking or inertial may allow firms to provide better recommendations to consumers by taking their ratings on past purchase or consumption occasions into account as a measure of their consumption outcomes. Our findings suggest that, in general, it might be effective for a firm to target satisfied customers with products that are similar to those they consumed in prior occasions, but target dissatisfied customers with products that are very different. Lastly, considering the heterogeneity across consumers observed, our model enables firms to examine each customer individually, and design efficient customized marketing strategies (e.g., recommendations, promotions, etc.) accordingly.

Chapter 4

Conclusion

In this dissertation, we extend prior research on inferring individual preferences by incorporating intragroup dynamics and variety-seeking behavior respectively. This dissertation provides examples of incorporating important latent behavioral factors (besides intrinsic preferences), such as possible social influence and interactions or interactions across consumption occasions, into account when inferring individual preferences. Specifically, in Chapter 2, we develop a new method to infer heterogeneous individual preferences and state dependence during joint consumption, with consideration of three potentially confounded components of intragroup dynamics (including preference revision, behavioral interaction and decision power). In Chapter 3, we propose an innovative method to infer heterogeneous individual preferences and consumption order, with consideration of variety-seeking behavior as well as the effect of the consumption outcomes. In both chapters, we show that ignoring important latent behavioral factors, such as intragroup dynamics or variety-seeking behavior, could lead to biased estimation on individual preferences and have several consequences. With an application to the TV viewing and targeted TV advertising, we demonstrate that our model in Chapter 2 (Essay 1) could significantly improve the efficiency of TV ad targeting in marketing practice. Also, in Chapter 3, in a context of video-game playing, we find that positive consumption outcomes lead to inertial preferences, while negative consumption outcomes lead to variety-seeking.

One of the novel aspects of this dissertation is that we allow incomplete information in the estimation process. In Chapter 2, we allow missing information on individual's behavior and identity (where we observe the total consumption for a group of customers, but we do not know how many and which of them consume which product); whereas, in Chapter 3, we tackle the challenge arising from missing information on the order of the consumption (where we observe the consumption occasions in a period, but we do not observe their order). With incomplete data, we use two-step iterative MCMC processes to estimate the models: a forward-backward algorithm in Chapter 2 and a switching algorithm in Chapter 3 to augment and impute missing information. We then estimate the individual-level Bayesian models with augmented data. In both chapters, we use simulation studies to validate the proposed methods, identification strategies, and estimation algorithms.

Our work in this dissertation sheds light on a variety of marketing contexts. First of all, both chapters focus on better inferring heterogeneous individual preferences. Gaining insightful information and understanding individual preferences properly are the premise of personalization, such as ad targeting and so on. With advances in new technology, it becomes easier to collect relatively rich data compared to the past. The data collected (even for big data), however, often lack important and useful information. To remove the barrier and overcome the challenges of underpinning new waves of marketing innovation, there is a need to develop advanced methods and algorithms to infer heterogeneous individual preferences with incomplete information. For example, the model proposed in Chapter 2 enables marketers to target a much narrowly defined customer more efficiently when only aggregate and de-identified data are observed, including household TV viewing and targeted TV advertising discussed in this chapter and other applications such as retailing, membership enrollment, and so on. In addition, the method proposed in Chapter 3 can be used to learn about each customer's latest preferences, taking into account his/her previous consumption outcomes and variety-seeking behavior, to improve the effectiveness of customized recommendation or promotion campaigns. In sum, this dissertation contributes to marketing practice by improving ways to infer individual preferences and consumption behavior, which enable marketers to better understand their customers and make customized marketing strategies correspondingly.

Our work in this dissertation can be extended in several aspects in the future. First, we may further take into account some other factors or behavioral patterns which may affect customers' consumption behavior. For example, there is an emerging type of behavior increasingly appearing during joint consumption, namely binging within a group behavior setting. Although previous research, mostly from a psychological standpoint, has been devoted to studying consumer's binge behavior, rather less attention has been paid to learn about whether and how group interaction affects consumers' binge behavior, as well as its implications to firms' marketing strategy. In the future, we may empirically document how group interaction during joint consumption of a product affects both consumer's binge behavior in the short run and consumption behavior in the long run. Taking into account binge behavior, we might be able to better infer individual preferences in some marketing contexts whether people may binge, such as online video watching, online video games, and so on.

Second, we may look into more ways to incorporate incomplete data. Although firms have more and more data, there are some areas that are difficult to collect data. For example, the offline consumption behavior within a household. It is usually difficult to have information about which member(s) in a household consumes, for example, the yogurt purchased. Even with online behavior, there are more and more privacy regulations which restrict tracking consumers' online behavior. Simply put rich data does not mean rich information. In the future, we may need more work on better extracting information either from incomplete data or from data with poor information (even though the data are complete).

In sum, in this dissertation, we examine how to better infer heterogeneous individual preferences and important latent behavioral behavior, especially when the information is incomplete. There are some aspects which we may extend our work in this dissertation in the future. We hope our work in this dissertation provides the building blocks of future research on inferring individual preferences as well as relevant marketing strategies.

Appendix

Appendix A. Model Estimation of Chapter 2

Following the discussion in Section 2.3.1, we define the missing data (Y_{ht}^{mis}) as the difference between the complete information when everything was observed (Y_{ht}^{comp}) and observed information for household h at time t (Y_{ht}^{obs}). That is, $Y_{ht}^{comp} = Y_{ht}^{mis} \bigoplus Y_{ht}^{obs}$ where operator $\bigoplus$ means the combination of two information sources. For example, the combination of a (missing) information that "member A consumes alone", with an observed information that "$Y_{ht}^{obs} = g$" will lead to a complete information that $\left\{Y_{ht}^{A} = g, Y_{ht}^{B} = 0\right\}$. Note that, we need Y_{ht}^{comp} to estimate the individual-level model. To do so, we first augment the missing information (Y_{ht}^{mis}) and then estimate the model using augmented information. Specifically, we iteratively draw samples of parameter set Ψ and missing information Y_{ht}^{mis} in the following two steps, which is similar to a Gibbs sampling framework:

Step 1 (data augmentation): Generate a sample of Y_{ht}^{mis}(for each household h and time t) based on $\left\{Y_{h}^{obs}, \Psi_i\right\}$ where Ψ_i is a draw of parameters in current iteration, i.e., sampling $p\left(Y_{ht}^{mis}|Y_{h}^{obs}, \Psi_i\right)$.

Step 2 (parameter estimation): Generate a sample of Ψ_{i+1} based on $\left\{Y^{mis}, Y^{obs}\right\}$ using the augmented individual-level data, i.e., sampling $p\left(\Psi_{i+1}|Y^{mis}, Y^{obs}\right)$.

In the description above, Y_{h}^{obs}(where we omit subscript t) stands for observed information of household h in all periods. Similarly, $Y^{obs}(Y^{mis})$ refers to observed information (augmented missing information) for all households and all periods. Note that in our setting, we need to sample $p\left(Y_{ht}^{mis}|Y_h^{obs},\Psi_i\right)$, where the distribution of the missing individual-level information of household h at time $t(Y_{ht}^{mis})$ depends on observed information of household h in all periods (Y_h^{obs}), instead of observed information of household h at time t only (Y_{ht}^{obs}). Specifically, a member's consumption choice at time t depends on his or her consumption choice in the previous period $t-1$, while her or his consumption choice at time $t-1$ further depends on that at time $t-2$, and so on. That is, the missing individual-level information of household h at time $t(Y_{ht}^{mis})$ is correlated across time, which makes it challenging to implement Step 1 described above. Moreover, since we estimate the model iteratively using a Gibbs sampling framework, in each iteration, it is likely that Ψ_i highly depends on augmented Y_{ht}^{mis} while augmented Y_{ht}^{mis} in the next iteration will highly and circularly depend on Ψ_i in the current iteration. Because of the uncertainty of estimated values, a small deviation from augmented Y_{ht}^{mis} to Y_{ht}^{true} may affect the estimation significantly. To overcome these challenges, we propose a forward-backward algorithm to simultaneously incorporate the dependence of time-correlated missing information and the uncertainty of estimated values, which significantly facilitates our data augmentation process and improves the speed of convergence in the estimation. We discuss data augmentation in Step 1 and parameter estimation in Step 2 above respectively in the following subsections.

A.1 Step 1: Data Augmentation: A Forward-Backward Algorithm

Building on the proposed joint consumption model, we discuss how we estimate and identify our model using aggregate and de-identified data. We start our discussion with a data augmentation process (Step 1) to impute missing individual-level data for household h. For the purpose of notation simplicity, we omit subscript h in this discussion. As we discussed, the dependence of time-correlated missing individual-level information significantly increases the difficulty of implementing data augmentation for this model.

To handle the dependence of time-correlated missing information, we propose a general forward-backward algorithm. First, since Y_t^{mis} depends on the missing information in all previous periods, $Y_{t'<t}^{mis}$, we need to sequentially augment missing information from the beginning to the end (from time $t = 1$ to $t = T$). Therefore, we have a "forward" process where we augment forwardly based on information in all previous periods. Using this forward process, we utilize all information before time t to sample Y_t^{mis}, which can be seen as an approximation of sampling $p\left(Y_t^{mis}|Y_{1:t}^{obs},\Psi\right)$ where $Y_{1:t}^{obs}$ refers to observed information by time t. Second, observed information after time t ($Y_{t+1:T}^{obs}$) depends on Y_t^{mis} and thus presumably it also should be taken into account to update the posterior of Y_t^{mis}. We achieve this goal by a backward process where, after the forward process, we reverse augmentation of Y_t^{mis} from the end of periods ($t = T$) to the beginning of periods ($t = 1$).

The idea behind the process described above is that sampling of $p\left(Y_t^{mis}|Y^{obs},\Psi\right)$ will be significantly simplified if complete information at time $t-1$ and $t+1$ is available. Because,

$$\begin{aligned} p\left(Y_t^{mis}|Y^{obs},\Psi\right) &\propto p\left(Y_t^{mis}|Y_{t-1}^{comp},Y_t^{obs},Y_{t+1}^{comp},\Psi\right) \\ &\propto p\left(Y_{t-1}^{comp}|Y_t^{mis},Y_t^{obs},\Psi\right)\cdot p\left(Y_{t+1}^{comp}|Y_t^{mis},Y_t^{obs},\Psi\right)\cdot p\left(Y_t^{mis}\right) \quad (A.1-1) \\ &\propto p\left(Y_t^{mis}|Y_{t-1}^{comp},Y_t^{obs},\Psi\right)\cdot p\left(Y_{t+1}^{comp}|Y_t^{mis},Y_t^{obs},\Psi\right) \end{aligned}$$

where $p\left(Y_t^{mis}|Y_{t-1}^{comp},Y_t^{obs},\Psi\right)$ refers to the forward step that we use (approximated) complete information in the previous period $(t-1)$ while $p\left(Y_{t+1}^{comp}|Y_t^{mis},Y_t^{obs},\Psi\right)$ refers to the backward step that we use (approximated) complete information in the next period $(t+1)$.

Specifically, the forward-backward algorithm is implemented as follows:

Algorithm A.1 A Forward-Backward Algorithm

\# Forward process

for $t = 1, ..., T$

(i) augment Y_t^{miss} using $p\left(Y_t^{miss}|Y_{t-1}^{comp}, Y_t^{obs}, \Psi\right)$, denoted as $Y_{t,forward}^{miss}$

(ii) update $Y_t^{comp} = Y_{t,forward}^{miss} \bigoplus Y_t^{obs}$

end;

\# Backward process

for $t = T-1, ..., 1$

(i) augment Y_t^{miss} using $p\left(Y_t^{miss}|Y_{t-1}^{comp}, Y_t^{obs}, \Psi\right) \cdot p\left(Y_{t+1}^{comp}|Y_t^{miss}, Y_t^{obs}, \Psi\right)$, denoted as $Y_{t,backward}^{miss}$

(ii) update $Y_t^{comp} = Y_{t,backward}^{miss} \bigoplus Y_t^{obs}$

end;

A.2 Step 2: Parameter Estimation

After the imputation step (Revised Step 1) using data augmentation, we obtain individual-level data with augmented Y_{ht}^{mis} , which we use for parameter estimation, i.e., sampling $p\left(\Psi_{i+1}|Y^{mis}, Y^{obs}\right)$. In particular, we use a Bayesian hybrid Metropolis-Gibbs sampling framework for parameter estimation, where parameters are drawn individually conditional on others ("Gibbs"). while Metropolis is used to sample individual parameters where it is infeasible to directly sample from the posterior. The framework begins with a step to generate latent preferences z_{hg}^A and z_{hg}^B using an augmentation process. Specifically, conditional on all other parameters, for each period t, we augment $z_{hgt}^A = z_{hg}^A + \varepsilon_{hgt}^A$ and $z_{hgt}^B = z_{hg}^B + \varepsilon_{hgt}^B$ such that, following the group decision mechanism, observed choice$\left\{Y_{ht}^A = g^A, Y_{ht}^B = g^B\right\}$ would be chosen by the members. Then, we draw z_{hg}^A and z_{hg}^B based on augmented z_{hgt}^A and z_{hgt}^B.

After drawing latent preferences, we then use equations (5) and (6) to estimate δ_h^A, δ_h^B,α_h^{BA}, α_h^{AB}, β and Σ_z. Specifically, equations (5) and (6) form a simultaneous equation model which can be rewritten as the follows:

$$\begin{bmatrix} z_{hg}^{A} & z_{hg}^{B} \end{bmatrix} = \begin{bmatrix} \delta_h^A \; \delta_h^B \end{bmatrix} + \begin{bmatrix} X_{hg}^{A} & X_{hg}^{B} \end{bmatrix} \begin{bmatrix} \beta_g^A & 0 \\ 0 & \beta_g^B \end{bmatrix} \left(W^{-1}\right) + \begin{bmatrix} u_{hg}^{zA0} & u_{hg}^{zB0} \end{bmatrix} \quad (A.2-1)$$

where $W = \begin{bmatrix} 1+\alpha^{BA} & -\alpha_h^{AB} \\ -\alpha^{BA} & 1+\alpha_h^{AB} \end{bmatrix}$ and $\left(u_{hg}^{zA0}, u_{hg}^{zB0}\right) \sim MVN\left[0, \left(W^{-1}\right)' \Sigma_z \left(W^{-1}\right)\right]$.

We estimate the above equation using a 3-step process:

Step 1: Conditional on δ_h^A, δ_h^B, α_h^{BA},α_h^{AB}, and Σ_z, we draw from the posterior distribution of β. Note that after using the decomposition step described above, the equations (5) and (6) become a seemingly uncorrelated regression (SUR) model which is conditional on δ_h^A, δ_h^B, α_h^{BA}, α_h^{AB}, Σ_z .

Step 2: We numerically draw α_h^{BA} and α_h^{AB} conditional on β and Σ_z. Specifically, for each $\left(\alpha_h^{BA}, \alpha_h^{AB}\right)$, we calculate $\left(u_{hg}^{zA0}, u_{hg}^{zB0}\right)$ after the decomposition step described above and obtain the likelihood of $\left(u_{hg}^{zA0}, u_{hg}^{zB0}\right) \sim MVN\left[0, \left(W^{-1}\right)' \Sigma_z \left(W^{-1}\right)\right]$.

Step 3: We update δ_h^A, δ_h^B, Σ_z conditional on $\alpha_h^{BA}, \alpha_h^{AB}$ and β.

After that, Metropolis-Hastings steps are used to update other parameters including behavioral interaction (θ_{hg}^A, θ_{hg}^B), state dependence (τ_{hg}^A, τ_{hg}^B), sensitivity to time-varying factors (λ_h^A, λ_h^B), as well as decision power (γ_h).

Finally, we estimate the heterogeneity structures on intragroup dynamics and state-dependence across households using regressions in the respective equations.

Appendix B. Identification of Chapter 2

In this appendix, we provide simulated examples to show how variation across time helps us disentangle the three group dynamics as well as preference heterogeneity. Note that, with single-member households, we would be able to get expected preferences for certain demographics. For example, the expected preference on drama for 40-year old male and so on. For any two-member households, we then can have their expected preferences based on their demographics. We can run a simulation to see what we expect the viewing behavior would be based on their preferences (assuming no group dynamics). We call this as a "base scenario". As mentioned in Section 2.3, four factors may cause difference on observed viewing behavior of the two-member household from the base scenario: any one or more of the three group dynamics and/or preference heterogeneity. To identify these four factors and know what causes the difference between observed household behavior and expected behavior (i.e. base scenario), we would need variation to see that these four factors would have different effects across different cases.

Specifically, we simulate a two-member household. We assume there are three choices (g1, g2, g3) and there is a time-varying factor for each choice (Q1, Q2, Q3). We then simulate 10 cases of these time-varying factors: case 0 is a base case where Q1, Q2, Q3 are all 0; case 1 to case 3 to have different values for Q1 respectively; case 4 to case 6 have different values for Q2 respectively; and case 7 to case 9 we have different values for Q3 respectively. We then simulate the viewing behavior at household level for the following five scenarios: (1) a base scenario without any group dynamics and preference heterogeneity; (2) with preference revision; (3) with behavior interaction; (4) with decision power; and (5) with preference heterogeneity. For each scenario, we simulate the household-level viewing behavior on the three choices under each of the 10 cases, then we compare and look into the difference between each of scenarios 2 to 5 and the base scenario, scenario 1, in the viewing behavior of the three choices.

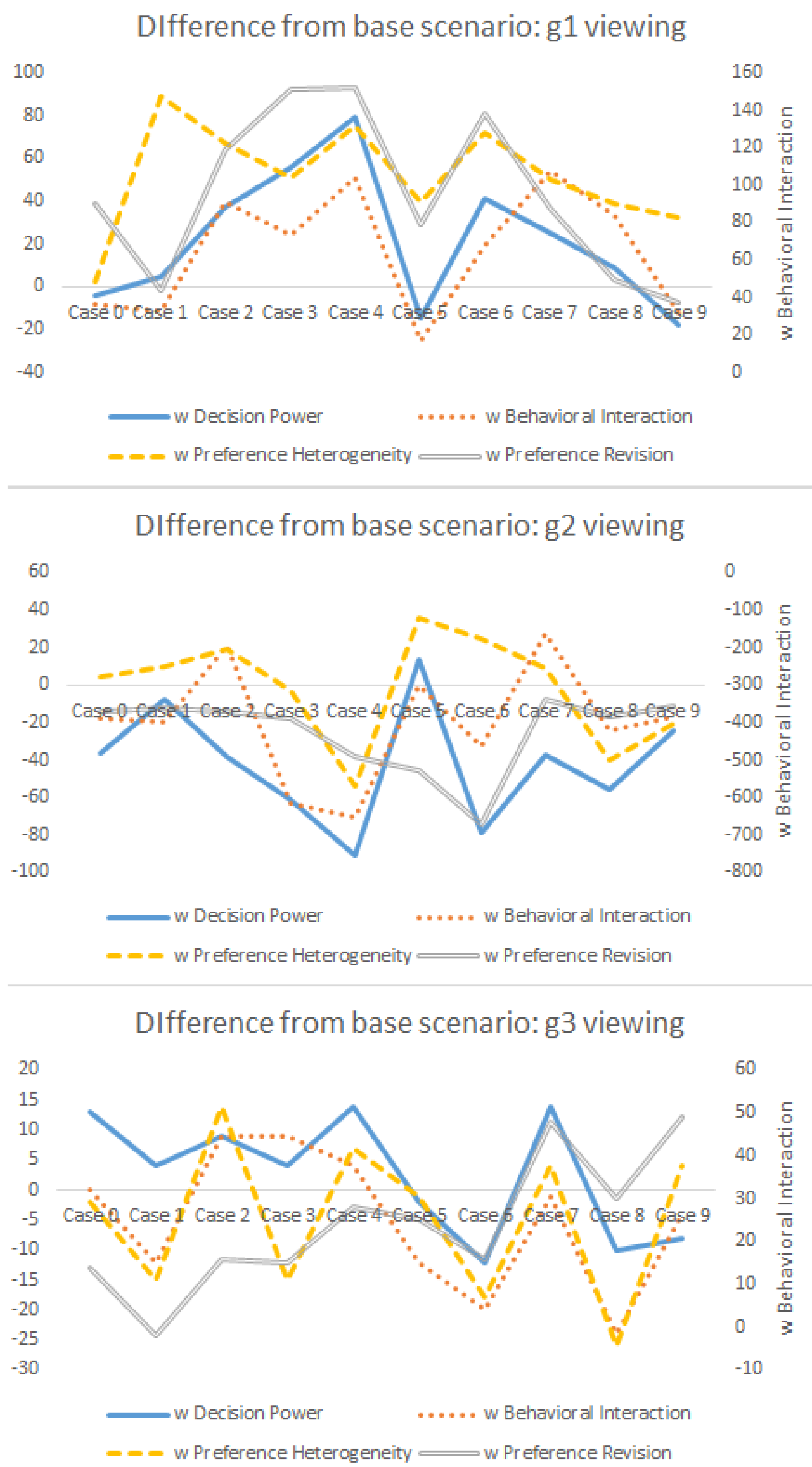

Figure B.1: Comparison of Viewing Behavior in Four Simulated Scenarios

Figure B.1 shows that difference between viewing behavior of the four scenarios with the base scenario on three choices respectively. For example, the solid line in the graph on the left shows the difference between the scenario with decision power and the base scenario on g1 viewing in the ten different cases. Note that, the variation across cases (which is the variation across time-varying factors) give us variation to identify and separate the four scenarios. For example, if there is a preference heterogeneity, we would expect a big difference in g1 viewing in case 1 while if it is caused by other reasons, we would expect a small difference in g1 viewing in case 1. Another example is that, if there is preference revision, we would expect a big difference in case 6 for g2 viewing and the difference increases from case 3 to case 6. As shown in Figure B.1, the four scenarios (the three group dynamics and preference heterogeneity) have a different impact on viewing behavior under these ten cases, which provides us information to disentangle these four factors.

Appendix C. Estimation Results of Chapter 2

In this appendix, we present the estimation results of the full model. First, we review the setting we have. We focus on the top 6 genres, and will consider all others as a genre called "others". We have 9 covariates, including household income, size of geographic location (whether it is in a big city), age, gender, working hours per week, education level (whether it is college or above), time spent on internet, whether is a household head, and whether the household consist of a married couple. Note that, (1) the first three covariates and the last one are at the household level, while the others are at the individual level; (2) the last two covariates are for two-member households only. Accordingly, in the estimation of ζ_γ we do not have coefficients for household level covariates since the regression estimation for γ is based on the difference between two members. Household level covariates are the same for the two members in the same household, so they are excluded from the regression estimation for γ. For the last two covariates which are for two-member households only but not single-member households, we do not have regression coefficients for β, ζ_τ and ζ_λ since these regressions are based on both two-member and single-member households.

C.1 Estimation Results of Parameters

Table C.1: Estimation Results of the Full Model of Chapter 2: β

β	CS	DN	GD	GV	PV	SE	Os
int	-1.66 (0.65)	-1.33 (0.68)	-1.82 (0.54)	-1.54 (0.51)	-1.33 (0.46)	-2.51 (0.67)	-1.15 (0.46)
income	-0.14 (0.05)	-0.11 (0.05)	-0.15 (0.03)	-0.09 (0.05)	-0.07 (0.04)	-0.03 (0.03)	-0.1 (0.05)
size	0.02 (0.08)	0.02 (0.08)	0.02 (0.06)	0 (0.06)	0.05 (0.06)	0.01 (0.08)	-0.03 (0.07)
age	-0.02 (0.16)	-0.08 (0.15)	0.18 (0.13)	-0.08 (0.12)	-0.04 (0.11)	0.07 (0.15)	-0.19 (0.12)
gender	0 (0.08)	-0.13 (0.11)	-0.09 (0.09)	-0.06 (0.09)	-0.17 (0.08)	0.31 (0.09)	-0.04 (0.1)
work	0.2 (0.1)	0.02 (0.09)	0.09 (0.07)	0.09 (0.08)	0.11 (0.07)	0.13 (0.09)	0.08 (0.07)
edu	-0.18 (0.09)	-0.27 (0.07)	-0.1 (0.07)	-0.32 (0.07)	-0.26 (0.07)	-0.03 (0.07)	-0.26 (0.09)
internet	-0.03 (0.03)	-0.05 (0.03)	-0.02 (0.02)	-0.06 (0.03)	-0.01 (0.03)	-0.03 (0.03)	-0.03 (0.04)

Table C.1 shows the estimation results for β, which describes how demographics affect the base preferences. As shown, higher income, higher education level and lower working hours are associated with less time on TV. Males watch more sports while females watch more participation variety (PV).

Table C.2: Estimation Results of the Full Model of Chapter 2: ζ_τ

ζ_τ	CS	DN	GD	GV	PV	SE	Os
int	4.22 (1.04)	2.33 (0.97)	2.78 (0.63)	2.61 (0.88)	2.18 (0.68)	2.68 (0.96)	3.07 (1.07)
income	0.12 (0.09)	0.19 (0.06)	0.09 (0.05)	0.17 (0.06)	0.08 (0.05)	0.06 (0.07)	0.03 (0.07)
size	-0.2 (0.1)	0.05 (0.14)	0.01 (0.09)	-0.01 (0.14)	-0.01 (0.1)	0.01 (0.11)	-0.02 (0.23)
age	-0.74 (0.27)	-0.19 (0.22)	-0.39 (0.14)	-0.28 (0.21)	-0.21 (0.16)	-0.18 (0.22)	-0.24 (0.24)
gender	0.1 (0.1)	0.03 (0.1)	0.09 (0.07)	-0.06 (0.12)	0.06 (0.07)	0.1 (0.11)	0.02 (0.12)
work	-0.19 (0.17)	-0.13 (0.12)	-0.04 (0.08)	-0.17 (0.13)	0.06 (0.09)	0.15 (0.13)	-0.05 (0.1)
edu	0.01 (0.14)	-0.14 (0.14)	-0.04 (0.09)	0.07 (0.15)	0.02 (0.09)	-0.16 (0.12)	0.11 (0.1)
internet	0.03 (0.05)	0.08 (0.05)	0.02 (0.03)	0 (0.04)	0.02 (0.03)	0.04 (0.05)	-0.03 (0.04)

Table C.2 shows the estimation results for ζ_τ, which describes how demographics affect state dependence. As shown, higher income and younger people are associated with higher state dependence.

Table C.3: Estimation Results of the Full Model of Chapter 2: ζ_θ

ζ_θ	CS	DN	GD	GV	PV	SE	Os
int	0.2 (0.83)	-0.61 (0.75)	-0.53 (0.85)	-0.03 (0.74)	-0.58 (0.65)	-1.37 (0.72)	-0.92 (0.77)
income	-0.06 (0.1)	0.02 (0.07)	0.02 (0.06)	-0.21 (0.08)	-0.01 (0.08)	0.24 (0.08)	-0.04 (0.06)
size	0.03 (0.11)	0.29 (0.12)	0.04 (0.09)	-0.02 (0.1)	0.07 (0.09)	-0.02 (0.12)	0.16 (0.12)
age	0.09 (0.21)	0.3 (0.16)	0.31 (0.2)	0.36 (0.17)	0.33 (0.15)	0.27 (0.16)	0.44 (0.19)
gender	0.2 (0.1)	0.14 (0.09)	-0.08 (0.09)	0.03 (0.09)	0.05 (0.09)	0.1 (0.09)	0.14 (0.09)
work	0.1 (0.14)	0.24 (0.14)	0.15 (0.11)	0.03 (0.09)	0.07 (0.1)	-0.06 (0.1)	0.06 (0.08)
edu5	0.17 (0.19)	0.24 (0.16)	-0.01 (0.09)	0.13 (0.13)	0.16 (0.08)	-0.17 (0.16)	0.2 (0.15)
internet	0.06 (0.06)	-0.02 (0.05)	-0.04 (0.05)	0.04 (0.04)	-0.02 (0.04)	-0.04 (0.04)	0.05 (0.05)
hh	0.21 (0.15)	-0.54 (0.11)	0.14 (0.1)	-0.19 (0.11)	0.07 (0.19)	0.09 (0.08)	-0.43 (0.15)
couple	-0.01 (0.11)	0.13 (0.1)	0.03 (0.08)	0.07 (0.1)	0.1 (0.09)	-0.1 (0.1)	-0.15 (0.1)

Table C.3 shows the estimation results for ζ_θ, which describes how demographics affect behavioral interaction. As shown, higher income is associated with less behavioral interaction in GV and higher behavioral interaction in Sports. People in large cities are associated with high behavioral interaction in News. Older people are associated with high behavioral interaction in most genres. Household heads are associated with less behavioral interaction in News.

Table C.4: Estimation Results of the Full Model of Chapter 2: ζ_α, ζ_γ, ζ_δ, ζ_λ

	ζ_α	ζ_γ	ζ_δ	ζ_λ
int	0.31 (0.62)	NA	-1.16 (0.35)	0.52 (0.55)
income	0.03 (0.04)	NA	0.05 (0.03)	-0.04 (0.03)
Size	-0.04 (0.06)	NA	-0.01 (0.06)	-0.03 (0.06)
Age	-0.03 (0.15)	0.02 (0.32)	0.08 (0.08)	0.13 (0.12)
gender	0.04 (0.07)	0.02 (0.07)	-0.02 (0.07)	-0.1 (0.06)
work	0.03 (0.06)	0.02 (0.12)	-0.1 (0.05)	0.03 (0.05)
edu	-0.02 (0.06)	0.03 (0.14)	0.06 (0.05)	0.08 (0.06)
internet	0 (0.02)	-0.03 (0.05)	0.03 (0.02)	0.01 (0.02)
hh	-0.02 (0.05)	0 (0.07)	0.19 (0.04)	NA
couple	-0.03 (0.07)	NA	0.06 (0.04)	NA

Table C.4 shows the estimation results for ζ_α, ζ_γ, ζ_δ, and ζ_λ, which describe how demographics affect the preference revision, decision power, preference shift and sensitivity to time-varying factors respectively.

C.2 Distribution of Heterogeneity across Households

We then look into the heterogeneity of individuals in two-member households in terms of their preferences (shown in Figure C.1 and Figure C.2), state dependence (Figure C.3 and Figure C.4), behavioral interaction (Figure C.5 and Figure C.6), as well as their preference revision, preference shift, decision power and sensitivity to time-varying factors (Figure C.7).

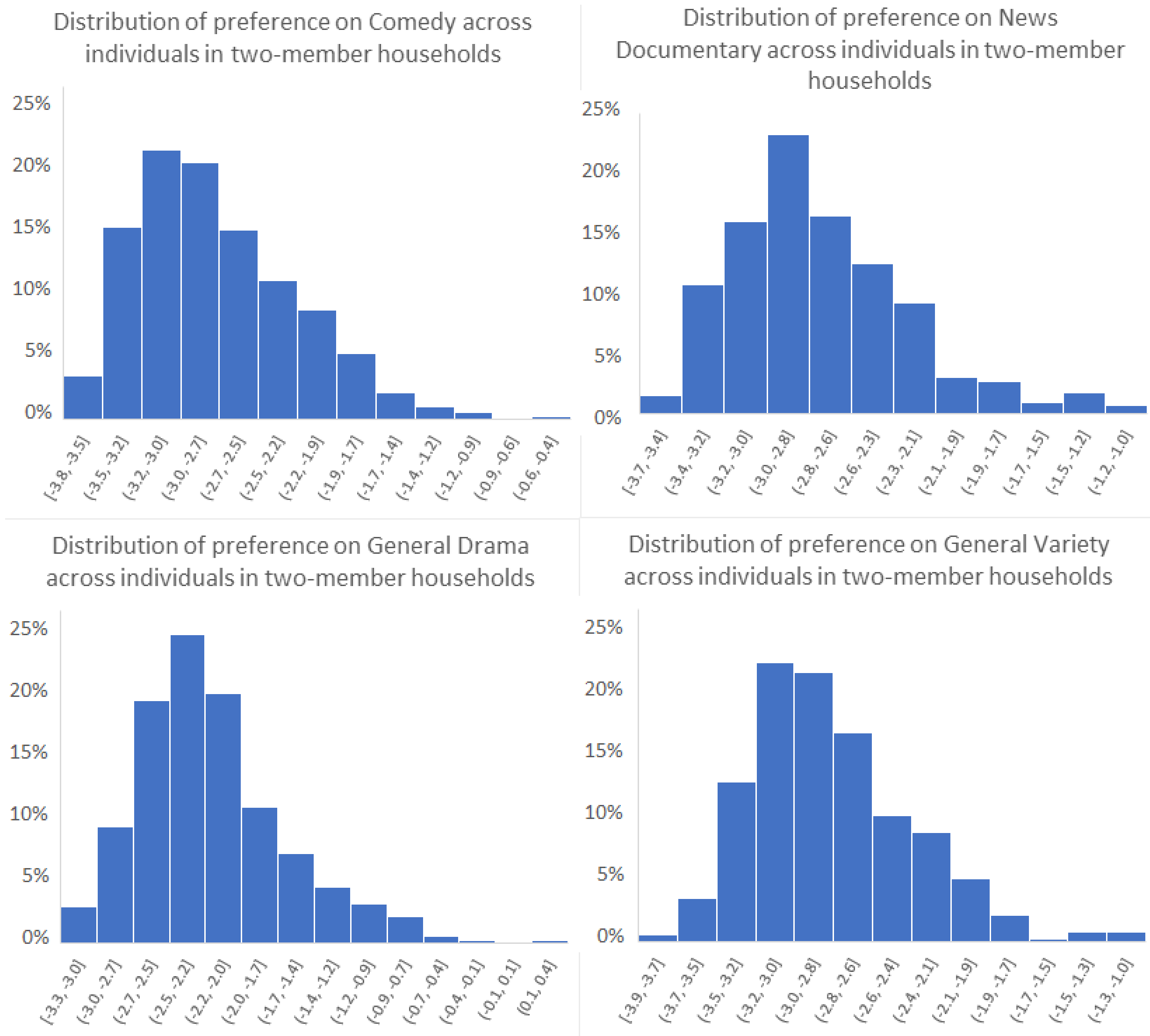

Figure C.1: Heterogeneity of Preference in Genres: Comedy, News Documentary, General Drama, and General Variety

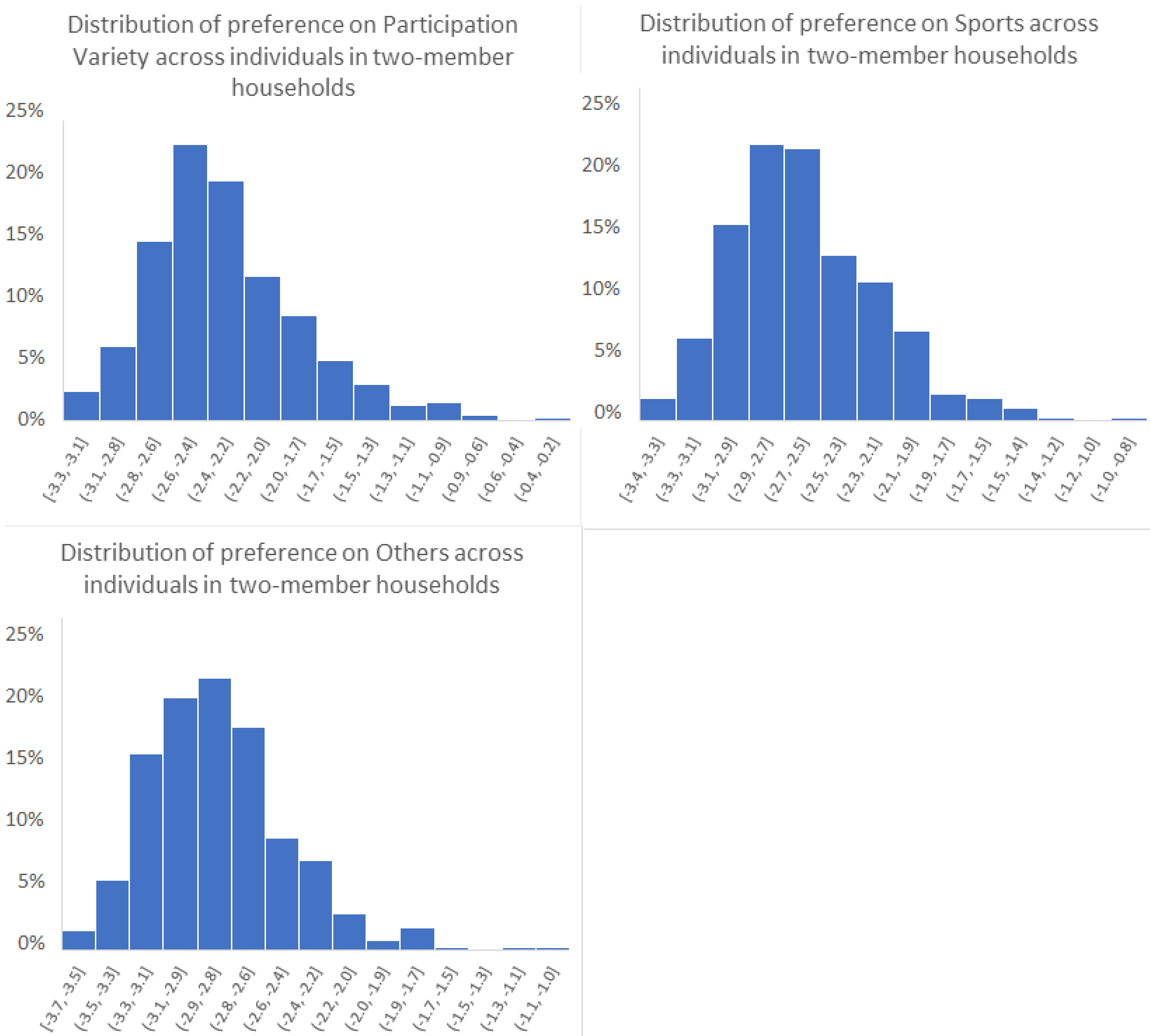

Figure C.2: Heterogeneity of Preference in Genres: Participation Variety, Sports, Others

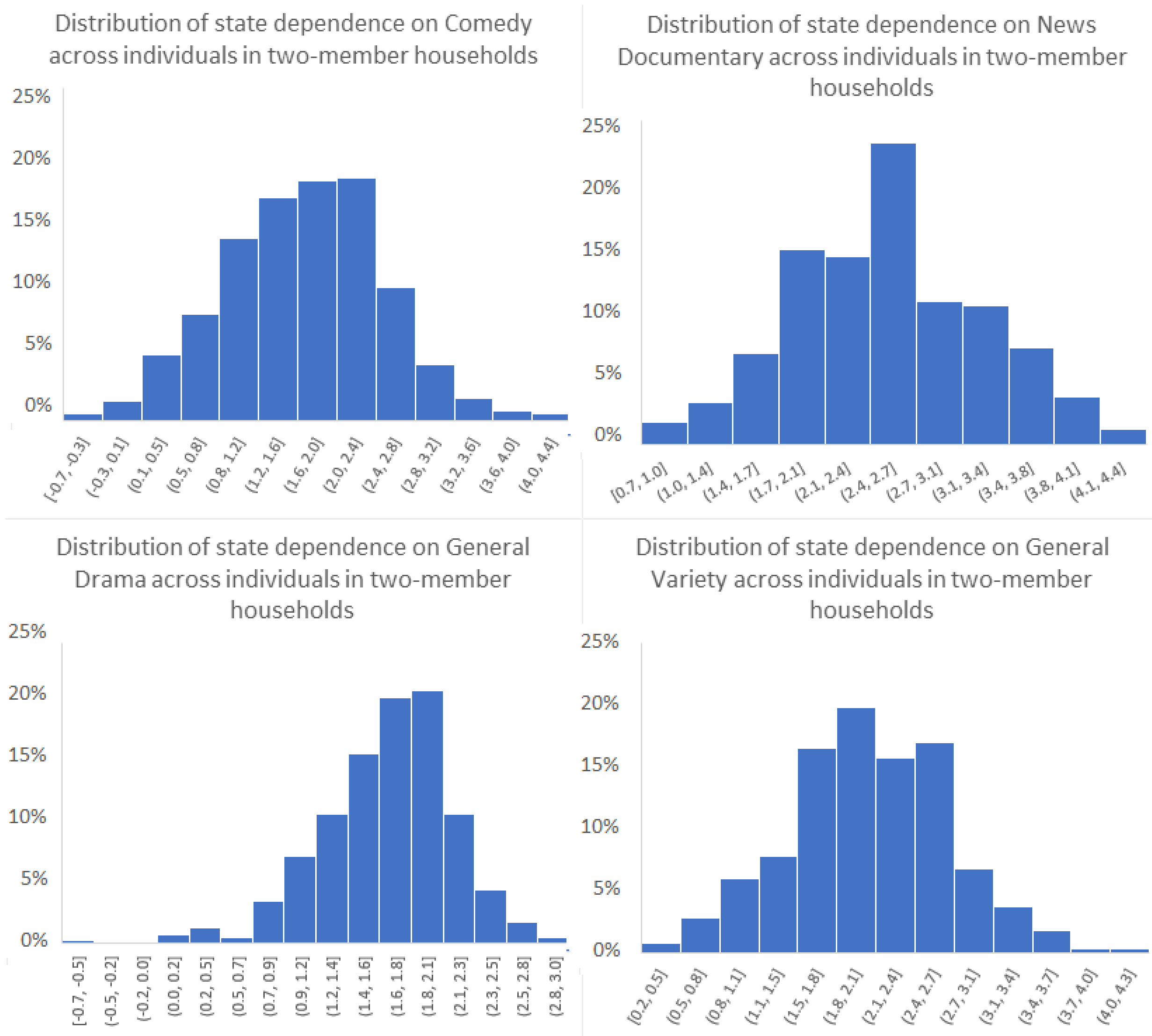

Figure C.3: Heterogeneity of State Dependence in Genres: Comedy, News Documentary, General Drama, and General Variety

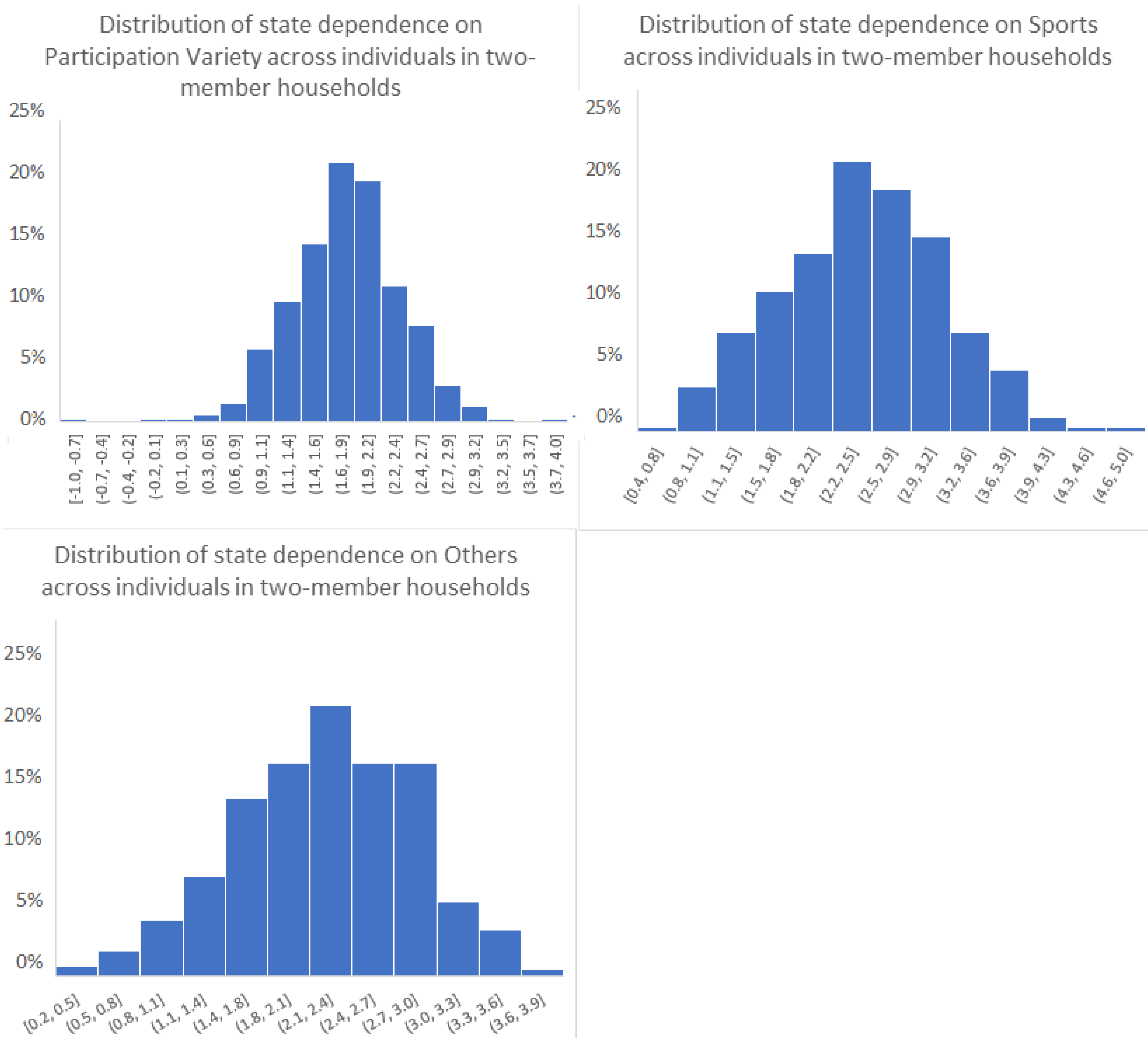

Figure C.4: Heterogeneity of State Dependence in Genres: Participation Variety, Sports, Others

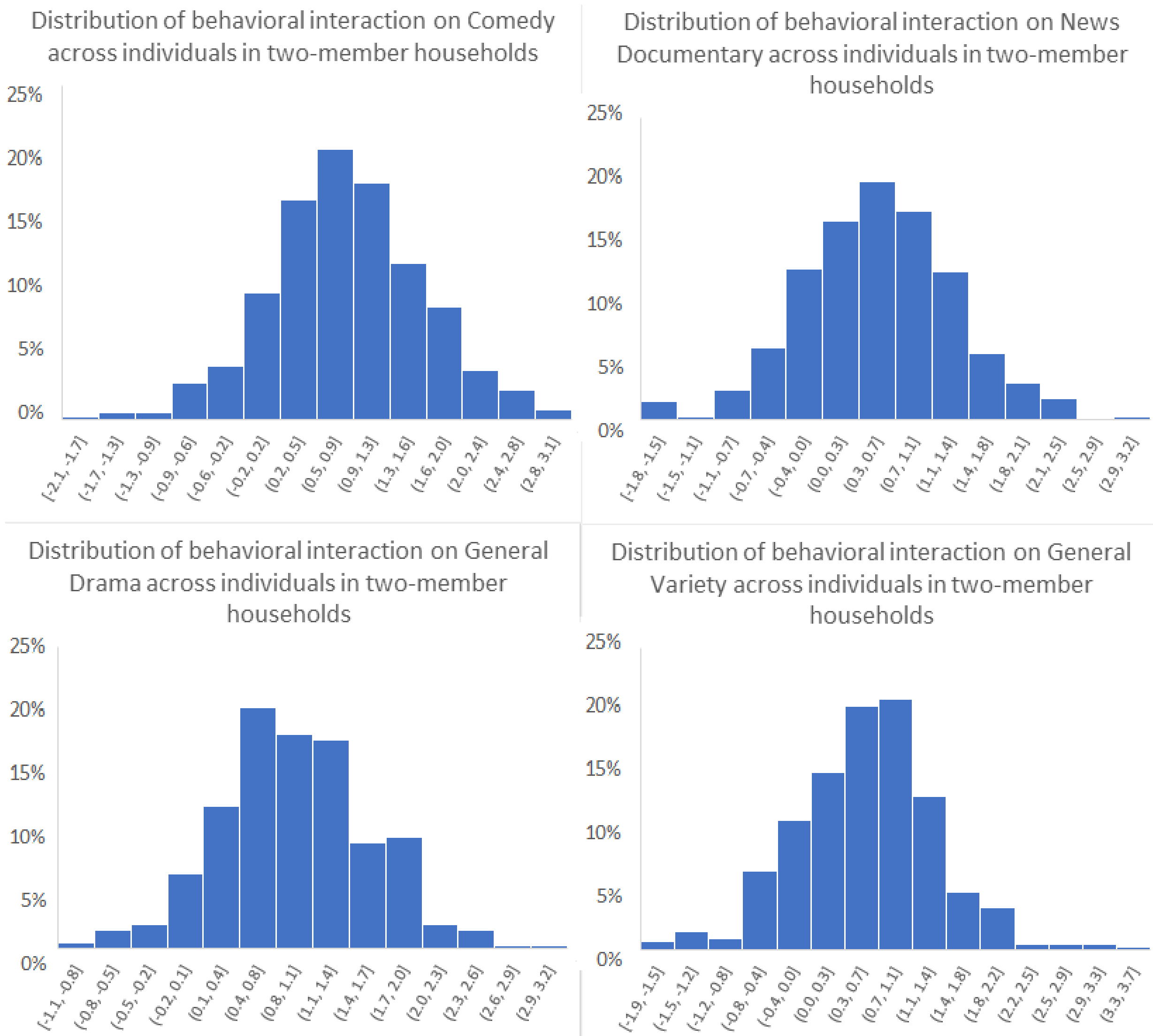

Figure C.5: Heterogeneity of Behavioral Interaction in Genres: Comedy, News Documentary, General Drama, and General Variety

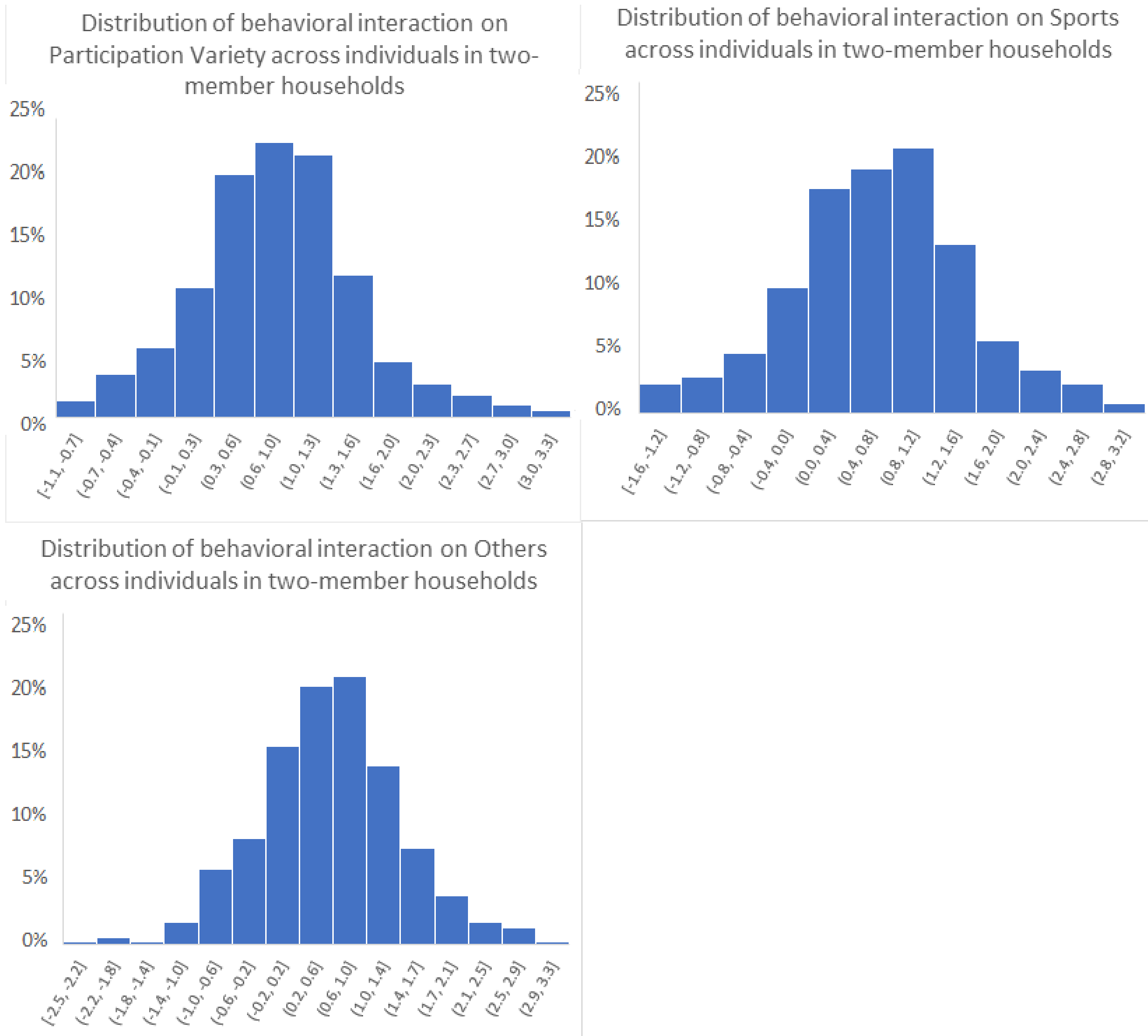

Figure C.6: Heterogeneity of Behavioral Interaction in Genres: Participation Variety, Sports, Others

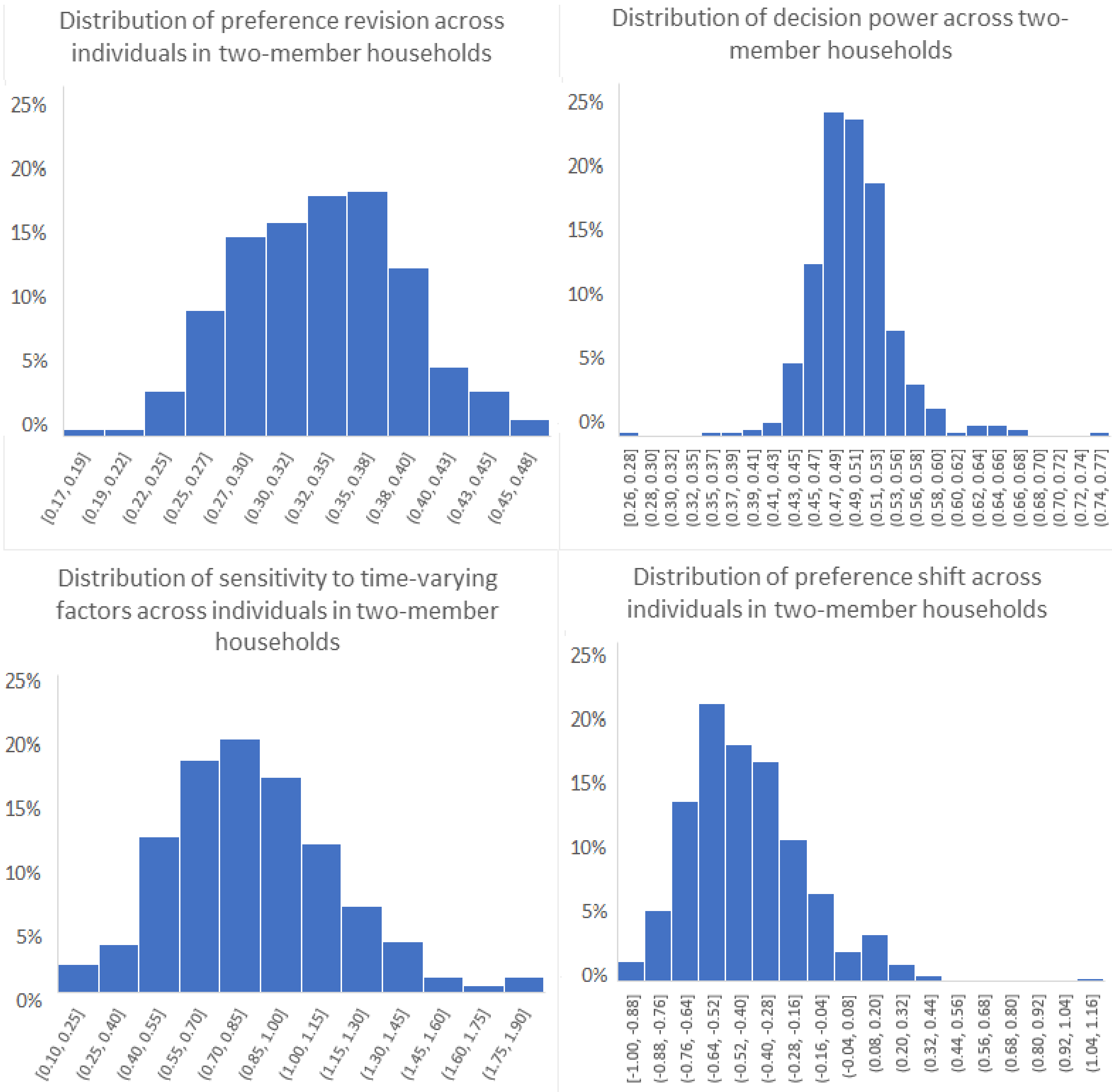

Figure C.7: Heterogeneity of Preference Revision, Decision Power, Sensitivity, and Preference Shift

C.3 Comparison with Model Estimation Using Individual-Level Data

We further compare the estimation results using aggregate data which are shown above with the estimation results of using individual-level data. If these two sets of results are close to each other, then it means we are able to get similar results using aggregate data only (instead of using individual level data). Figure C.8 compares these two sets of estimation results, where each point on the graph stands for one of interest parameter (for example, intercept term of β for genre 1); the horizontal axis describes estimation results using individual-level data while the vertical axis denotes estimation results using aggregate data. The dashed line shows where estimation using the two data sets generate the same results. As shown in Figure C.8, most parameters are around the dashed line, which means estimation results for most parameters are close to each other using these two data sets.

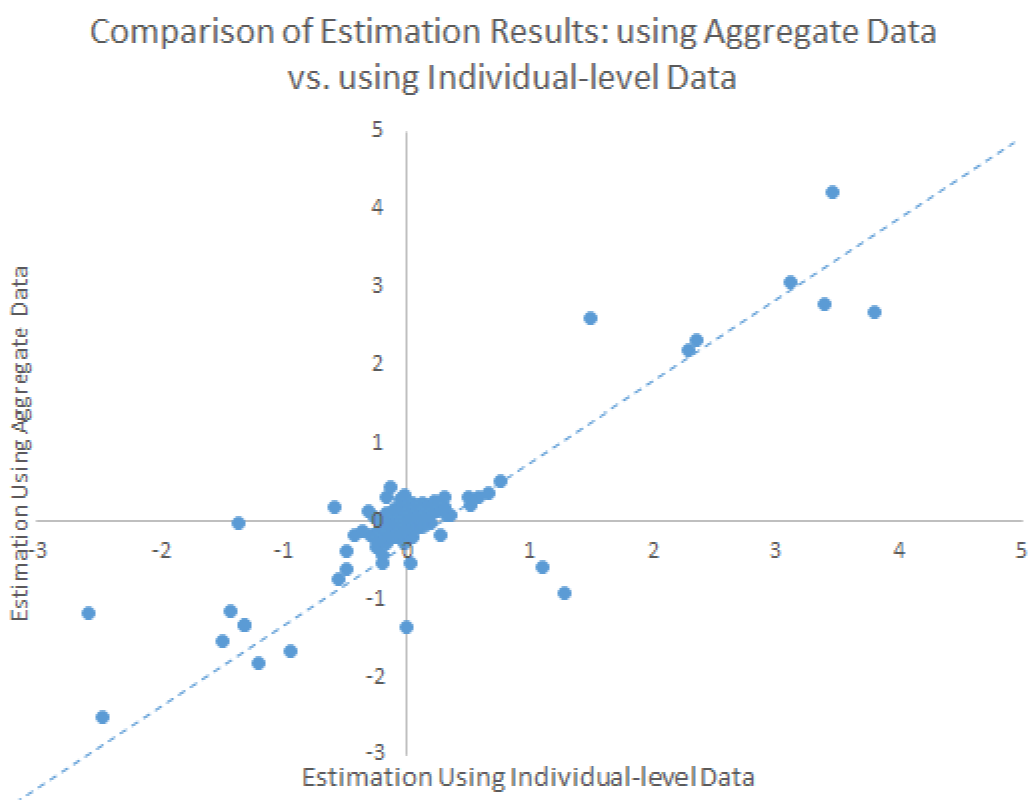

Figure C.8: Comparison of Estimation Results with Winner-Maximization Mechanism: Using Aggregate Data Versus Using Individual-level Data

Appendix D. Joint-Optimization Decision Mechanism of Chapter 2

In this appendix, we discuss another decision mechanism which we call "Joint Optimization." For the "Joint Optimization" mechanism, group members cooperate and work together to maximize the total group utility which is a weighted average of each group member's individual utility. This decision-making mechanism shares the same intuition of Samuelson's (1956) consensus model. According to Samuelson's consensus model, group members reach a consensus on maximizing a total welfare function of their individual utilities. It is acting like there is a hypothetical single agent (i.e., group head) who maximizes the consensus social welfare function of the entire group. It is also consistent with previous literature of collective (cooperative) models (e.g., Apps and Rees 1988; Browning and Chiappori 1998; Chiappori 1988, 1992), which assumes individuals make Pareto-efficient decisions resulting from an intragroup bargaining process (Browning and Chiappori 1998; Cherchye, De Rock, and Vermeulen 2012; Chiappori 1988; Zhang et al. 2009).

Following previous literature (e.g., Samuelson 1956, Kato and Matsumoto 2009), we assume that the two members follow the Harsanyi decision heuristic (Arora and Allenby 1999, Krisnamurthi 1988) for group decision making. In this case, the household chooses a choice to maximize the weighted total utility as follows:

$$U_{ht}\left(g^A, g^B\right) \quad = \quad \gamma_h \cdot U_{ht}^A\left(g^A, g^B\right) + (1-\gamma_h) \cdot U_{ht}^B\left(g^A, g^B\right) \tag{4.1}$$

where $U_{ht}\left(g^A, g^B\right)$ is the weighted total utility when member A consumes g^A and member B consumes g^B. γ_h refers to the weight of member A's utility.

Following the decision mechanism, the probability of household members to choose a choice of $\left\{y^A_{ht} = g^A, y^B_{ht} = g^B\right\}$ is

$$p\left\{\left(y^A_{ht} = g^A, y^B_{ht} = g^B\right)\right\} \quad = \quad p\left\{U_{ht}\left(g^A, g^B\right) \geq U_{ht}\left(\widetilde{g}^A, \widetilde{g}^B\right)\right\} \tag{4.2}$$

where $\left(\widetilde{g}^A, \widetilde{g}^B\right)$ refers to any other feasible consumption choices available to household members.

We apply the model to the same empirical setting in Section 2.4 with Joint-Optimization decision mechanism instead of Winner-Maximization.We further compare the estimation results using aggregate data with the estimation results of using individual-level data. Figure D.1 compares these two sets of estimation results, where each point on the graph stands for one interested parameter; the horizontal axis describes estimation results using individual-level data while the vertical axis denotes estimation results using aggregate data. The dashed line is a line to show where estimation using the two data sets generate the same results. As shown in Figure D.1, most parameters are around the dashed line, which means estimation results for most parameters are close to each other using these two data sets.

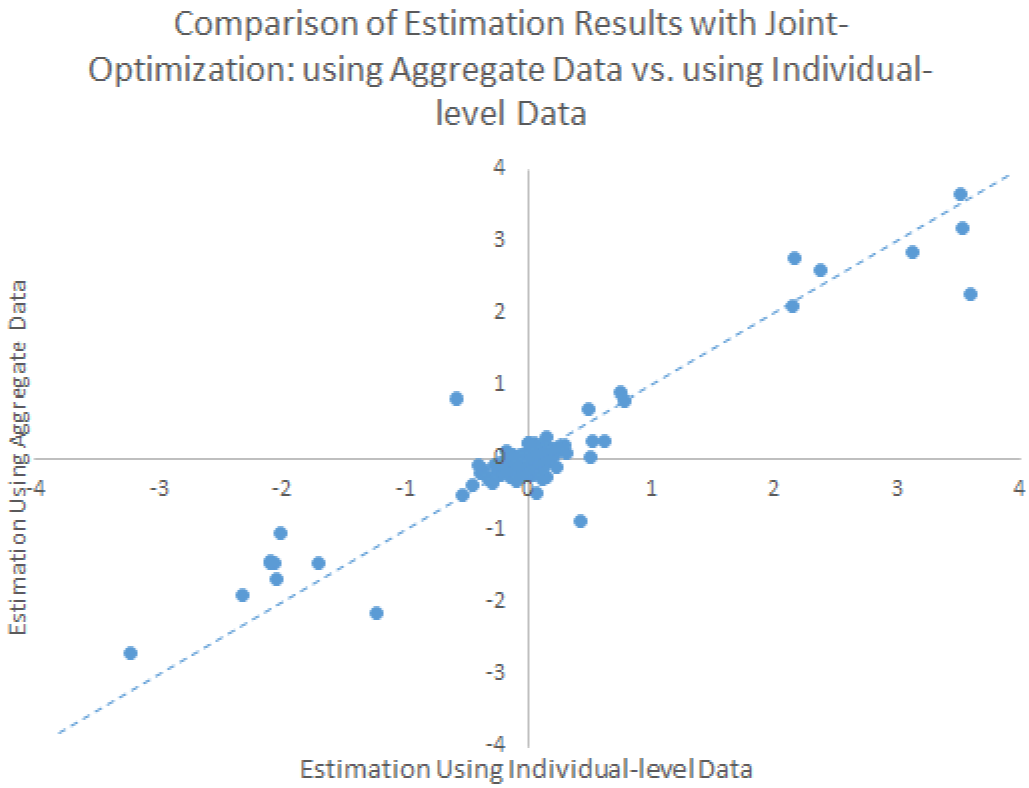

Figure D.1: Comparison of Estimation Results with Joint-Optimization Mechanism: Using Aggregate Data Versus Using Individual-level Data

Appendix E. Identification of Chapter 3

In this appendix, we use a numerical example to show the identification source of the proposed model. We simulate consumption data for a consumer in the following three scenarios: (1) the consumer is not variety-seeking and not inertial; (2) the consumer is variety-seeking, and (3) the consumer is inertial. Then we examine the variation we can observe from the non-ordered data to identify the variety-seeking/inertia and how consumption outcomes affect the variety-seeking/inertia. Specifically, we simulate the following: a consumer consumes three options (A, B and C) with one attribute. The distance on the attribute between A and B, B and C, A and C are 1, 1, 2 respectively. We further assume the following parameter values: $z_A = z_B = z_c = 0$, so the consumer equally prefers the three options; $\alpha = 0$, $\alpha = 2$, and $\alpha = -2$ for the three scenarios respectively, where a positive α stands for variety-seeking while a negative one stands for inertia; $\beta = 1$, so we assume that consumption outcomes are positively related with the variety-seeking behavior. Consumption outcomes are randomly generated from a standard normal distribution.

For each of the three scenarios, we simulate the consumption behavior of the consumer in 60 days; each day has five consumption occasions. The five consumption occasions within each day are non-ordered. For each scenario, we do the following: (1) summarize the number of times each option is consumed (which is from the non-ordered data). For example, we observe the consumer consumes A twice, B once, and C twice on day 1 in the first scenario; (2) for each option, we calculate the average number of times it is consumed each day across the 60 days, as well as the variance. We then compare the summary statistics from these three scenarios in the following table (for example, in scenario 1 where $\alpha = 0$, option A is consumed 1.87 times each day on average, with a variance of 1.03):

Table E.1: Mean and Variance of the Number of Times Consumed in Three Scenarios of Chapter 3

		A	B	C
$\alpha = 0$	mean	1.87	1.62	1.52
	variance	1.03	1.02	1.20
$\alpha = 2$	mean	2.20	0.60	2.20
	variance	0.37	0.45	0.26
$\alpha = -2$	mean	1.40	1.90	1.70
	variance	3.06	2.77	3.50

The above table shows that, the variance (of the number of times each option consumed across the 60 days) are different in the three scenario, which provides us the information to identify whether the consumer is variety-seeking or inertial. For example, option A has a variance of 1.03, 0.37 and 3.06 in the three scenarios respectively. The other two options have similar variances also. When $\alpha = 2$, the consumer is variety-seeking, the variance is smaller (so we observe that option A is consumed almost every day around 2 to 3 times); while when $\alpha = 3$, the consumer is inertial, the variance is much bigger (so we observe that option A is consumed five times on some days while not consumed at all on some other days). This is because, when the consumer is variety-seeking, she widely explores different options within each day; when the consumer is inertial, she usually focuses on one or two options each day (so it is more likely for us to observe that an option is consumed five times in some days). Therefore, the variation of the number of times that each option is consumed across days allow us to identify the variety-seeking/inertia of the consumer.

Then, we further examine how we can identify the effect of consumption outcomes on the variety-seeking/inertia. To do so, we take the third scenario, $\alpha = -2$, as an example. In this scenario, we illustrate how the consumption outcomes affect the non-ordered data observed. We follow the same way to calculate the mean and variance of the number of times that each option is consumed (as in the table above), but this time, we separate the 60 days into two groups according to the average consumption outcomes each day. Specifically, for each of the 60 days, we calculate the average consumption outcomes of the five consumption occasions. We then consider the 30 days with lower average consumption outcomes in a group, and the other 30 days in another group.

Table E.2: Mean and Variance of the Number of Times Consumed in Two Groups of Chapter 3

		A	*B*	*C*
lower consumption outcomes	mean	1.67	1.80	2.10
	variance	3.97	3.22	4.69
higher consumption outcomes	mean	1.67	1.77	1.38
	variance	2.46	2.17	2.16

As shown in the table above, the group with lower average consumption outcomes has a higher variance (which indicates more inertia) than the other group. So on days with low average consumption outcomes, it is more likely to observe a large variation on the number of times that each option is consumed. This provides us the information that the consumption outcomes is positively related to the variety-seeking (and negatively related to the inertia).

Bibliography

Aribarg A, Arora N, Bodur H O (2002) Understanding the role of preference revision and concession in group decisions. *Journal of Marketing Research*. 39(3): 336-349.

Arora N, Allenby G M (1999) Measuring the influence of individual preference structures in group decision making. *Journal of Marketing Research*. 36(4): 476-487.

Ashraf N (2009) Spousal control and intra-household decision making: an experimental study in the Phillippines. *American Economic Review*. 99(4): 1245-1277.

Berridge, K C (2012). From prediction error to incentive salience: mesolimbic computation of reward motivation. *European Journal of Neuroscience*, 35(7), 1124-1143.

Blundell R, Chiappori P A, Magnac T, Meghir C (2000) *Discrete choice and collective labor supply*. Mimeo, UCL.

Bourguignon F, Browning M, Chiappori P A (2009) Efficient intra-household allocations and distribution factors: implications and identification. *Review of Economic Studies*, 76: 503-528.

Browning M, Chiappori P A (1998) Efficient intra-household allocations: a general characterization and empirical test. *Econometrica*, 66: 1241–1278.

Browning M, Chiappori P A, Lechène V (2006) Collective and unitary models: a clarification. *Review of Economics of the Household*, 4(1):5-24.

Browning M, Chiappori P A, Lewbel A (2013) Estimating consumption economies of scale, adult equivalence scales, and household bargaining power. *The Review of Economic Studies*, 80 (4): 1267-1303.

Case A (1991) Spatial patterns in household demand. *Econometrica*, 59(4):953-965.

Chen Y, Yang S (2007) Estimating disaggregate models using aggregate data through augmentation of individual choice. *Journal of Marketing Research*. 44(4): 613-621.

Cherchye L, De Rock B, Vermeulen F (2012) Married with children: A collective labor supply model with detailed time use and intrahousehold expenditure information, *American Economic Review*, 102(7):3377-3405.

Chiappori P A (1988a) Rational Household Labor Supply. *Econometrica*, 56(1):63-89.

Chiappori P A (1988b) Nash-bargained Household Decisions: A Comment. *International Economic Review*, 29:791-796.

Chiappori P A (1992) Collective labor supply and welfare. *Journal of Political Economy*, 100:437-467.

Chiappori P A, Ekeland I (2009) The micro-economics of efficient group behavior: Identification. *Econometrica*, 77:763-794.

Chiappori P A, Fortin B, Lacroix G (2002) Household labor supply, the sharing rule and the marriage market. *Journal of Political Economy*, 110:3772.

Chintagunta, P K (1998). Inertia and variety seeking in a model of brand-purchase timing. *Marketing Science*, 17(3), 253-270.

Chu C (1989) A paired combinatorial logit model for travel demand analysis. *Proceedings of the Fifth World Conference on Transportation Research*, 4: 295-309.

Couprie H (2007) Time allocation within the family: welfare implication of life in couple. *Economic Journal*, 117:287-305.

Donni O, Chiappori P A (2011) *Nonunitary models of household behavior: A survey of the literature*. Household Economic Behaviors. Springer New York, 1-40.

Dube J P, Hitsch G, Rossi P (2010) State dependence and alternative explanations for consumer inertia. *The Rand Journal of Economics*, 41(3): 417-445.

Erdem T (1996) A dynamic analysis of market structure based on panel data. *Marketing Science*, 15(4), 359-378.

Erdem T, Keane M P (1996) Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science*, 15(1), 1-20.

Feit E, Wang P, Bradlow E, Fader P (2013) Fusing aggregate and disaggregate data with an application to multiplatform media consumption. *Journal of Marketing Research*, L(June):348-364.

Fong Y, Zhang J (2001) The identification of unobservable independent and spousal leisure. *Journal of Political Economy*, 109:191-202.

Gelman A, Rubin D B (1992) Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science,* 7: 457–511.

Givon M (1984) Variety seeking through brand switching. *Marketing Science*, 3(1), 1-22.

Guadagni P M, Little J D (1983) A logit model of brand choice calibrated on scanner data. *Marketing Science*, 2(3), 203-238.

Gupta S, Chintagunta P, Wittink D (1997) Household heterogeneity and state dependence in a model of purchase strings: empirical results and managerial implications. *International Journal of Research in Marketing*, 14: 341-357.

Hartmann WR (2010) Demand estimation with social interactions and the implications for targeted marketing. *Marketing Science*, 29 (4): 585-601.

Heckman J (1991) Identifying the hand of past: distinguishing state dependence from heterogeneity. *American Economic Review*, 81(2): 75-79.

Inman J J (2001) The role of sensory-specific satiety in attribute-level variety seeking. *Journal of Consumer Research*, 28(1), 105-120.

Jacoby J, Kyner D B (1973) Brand loyalty vs. repeat purchasing behavior. *Journal of Marketing Research*, 10(1), 1-9.

Jeuland A P (1979) Brand choice inertia as one aspect of the notion of brand loyalty. *Management Science*, 25(7), 671-682.

Johnson E (2004) Identification in discrete choice models with fixed effects, Working paper, *Bureau of Labor Statistics*.

Kahn B, Morrison D, Wright G (1986) Aggregating individual purchases to the household level. *Marketing Science*, 5(3): 260-268.

Kahn B E (1995) Consumer variety-seeking among goods and services: An integrative review. *Journal of Retailing and Consumer Services*, 2(3), 139-148.

Kahn B E, Isen A M (1993) The influence of positive aspect on variety seeking among safe, enjoyable products. *Journal of Consumer Research*, 257-270.

Kahn B E, Kalwani M U, Morrison D G (1986) Measuring variety-seeking and reinforcement behaviors using panel data. *Journal of Marketing Research*, 23(2), 89-100.

Kahn B E, Raju J S (1991) Effects of price promotions on variety-seeking and reinforcement behavior. *Marketing Science*, 10(4), 316-337.

Kato H, Matsumoto M (2009) Intra-household interaction in a nuclear family: A utility-maximizing approach. *Transportation Research*, B(43): 191–203.

Keltner D, Ellsworth P C, Edwards K (1993) Beyond simple pessimism: effects of sadness and anger on social perception. *Journal of Personality and Social Psychology*, 64(5), 740.

Kent R J, Schweidel D A (2011) Introducing the Ad ECG: how the set-top box tracks the lifeline of television. *Journal of Advertising Research*, 51 (4), 586-593.

Koppelman F, Wen C (2000) The paired combinatorial logit model: properties, estimation and application. *Transportation Research*, B(34): 75-89.

Lancaster G, Maitra P, Ray R (2006) Endogenous intra-household balance of power and its impact on expenditure patterns: evidence from India. *Econometrica*, 73: 435-460.

Lazarus R S (1991) Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist*, 46(8), 819.

Lerner J S, Small D A, Loewenstein G (2004) Heart strings and purse strings carryover effects of emotions on economic decisions. *Psychological Science*, 15(5), 337-341.

Manser M, Brown M (1980) Marriage and household decision-making: a bargaining analysis. *International Economic Review*, 21(1): 31–44.

Mansour H, McKinnish T (2014) Couples' time together: complementarities in production versus complementarities in consumption. *Journal of Population Economics*, 27: 1127-1144.

McAlister L (1982). A dynamic attribute satiation model for choices made across time. *Journal of Consumer Research*, 9(3), 141-150.

McLachlan G J, Peel D (2000) *Finite Mixture Model*s. Wiley.

Michaud P C, Vermeulen F (2011) A collective labor supply model with complementarities in leisure: Identification and estimation by means of panel data. *Labour Economics*, 18:159-167.

Musalem A, Bradlow E, Raju J (2008) Who's got the coupon? estimating consumer preferences and coupon usage from aggregate information. *Journal of Marketing Research*, 45 (December): 715–730.

Musalem A, Bradlow E, Raju J (2009) Bayesian estimation of random-coefficients choice models using aggregate data. *Journal of Applied Econometrics*, 24 (3): 490–516.

Poor M, Duhachek A, Krishnan S (2012) The moderating role of emotional differentiation on satiation. *Journal of Consumer Psychology*, 22(4), 507-519.

Roehm H A, Roehm M L (2005) Revisiting the effect of positive mood on variety seeking. *Journal of Consumer Research*, 32(2), 330-336.

Roy R, Chintagunta P K, Haldar S (1996) A framework for investigating habits, "The Hand of the Past", and heterogeneity in dynamic brand choice. *Marketing Science*, 15(3), 280-299.

Samuelson P (1956) Social indifferent curves. *Quarterly Journal of Economics*, 70(1):1–22.

Seetharaman P B (2004) Modeling multiple sources of state dependence in random utility models: A distributed lag approach. *Marketing Science*, 23(2), 263-271.

Seetharaman P B, Chintagunta P (1998) A model of inertia and variety-seeking with marketing variables. *International Journal of Research in Marketing*, 15(1), 1-17.

Shachar R, Anand B (1998) The effectiveness and targeting of television advertising. *Journal of Economics and Management Strategy*, 7(3):363–396.

Shaffer G, Zhang Z J (2000) Pay to switch or pay to stay: Preference?based price discrimination in markets with switching costs. *Journal of Economics & Management Strategy*, 9(3), 397-424.

Siddarth S, Chattopadhyay A (1998) To zap or not to zap: a study of the determinants of channel switching during commercials. *Marketing Science*, 17:124-138.

Stephens M (2000) Dealing with label switching in mixture models. *Journal of the Royal Statistical Society*, B(62): 795-809.

Su C, Fern E F, Ye K (2003) A temporal dynamic model of spousal family purchase-decision behavior. *Journal of Marketing Research*, 40(3): 268-281.

Sullivan O (1996) Time co-ordination, the domestic division of labour and affect relations: time-use and the enjoyment of activities within couples. *Sociology,* 30: 97-100.

Tanner M, Wong W H (1987) The calculation of posterior distributions by data augmentation: rejoinder. *Journal of the American Statistical Association,* 82(298):548-550.

Trivedi M, Bass F M, Rao R C (1994) A model of stochastic variety-seeking. *Marketing Science*, 13(3), 274-297.

Vranica S, Perlberg S (2015) In advertising Battle, TV fights back against the web. *The Wall Street Journal* (April 28 2015).

Wadhwa M, Shiv B, Nowlis S M (2008). A bite to whet the reward appetite: The influence of sampling on reward-seeking behaviors. *Journal of Marketing Research*, 45(4), 403-413.

Walsh J W (1995) Flexibility in consumer purchasing for uncertain future tastes. *Marketing Science*, 14(2), 148-16.

Wang P (2012) *Informative Bayesian modeling with applications to media data*. Dissertation.

Wen C, Koppelman F (2001) Then generalized nested logit model. *Transportation Research*, B(35): 627-641.

Yang S, Allenby G (2003) Modeling interdependent consumer preferences. *Journal of Marketing Research*, 40(3): 282-294.

Yang S, Narayan V, Assael H (2006) Estimating the interdependence of television program viewership between spouses: a Bayesian simultaneous equation model. *Marketing Science*, 25(4): 336-349. Yang S, Zhao Y,

Erdem T, Zhao Y (2010) Modeling the intra-household behavioral interaction. *Journal of Marketing Research*, 47(3): 470-484.

Zhang J, Kuwano M, Lee B, Fujiwara A (2009) Modeling household discrete choice behavior incorporating heterogeneous group decision-making mechanisms. *Transportation Research*, B(43): 230–250.