Ph. D. Dissertation in Economics

Incorporating Latent Psychological Factors and Social Interaction in A New Generalized Heterogeneous Data Model (GHDM)

잠재 심리학 변수 및 사회적 상호작용을
고려한 새로운 일반화된 이질적 데이터 모형

August 2020

Graduate School of Seoul National University

Technology Management, Economics, and Policy Program

Meihan He

Incorporating Latent Psychological Factors and Social Interaction in A New Generalized Heterogeneous Data Model (GHDM)

지도교수 이종수

이 논문을 경제학박사학위 논문으로 제출함

2020 년 8 월

서울대학교 대학원
협동과정 기술경영경제정책 전공
하미함

하미함의 경제학박사학위 논문을 인준함

2020 년 6 월

위 원 장 구 윤 모 (인)

부위원장 이 종 수 (인)

위 원 신 정 우 (인)

위 원 안 중 하 (인)

위 원 김 연 중 (인)

Abstract

Incorporating Latent Psychological Factors and Social Interaction in A New Generalized Heterogeneous Data Model (GHDM)

Meihan He
Technology Management, Economics, and Policy
Engineering
Seoul National University

Social interaction has enormous effects on an individual's attitude and opinion, preferences, and behaviors. Social interaction is one of the most important fields of study within social sciences. There are generally two types of social interaction: word-of-mouth and observed learning. Observed learning is considered as social interaction in the majority of choice models in economic studies. However, word-of-mouth is largely investigated in attitude and behavior propensity-related studies, and has hardly ever been incorporated in choice models due to its characteristics. On the other hand, an attempt has been made to incorporate an individual's psychological variables into choice models in order to facilitate a better understanding of an individual decision process and improve the forecasting ability of the models. However, limited studies have considered the effect of social interaction on an individual's psychological variables which are some of the major mechanisms of social interaction. Moreover, since an individual's behaviors endogenously correlate with each other, simultaneous consideration of endogenously correlated outcomes is necessary for many choice situations. Although there are some studies that have derived a handful of models for multiple choices, social interaction has been incorporated into only a few models. This study proposes a new multiple endogenous choice model incorporating both types of social interaction. Furthermore, the proposed model is capable of dealing with multiple endogenous heterogenous dependent variables. In this dissertation, a simulation study has been conducted to confirm the performance of the proposed model, and an empirical study has been conducted to provide evidence of how a social interaction effect on an individual's choice and ignoring such an effect may lead to inconsistent estimation and over-estimation of the variable effects.

Keywords: Social interaction, Integrated Choice and Latent Variables (ICLV), Heterogenous dependent variable, Composite marginal likelihood (CML)

Student Number: 2015-30854

Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Research Background

The most basic units of an economy are firms, households, and the government. Firms make production decisions, households make consumption decisions and production of own factors, and the government collects and redistributes taxes. Individuals, as the most basic unit of households, are conceptualized as decisionmakers who have certain preferences, form expectations, and face constraints (Manski, 2000). In random utility theory, the coefficients of observed factors in the utility function are interpreted as personal preference toward certain factors which are usually the parameters that need to be estimated. Many economic studies assume that the individuals can always reach the optimal choice with the inherent preferences, and do not explain how the individuals obtain the preference structure and how they maximize their utilities (Manski, 2000). Economic studies, especially choice-related studies, usually attempt to estimate the true preferences of individuals, and then adjust the observed factors to predict the future market with different scenarios. In choice applications, parameters associated with explanatory variables and covariates are assumed to be fixed. However, marketing studies indicate that there are variations in preferences either in the long-term or short-term (Guhl, Baumgartner, Kneib, & Steiner, 2018).

In psychology, individual preferences are either inherited through genetic transmission or determined by imitation processes (Bisin & Verdier, 2001). Take pro-social behaviors as an example. Some studies argue that pro-social behaviors, including helping, sharing, and other moral behaviors among individuals, are taught by adults and learned by children (Bandura, 1986; Bisin & Verdier, 2001; Duesenberry, 1949); Jean Piaget argues that pro-social behaviors among children are learned through interactions with peers (Flavell, 1963). Similar to pro-social behaviors, the "inherent" preferences of individuals are also learned from either adults or peers through social interactions.

Human beings are gregarious and generally live in groups. An individual's behaviors are affected by the surrounding environment. According to the theory of human needs (Maslow, 1943), five types of human needs are defined hierarchically: physiological needs, safety needs, belongingness and love needs, esteem needs, and self-actualization needs. Physiological and safety needs are categorized as basic needs; the other needs are defined as psychological needs. All kinds of human needs highly depend on social environments. Human beings fulfill their needs and thrive by continually interacting with others.

In economics, individuals tend to interact with others who share similar characteristics, which is defined as homophily. They also tend to favor others who have a similar social identity, which is referred to as in-group bias. Therefore, it is no surprise that due to different types of social interactions, the behaviors of individuals with similar characteristics are correlated. Manski (2000) states that an individual's behavior might not have a causal effect on another. The correlated behavior may be due to the endogenous group formation phenomenon (which can be seen as another type of social interaction), or just due to simultaneity.

Hogg (2000) states that homophily is a way of reducing subjective uncertainty toward a situation. Both homophily and in-group bias indicate the propensity of segregation, and are intercorrelated. One of the most popular approaches in identifying homophily and in-group bias is to conduct a controlled laboratory experiment; see the study conducted by Currarini and Mengel (2016). To capture a pure social interaction effect with panel data, add individual fixed (Nair, Manchanda, & Bhatia, 2010) or random effects (Hartmann, 2010), and control for homophily and other correlated unobserved factors. Shalizi and Thomas (2011) indicate that even when homophily is teased out by conditioning on a previous decision and observed individual characteristics, there still exists unobserved individual characteristics that confound the social interaction effects with unobserved homophily which is defined as latent homophily. Regarding latent homophily, Wang, Aribarg, and Atchade (2013) indicate that the correlated unobserved characteristics may be due to exposure to the same external stimuli. The use of geographic and demographic proximities to infer social relations can further exacerbate this type of endogeneity. Correlated behavior within a reference group might be due to propensity to maintain one's status against that of their peers (Bernheim, 1994), or conform to the social norm of the reference group (Akerlof & Kranton, 2000), or social interaction as a form of observational learning (Banerjee, 1992; Bikhchandani, Hirshleifer, & Welch, 1992).

No matter what kind of mechanisms drive the correlated behaviors within individuals who share similar characteristics, it is obvious that human behaviors are influenced by other individuals. To further understand the correlated behaviors of individuals who share similar characteristics, it is necessary to pay attention to social cognition, which is a sub-topic of social psychology. Social cognition focuses on how people process, store, and apply information within a social context, and on how the cognitive processes influence social interaction. Social cognition relies on the perspective-taking ability of an individual, which is developed during childhood.

Individuals rely on self-information, and are biased in viewing their own position as normative when they have already formed their preferences toward a certain situation (Rabin, 1993). On the other hand, when individuals face a new situation that they have never experienced before, people who have similar preferences and values can serve as surrogates. When individuals notice that their opinion is idiosyncratic, they tend to rely on the opinion of others especially of those who clearly have more expertise about the situation. Whenever people lack knowledge of others, they first categorize them with respect to their similarity with self, then they rely on self-information to anticipate the behavior of similar others (Gramzow, Gaertner, & Sedikides, 2001). An individual's "inherent" preferences are as a result of taking others' opinions and observing behavioral outcomes. Therefore, it is necessary to consider the effect of other individuals on an individual's formation of preference. Exchanging information with others and observing their behaviors is considered as social interaction.

1.2 Research Object

Social interactions affect individual behaviors. In an economic realm, they affect an individual's purchasing behavior. Economic growth is one of the main engines of social development which depends on diffusion of innovations. The diffusion process is driven by aggregated individual behaviors as well as the interactions between individuals.

According to Roger's innovation diffusion theory, four main factors affect innovation diffusion: (1) the characteristics of the innovation, (2) time, (3) the communication channel, and (4) the social system (Rogers, 2010). Communication channels are divided into external (e.g., mass media) and internal communication channels (e.g., word-of-mouth). An individual is informed about the innovation from either an external or internal source. The communication between an individual and mass media is typically one-way. The information flows from mass media to the individual, and the information is not varied according to the feedback from individuals. Therefore, there is no interaction between an external communication channel and individuals. On the other hand, the communication between an individual and entities through an internal communication channel is two-way. The information flows with feedbacks in the internal communication channel. The communication in internal channels does not have to be verbal. For example, an individual's expected utility of an innovation may be affected by the adoption behavior of other individuals. This kind of communication does not have to be conducted verbally, and it is generally referred to as observed learning. A typical two-way verbal communication is known as word-of-mouth. Human communication is designed to influence others by modifying their beliefs, values, or attitude (Simons, 2001). Individuals exchange information about the innovation. During the communication process, individuals are informed about the innovation, gain knowledge of the innovation, form an attitude toward the innovation, exchange their personal opinion regarding the innovation, and adjust their personal attitude or opinion at the end of the communication. Further, individuals form their expected utility according to the knowledge that they possess, and their attitude and opinion. Internal communication plays an essential role in an individual's expected utility formation process.

Individuals "tend to be linked to others who are close to them in physical distance and who are relatively homophilous in social characteristics" (Rogers, 2010). Individuals are more likely to communicate with individuals who share similar characteristics such as beliefs, education, socio-economic status, and so forth, and pay attention to objects that attract other individuals' attention (Salganik, Dodds, & Watts, 2006).

When individuals form their expected utility, they rely on the knowledge that they possess, and their attitude and opinions. All information through either an external or internal communication channel is converted to personal knowledge, attitude, and opinion. According to classical economic theory, individuals maximize their utility with given restrictions. When individuals face a choice situation with multiple alternatives, they chose the alternative that maximizes their utility. People tend to presume that they make their decisions by their own will, but the knowledge, attitude, and opinion are never their own. That is why we say individuals are independent beings, but they correlate with each other.

Information exchanged through a communication channel includes the characteristics of the innovation such as its relative advantage, compatibility, complexity, observability, and triability. This information helps individuals gain knowledge about the innovation. During the communication, individuals also obtain information about the attitude and opinions of others. Such information further has an effect on the individual's attitude and opinion. An individual adjusts his or her attitude and opinion to be closer to those of the communicated other if he or she believes that the communicated other is similar to him or her in some aspect, and more rational or more professional toward the innovation, otherwise, he or she would probably maintain his or her original attitude or opinion.

Rogers (2010) state that the "innovation adoption decision is the process through which an individual pass in sequence from (1) gaining initial knowledge of an innovation, (2) to forming an attitude toward the innovation, (3) to making the adopt-or-reject decision, (4) to implementing the new idea, and (5) to finally confirming the decision". The persuasion stage is extremely important in individual's decision process. The persuasion stage is extremely important in an individual's decision process. The individual considers whether to adopt the innovation only if he or she is persuaded about the advantage of the innovation. Similarly, the individual adjusts his or her attitude or opinion only when he or she is persuaded about the other's attitude or opinion.

In psychology, there are multiple theories that deal with persuasion. Expectation disconfirmation theory states that attitudes change over time and are strongly affected by the initial attitude. Mere-exposure effect was discovered by Zajonc (1960), and it argues that an individual's attitude toward an alternative becomes more positive by purely increasing the exposure level. This is also known as familiarity principle. Festinger (1954) presented social comparison theory that insists that individuals' attitudes and opinions are affected by communication between individuals. Individuals' choices are influenced by social dependency effects. As individuals interact and exchange information with, or observe the behavior of, those in close proximity to themselves, they are likely to shape their behaviors. Particularly, geographic proximity is the most studied in econometric choice models. However, with the development of information technology, the strength of associations among individuals is not solely dependent on spatial proximity. Rather, as discussed above, individuals are more likely to be affected by others who share similar attitudes, values, lifestyle, and so forth. Ignoring such effects may lead to misunderstanding of individuals' adoption and usage behaviors.

Traditionally, social interaction is considered as an independent factor from other explanatory variables, and it usually acts as an additive term, which indicates that social interaction (mostly the proportion of adopted individuals among the whole population or a certain neighborhood) directly affects an individual's choice. In general, economists have assumed that agents do not directly observe the expectation of other agents. In many cases, observational learning generates expectation interaction. Expectation interactions pervade the modern economics of information. Social interaction is largely limited to observed learning in consumer choice models. In case of innovations with strong network externality effects, the number of adopted individuals can be an ideal explanatory variable of expected utility. Otherwise, this kind of social interaction (which is defined as observed learning) leads to preference formation of other related explanatory variables. The information of a number of adopted individuals helps an individual to infer about the factors that he or she cares about. Both observed learning and word-of-mouth play an important role in social interaction related studies. However, word-of-mouth effects are seldomly considered in consumer choice models. From a conceptual standpoint, the underlying attitudes and opinions are the ones that are likely to be exchanged through interactions, and subsequently, these attitudes and opinions impact the choices.

1.3 Research Outline

Social interaction has an enormous effect on individual behaviors. To reflect the effect of social interaction on an individual choice behavior, it is necessary to simultaneously mimic the cognitive mechanism in social interaction and model the choice. Moreover, an individual's choices are correlated with other multiple behaviors. For example, a household's choice of residential location is affected by vehicle ownership and travel miles (Bhat, 2014). Joint modeling of multiple outcomes is of interest in many fields, including clinical biology, health, transportation, and so forth (Bhat, 2015a).

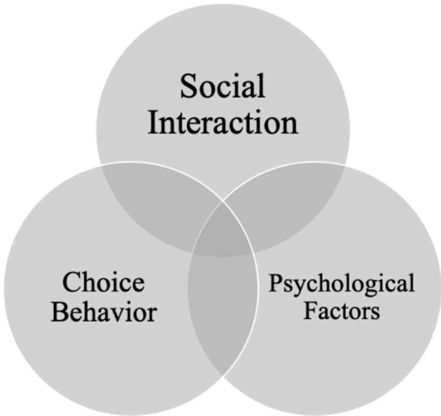

Figure 1. Research Objects

The research objects of this study are shown in Figure 1. This study proposes a multi-choice model considering the effects of social interaction on an individual's attitude and opinion. The proposed model can be a valuable tool for modeling social dependencies in multidimensional mixed data outcomes. Empirically, the proposed model helps a researcher to gain a better understanding of the underlying motivation of individual behaviors. Chapter 2 includes a general review of previous literature. In Chapter 3, the proposed model is specified and simulated. Chapter 4 provides an empirical application of the proposed model, and Chapter 5 concludes the dissertation.

Chapter 2. Literature Review

The overview of this chapter is shown in Figure 2. Social interactions influence an individual's attitude and opinion, preferences, and behaviors. Since social interactions can take various forms, it is necessary to review major studies on social interaction. An individual's psychological activities, which have an effect on observed choice, are persistently incorporated in choice models. Moreover, the effect of social interaction on observed individual choices is also an important branch of choice models. This chapter includes a general review of previous studies on social interaction, choice models with psychological factors and social interaction, and mixed data models.

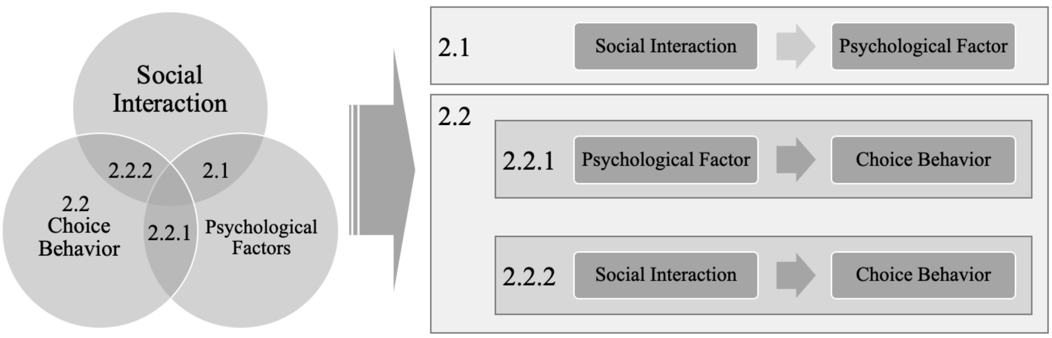

Figure 2. Instruction of Chapter 2

2.1 Theoretical Insights

2.1.1 Studies on Human Behavior

To understand the behaviors of individuals, it is necessary to borrow insight from behaviorism. Behaviorism assumes that an individual's behavior is either a reflex evoked by the pairing of certain antecedent stimuli, or is a consequence of the individual's history, along with the individual's current motivational state and controlling stimuli. Following the theory of behaviorism, most individual behaviors are learned and can be adjusted by learning. Moreover, continuing or abandoning a learned behavior is affected by the antecedent stimuli and the result that follows after the behavior. John B. Watson first proposed this idea in his psychology research in 1913. Following Pavlov's classical conditioning theory and Skinner's operant conditioning theory, Albert Bandura suggested observation learning (or social learning theory). Social learning theory, which combines conditioning theories of behaviorism and learning theory of cognitivism, is an effective psychological approach in the investigation of human behavior. Social learning theory claims that human beings are social mammals; human beings' behaviors are not just learned by reward and punishment but also by observing the behaviors and results of other beings (Bandura, 1986).

Traditional behavioral models assume a linear relationship between knowledge and attitude. However, although there is an extremely strong correlation between attitude and behavior, the relationship between knowledge and attitude is extremely complicated (Thompson & Mintzes, 2002). Ajzen (1980) presented the theory of reasoned action which aims to predict human behaviors by explaining the relationship between attitudes and behaviors. This theory states that knowledge is the only factor that has an effect on attitude, while external constraints and situations have more influence on change of behavior. The theory of reasoned action is based on the social learning theory, expectancy value theory, cognitive consistency, and attribution theory. One of the most important explanatory factors of an individual's behavior is whether he or she has the behavioral intentions which are based on personal attitude, important others' attitudes, perceived social pressure, and so on. The theory of reasoned action assumes that before a certain behavior, an individual searches for relative information, rationally considers the result of the behavior, and then makes a decision. However, in many occasions, individuals' behaviors cannot be explained by behavioral intention. Ajzen proposed the theory of reasoned action which assumes that intention toward attitude, subjective norm, and perceived behavioral control shape an individual's behavioral intentions and behaviors.

In an individual's decision process, after being informed about the innovation and before the decision stage, an individual needs a persuasion process where he or she gains knowledge, forms his or her own attitude toward the innovation, and is persuaded to adopt the innovation. Individuals can learn from and be influenced by social interactions with others in two ways: they can extract information directly from others' opinions, or they can infer information indirectly from observing others' previous product adoption decisions. Since social interaction plays an important role in the persuasion stage, it is necessary to understand how individuals get persuaded through social interaction.

Some studies have proven that the effect of social interaction varies according to many aspects of social contexts. Hartmann et al. (2008) distinguish social interaction as either "active" or "passive." The former takes place when two individuals are in a dyadic relationship and recognize the effect of their own outcome on the other. The latter takes place when one of the two individuals does not recognize the effect of his or her outcome on the other individual. "Passive" social interaction considers the impact of an individual's action on the other's action as being a sequential process. Kelman (2017) divides social influence into two types: "value-expressive" influence, which is driven by one's goal to maintain and enhance their self-perception or identity within a reference group, and "informational" influence, which is driven by one's goal to gain more knowledge to reduce the uncertainty associated with a decision. Wang et al. (2013) investigate how individuals' product choices are influenced by the product choices of their connected others and how the influence mechanism may differ for fashion- versus technology-related products. The effect of social interaction depends on the relationship between individuals, the purpose, and the object.

Some studies have attempted to investigate the correlated behaviors in aspects of social interaction. Narayan, Rao, and Saunders (2011) investigated multi-attribute product choices considering peer influence. They proposed a two-stage conjoint-based approach in examining behavioral mechanisms of peer influence. Lu and Tang (2019) investigated correlated behavior through social interaction effects on an individual's asset allocation decision. Ma, Krishnan, and Montgomery (2015) conducted an empirical study on purchases of caller ring-back tones. They attempted to measure latent homophily, social influence, and exogenous factors using dynamic panel data and the available detailed communication data.

Another effect of social interaction is to assist individuals in maintaining their social identity. To avoid isolation and envy, most people attempt not to be too far ahead in income comparisons within their reference group. Risk-taking is affected by the social context. The magnitude of this impact is likely to depend on individuals' sensitivity to social comparison. Individuals' choices are not only context-dependent but are also sensitive to their degree of inequality aversion. Müller and Rau (2019) address the question of how social contexts and the heterogeneity in risk and social preferences affect individual risk-taking. They theoretically identify a certain social preference as a potential driver of how social context may impact risk-taking. Some studies have established inequality aversion as an important dimension of social comparison. They make use of this concept and extend inequality aversion to an uncertain environment incorporating risk preferences.

The others' opinions, often termed word-of-mouth (WOM), can have a significant impact on consumers' purchase and adoption behaviors (Lovett & Staelin, 2016). On the other hand, others' purchase decisions can affect an individual's own decisions, leading to herding behavior (Zhang & Liu, 2012). An individual can be influenced by his or her friends' opinions and/or actions while concurrently observing product adoptions, online reviews, and ratings by users beyond his or her personal network. Some argue that WOM plays a more important role than observed adoption, because compared to observed adoption, WOM conveys more diagnostic information about product quality. On the other hand, actions are more powerful than words. However, as discussed above, the majority of economic studies that considered social interaction only investigated the effect of observed learning.

2.1.2 Studies on Word-of-Mouth

Word-of-mouth is one of the most well-known mechanisms of social interaction. Word-of-mouth communication consists of exchange of information, attitudes, and opinions about an innovation. Word-of-mouth communication is a dynamic process.

Opinion dynamics is one of the most explored branches of word-of-mouth communication. It is a fusion process of individual opinions in which interacting agents within a group continuously update and fuse their opinions on the same issue based on the established fusion rules. Opinion dynamics models are usually composed of a few basic elements: opinion expression formats, fusion rules, and opinion dynamics environments, and focus on three varieties of stabilized patterns: consensus, polarization, and fragmentation.

The DeGroot model (DeGroot, 1974) is one of the most classic models in opinion dynamics. It assumes that the weight of other individuals does not change over time or with opinions. It also assumes that individuals' opinions are continuous. The bounded confidence model (Dittmer, 2001) was developed from Krause's consensus formation model (Krause, 2000) which is a dynamic model for investigating consensus of opinions. The bounded confidence model assumes an individual's opinion will only be influenced by agents whose opinions differ from that of the individual no more than a certain confidence level. Considering the above psychological factors makes the bounded confidence model popular in opinion dynamics studies. Otherwise, there are heterogeneous bounded confidence models such as the Deffuant-Weisbuch (DW) model (Deffuant, Neau, Amblard, & Weisbuch, 2000) and the Hegselmann-Krause (HK) model (Hegselmann & Krause, 2002). The DW model considers disagreement dynamics. Heterogenous bounds of confidence assume that all individuals are exposed to external information, hence, the HK model additionally considers heterogenous bounds of confidence. In the DW model, two individuals are randomly selected from a set of agents; subsequently, based on bounded confidence, the two agents decide whether to communicate. The DW and the HK models both rely on the idea of repeated averaging under bounded confidence. In the DW model, agents meet in random pairwise encounters after which they do or do not compromise; in the HK model, each agent moves to the average opinion of all agents who lie in his/her area of confidence. The HK model is more suitable for modeling situations like formal meetings, where interaction occurs in a large group, while the DW model is better suited for pairwise interactions within large populations (Castellano, Fortunato, & Loreto, 2009).

There are some extensions of bounded confidence models that consider other factors that can influence an individual's opinion, such as propaganda as an external message (Carletti, Fanelli, Grolli, & Guarino, 2006), repulsive links and external factors (Martins, Pineda, & Toral, 2010), and disagreement and possibility of modulating external information/media effects both from one and multiple sources (Sîrbu, Loreto, Servedio, & Tria, 2013).

The relative agreement model (Deffuant, Amblard, Weisbuch, & Faure, 2002) is an extension of the bounded confidence model which assumes that people do not take into account opinions out of their range of uncertainty. Quattrociocchi, Conte, and Lodi (2011) extended the relative agreement model by introducing additional sources of information. A great deal of studies have attempted to investigate social interaction in the realm of individual attitude and opinion dynamics by considering different ideas, such as truth seekers (Hegselmann & Krause, 2006), external messages such as propaganda (Carletti et al., 2006), repulsive links and mass media (Martins et al., 2010), wise agents and televiewers (Quattrociocchi et al., 2011), truth as external information (Kurz & Rambau, 2011), disagreement and possibility of modulating external information/media effects (Sîrbu et al., 2013), and gossip leader and gossip follower (Quattrociocchi, Caldarelli, & Scala, 2014).

Summative model assumes that an individual has many beliefs toward an alternative, and only few beliefs that are salient form the individual's attitude toward the alternative. A salient belief is determined by its strength and evaluation. An individual's change in attitude regarding whether to add new salient beliefs, increases/decreases favorability of existing positive/negative beliefs. Information exchange is needed in the process of changing attitude. The flow of information can be through either external or internal communication channels. Change in attitude is affected by other individuals' attitudes because of the need for similarity (social comparison theory, see Festinger (1954)) and cognitive consonance.

Festinger (1957) proposed the theory of cognitive dissonance, and claims that human beings seek internal psychological consistency. Cognitive dissonance occurs when cognitive elements (such as beliefs, opinions, attitudes, and knowledge) contradict with each other. Individuals tend to become psychologically uncomfortable when they experience cognitive dissonance, and are willing to reduce the cognitive dissonance by adding new cognition or by avoiding circumstances and contradictory information likely to increase the magnitude of the cognitive dissonance.

The above studies have focused on dynamics of opinions among individuals and have no extra consideration of an individual's actual choice of behavior. The fuse of opinion was completed after occurrence of the stable patterns.

Another important branch of the mechanism of social interaction is observed learning. There are also numerous studies that have explored the effect of observed learning on an individual's decision behavior with regard to opinion dynamics. Contrary to the typical opinion dynamics models, the continuous opinion and discrete actions model (CODA) (Martins, 2008) assumes that individuals can only observe others' actions that have an effect on the individual's probability of conducting the action following Bayes' rule. Under circumstances where each individual notices only the choices of other individuals, but is not aware of their internal opinions, there is no way that interacting individuals will converge to a mean result as in the bounded confidence models

In the CODA model, individuals show discrete behaviors but express continuous opinions that are updated by interacting with other agents. Each agent changes its continuous internal probability toward the value of its peers. When someone faces a binary decision, the opinion about which option is the best one is not necessarily binary. For most problems, it is reasonable to assume that the person believes one of the alternatives is better with a probability p. If the consequences of being right or wrong are equivalent for both choices, the alternative with a higher probability, p or $1-p$, will be chosen as the best one.

Martins (2009) investigated the use of Bayesian updating rules in the CODA model, and analyzed mobility and social network effects on extremist opinions. Martins (2013) introduced the concept of trust in the CODA model, with agents holding an array of probabilities that the others are trustworthy, and explored the relationship between the CODA model and traditional discrete models.

There are many applications of opinion dynamics models and CODA models in investigating the effect of social interaction on either individual opinion formation or individual decision behavior. Sznajd-Weron and Weron (2003) proposed an Ising model to describe the mechanism of advertising in a duopoly market. Schulze (2003) argued that the influence of advertising can be simulated in a binary model by the probability to change opinion, irrespective of the normal convincing process. Li, Braunstein, Havlin, and Stanley (2011) investigated the strategy of competition between two groups based on an inflexible contrarian opinion model. Quattrociocchi et al. (2014) investigated how mainstream media signed interaction might shape the opinion space, focusing on how different numbers of media and interaction patterns of the information system affect collective debates and opinion distribution. Innes (2014) proposed an aggregation model incorporating ingroup-outgroup dynamics as well as media influence to establish potential causal relationships between various types of social interaction and social phenomena such as the occurrence of group consensus and a hostile media effect, then further applied the model to simplified commercial applications in advertisement optimization to determine the optimal proportion of a population to target with advertising in order to maximize opinion shift while fixing cost. Varma, Morarescu, Lasaulce, and Martin (2017) analyzed competition between two firms, where each firm attempts to sway public opinion to its own side by spending money on advertising or discounts for specific consumers, thus, capturing a larger market share. Castro, Lu, Zhang, Dong, and Martínez (2018) proposed a recommender system based on opinion dynamics to assist users in selecting the right products or services in information overload scenarios.

There are also some extensions considering noise and uncertainty within social interaction. When people express their opinions, sentiments, or emotions regarding different issues, such as politics, products, and events, they often cannot provide exact opinions, but express uncertain types of opinions. Liang, Dong, and Li (2016) investigated the dynamics of interval opinion formation within the framework of bounded confidence. Liang et al. (2016) focused on opinion formation in a linguistic environment, and proposed a linguistic opinion dynamic with bounded confidence. Wang and Mendel (2016) proposed a new mathematical framework for the evolution and propagation of fuzzy opinions.

The above studies have comprehensively investigated the mechanism of word-of-mouth. However, most of the studies did not consider an individual behavior. The individual behavior considered in CODA is not the behavior that was considered in choice behavior in the economic realm. The individual behavior considered in CODA does not face any constraints.

2.2 Choice Models

2.2.1 Choice Models with Psychological Factors

In random utility model (RUM), an individual's utility of an alternative is depicted in the form of $U = V + \varepsilon$, and the framework in shown in Figure 3. The solid arrows indicate the causal effect represented by structural equations, and the dashed arrows indicate underlying relation represented by measurement equations. The term V represents the observed factors, and the term ε represents the unobserved factors of the individual's utility.

Figure 3. Random Utility Choice Model

The RUM-based choice models do not consider the formation of an individual's attitudes and perceptions. Since researchers have no information about such unobserved factors, all unobserved factors are ε, which are treated as random vectors following a researcher's specific distribution. According to the density distribution of ε, $f(\varepsilon)$, the choice model can be divided into the logit model, the generalized extreme value (GEV) distribution, and a probit model. The observed factors are generally represented as observed individual characteristics and observed alternative specific characteristics which take a general form of $X\beta$, where X represents the observed factors and β represents the corresponding coefficients which need to be estimated.

Considering that there are unobserved individual characteristics that influence an individual's choice, researchers make an effort to develop choice models that can reflect the unobserved individual characteristics. Some studies include psychological factors such as attitudes and perceptions of individuals directly into choice models (Harris & Keane, 1998; Koppelman & Hauser, 1978). The framework is shown in Figure 4. The solid arrows indicate the causal effect represented by structural equations, and the dashed arrows indicate underlying relations.

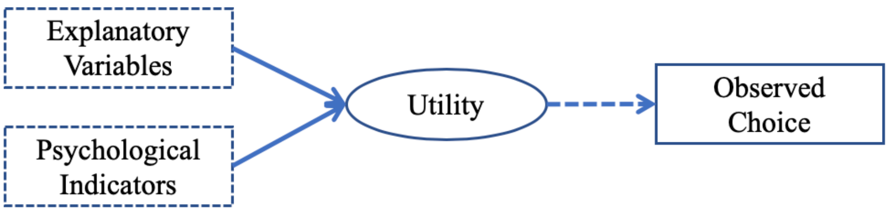

Figure 4. Choice Model with Psychological Indicators as Independent Variable

An example of this approach is the study conducted by Bhat, Schofer, Koppelman, and Bautch (1993). They considered the effect of individual attitudes on the choice outcomes. The study assumes that an individual's choice is determined by characteristics of alternatives and the individual's characteristics and attitudes. The latent utility for an ordered dependent variable of individual q is defined as:

$$U_q = \beta' x_q + \varepsilon_q \quad \cdots\cdots\cdots\cdots \text{Eq. (2.2.1)}$$

where ε_q is assumed to be normally distributed with a mean equals to zero and a variance equals to one. The observed choice of individual q is P_q:

$$P_q = \begin{cases} \text{definitely no,} & \text{if} \quad U_q < 0 \\ \text{probably no,} & \text{if} \quad 0 \le U_q < \psi_1 \\ \text{probably yes,} & \text{if} \quad \psi_1 \le U_q < \psi_2 \\ \text{definitely yes,} & \text{if} \quad U_q > \psi_2 \end{cases} \quad \cdots\cdots\cdots\cdots \text{Eq. (2.2.2)}$$

where ψ_1 and ψ_2 are threshold parameters that need to be estimated. The individual attitudes are treated as exogeneous variables. The study assumes that the attitudes of individuals are exactly represented by the indicator variables, and these attitudes have a direct effect on an individual's utility. Actually, the indicators are proxies to certain attitudes that cannot be observed directly by the researchers, and of course, there are measurement errors in the proxy procedure and the indicators may be correlated with other unobserved factors, leading to inconsistent estimation (Bhat & Dubey, 2014). Furthermore, ignoring the measurement errors may lead to inconsistent estimation (Ashok, Dillon, & Yuan, 2002).

Some studies have applied factor analysis to derive the psychological latent variables, and then included the latent variables into the choice model (Madanat, Yang, & Yen, 1995). This framework is shown in Figure 5. In this approach, the latent psychological variables act in a similar manner as the other exogeneous variables. The inclusion of such latent psychological variables can lead to a better understanding of an individual's decision process, and hence, better forecasting (Bolduc & Alvarez-Daziano, 2010; Temme, Paulssen, & Dannewald, 2008).

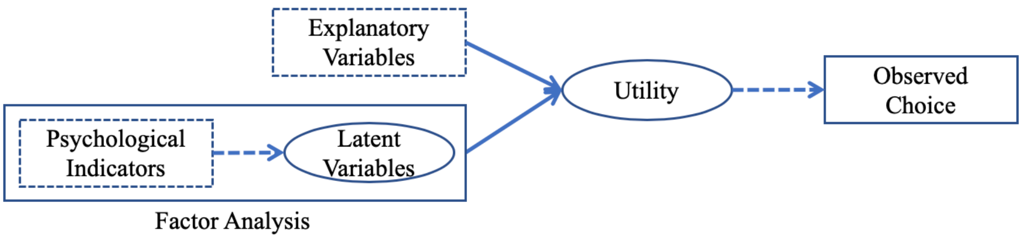

Figure 5. Choice Model with Latent Variable as "Error-Free" Independent Variable

However, this approach also treats the latent psychological variable as an "error-free" explanatory variable. Ignoring the correlation between the latent psychological variables leads to inconsistent estimation (Bhat, 2014).

A mixed logit model is one of the attempts to consider the taste variation among individuals. This model decomposes the unobserved factors into two parts: one part contains all kinds of correlation and heteroskedasticity and can follow any distribution; another part follows an independent identically distributed extreme value distribution. Bhat (1998) applied a mixed logit model to develop a random-coefficients logit (RCL) model to allow intrinsic preference for alternatives and variation in the sensitivity of attributes of the alternatives across individuals. The preference heterogeneity and response heterogeneity are decomposed into observed and unobserved individual characteristics, and defined as systematic and random heterogeneity, respectively.

Let U_{qi} indicate the utility of individual q on alternative i which can be decomposed into observed term $\eta'_{qi}x_{qi}$ and unobserved term Λ_{qi}. Subsequently, the utility of individual q on alternative i can be written as:

$$U_{qi} = \Lambda_{qi} + \eta'_{qi}x_{qi} \quad \cdots\cdots \text{Eq. (2.2.3)}$$

To account for systematic and random heterogeneity, the term Λ_{qi} is decomposed to 3 terms, and the utility function Eq. (2.2.3) is redefined as:

$$U_{qi} = \alpha_i + \delta'_i z_q + \varepsilon_{qi} + \eta'_q x_{qi} \quad \cdots\cdots \text{Eq. (2.2.4)}$$

where α_i represents an individual-invariant bias constant; $\delta'_i z_q$ represents systematic preference heterogeneity where z_q indicates the observed individual characteristics; ε_{qi} represents random preference heterogeneity which follows an independent and identically distributed type I extreme value distribution. The choice probability given η_q can be written as:

$$P_{qi}\big|\eta_q = \frac{\exp\left(\alpha_i + \delta'_i z_q + \eta'_q x_{qi}\right)}{\sum_{j=1}^{I}\exp\left(\alpha_j + \delta'_j z_q + \eta'_q x_{qj}\right)} \quad \cdots\cdots \text{Eq. (2.2.5)}$$

To allow for systematic response heterogeneity, the term η_q is considered as a vector with elements η_{qk} which indicate the corresponding coefficient of k th attribute of individual q. η_{qk} is defined as a function of observed individual characteristics: $\eta_{qk} = \gamma_k + \beta_k f\left(w_{qk}\right)$. To allow for random response heterogeneity over the systematic response heterogeneity, η_{qk} is redefined as $\eta_{qk} = \pm\exp\left(\gamma_k + \beta_k' w_{qk} + v_{qk}\right)$. The $\pm\exp(\cdot)$ term is applied to ensure a proper sign on the response coefficients, and w_{qk} as well as v_{qk} represent similar observed characteristics and random taste variation across individuals, respectively. The utility function in Eq. (2.2.4) can be written as:

$$U_{qi} = \alpha_i + \delta_i' z_q + \varepsilon_{qi} + \sum_k \left[\exp\left(\gamma_k + \beta_k' w_{qk} + v_{qk}\right)\right] x_{qik} \quad \cdots\cdots\cdots \text{Eq. (2.2.6)}$$

where the random taste term of individual q for attribute k, v_{qk}, follows a normal distribution with a mean that is equal to zero and variance that is equal to σ_k^2, and assumes that v_{qk} is independently and identically distributed across individuals. The RCL model is independent from irrelevant alternative (IIA) property which is the main shortcoming of the multinomial logit model. The choice probability given as v_{qk}, $k = 1,2,\ldots,K$ can be written as:

$$P_{qi}\left|\left(v_{q1},\ldots,v_{qK}\right)=\frac{\exp\left(\alpha_i+\delta_i' z_q+\sum_{k=1}^{K}\left[\exp\left(\gamma_k+\beta_k' w_{qk}+v_{qk}\right)\right]x_{qik}\right)}{\sum_{j=1}^{I}\exp\left(\alpha_j+\delta_j' z_q+\sum_{k=1}^{K}\left[\exp\left(\gamma_k+\beta_k' w_{qk}+v_{qk}\right)\right]x_{qjk}\right)}\right. \quad \cdots\cdots \text{ Eq. (2.2.7)}$$

The parameters are $\zeta_i=\left(\alpha_i,\delta_i'\right)'$ for each i, and $\xi_k=\left(\gamma_k,\beta_k',\sigma_k'\right)$ for each k. Stack all the parameters into vector $\theta=\left(\zeta_1',\zeta_2',\ldots,\zeta_I',\xi_1',\xi_2',\ldots,\xi_K'\right)'$ and the log-likelihood function for the RCL model can be written as:

$$\begin{aligned}L(\theta)&=\sum_{q=1}^{Q}\sum_{i=1}^{I}y_{qi}\log P_{qi}(\theta)\\&=\sum_{q=1}^{Q}\sum_{i=1}^{I}y_{qi}\log\left[\begin{array}{c}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left\{\dfrac{\exp\left(\alpha_i+\delta_i' z_q+\sum_{k=1}^{K}\left[\exp\left(\gamma_k+\beta_k' w_{qk}+\sigma_k u_{qk}\right)\right]x_{qik}\right)}{\sum_{j=1}^{I}\exp\left(\alpha_j+\delta_j' z_q+\sum_{k=1}^{K}\left[\exp\left(\gamma_k+\beta_k' w_{qk}+\sigma_k u_{qk}\right)\right]x_{qjk}\right)}\right\}\\ \cdot\phi\left(u_{q1}\right)\phi\left(u_{q2}\right)\cdots\phi\left(u_{qK}\right)du_{q1}du_{q2}\cdots du_{qK}\end{array}\right]\end{aligned} \quad \text{Eq. (2.2.8)}$$

where $\phi(\cdot)$ indicates the standard normal density function, and y_{qi} indicates the indicator function which equals to 1 if individual q chooses alternative i and equals to 0 otherwise.

The RCL model allows for intrinsic preference for alternatives and heterogeneous sensitivity of attributes to capture the taste variation across individuals. The main shortcoming of the RCL model (and similar approaches in Bhat (1997) and Revelt and Train (1998)) is that it cannot capture the correlation structure between individual characteristics and explanatory variables which can lead to inconsistent estimation and integrate out all the unobserved psychological factors (Bhat & Dubey, 2014).

To gain a better understanding of an individual decision making process, Ben-Akiva et al. (2002) developed an Integrated Choice and Latent Variable (ICLV) model which integrates latent psychological variables into a traditional choice model, taking the form of a hybrid model. An ICLV model considers both traditional explanatory and latent psychological variables which normally indicate an individual's attitude, propensity, perception, and so forth. A typical ICLV model incorporates latent psychological variables through a structural equation model (SEM). The SEM relates the latent psychological variables to the observed explanatory variables through a latent measurement equation model (MEM) considering any measurement errors in the SEM. The framework of an ICLV model is shown in Figure 6. The solid arrows indicate the causal effect represented by structural equations, and the dashed arrows indicate underlying relations.

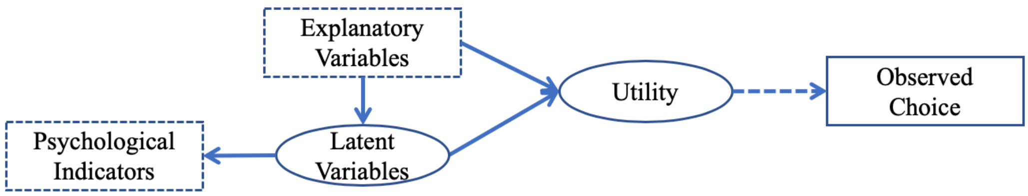

Figure 6. Integrated Choice and Latent Variable Model (ICLV)

General empirical studies of the ICLV model have applied an independent and identically distributed Gumbel error term for the choice model, and ignored the correlation between latent variables (Vij & Walker, 2014). Consideration of the correlation between latent psychological variables is necessary because there may be underlying unobserved individual characteristics that influence both the latent psychological variables and other explanatory variables, and hence, influence an individual's choice (Temme et al., 2008). A traditional ICLV model is estimated by a maximum simulated likelihood approach which is identical to a traditional mixed logit model. Since the integral of an ICLV model is a mixture of two probabilities, challenges are always encountered in the estimation of an ICLV model, which is extremely time consuming (Bhat & Dubey, 2014).

Bhat and Dubey (2014) propose a different model formulation for the ICLV model which is based on a multivariate probit (MVP) kernel. There are three components in the ICLV model: (1) a latent variable structural equation model, (2) a latent variable measurement equation model, and (3) a choice model. The lth latent variable z_l^* is defined as a linear combination of indicator variables:

$$z_l^* = \alpha_l' w + \eta_l \quad \text{Eq. (2.2.9)}$$

where w is a $(\tilde{D} \times 1)$ vector of observed indicator variables, and α_l is a $(\tilde{D} \times 1)$ vector of corresponding coefficients, and η_l is a normally distributed error term. Stack all the vectors together, the matrix form of Eq. (2.2.9) can be written as:

$$z^* = \alpha w + \eta \quad \text{Eq. (2.2.10)}$$

with $\eta \sim N(\mathbf{0}_L, \Gamma)$.

In the measurement equation model, consider there are H continuous variables $(y_1, y_2, \ldots, y_H)$. Subsequently, the h th continuous variable can be expressed as: $y_h = \delta_h + d_h' z^* + \xi_h$ where δ_h represents the intercept, d_h is a $(L \times 1)$ vector of latent variable loadings, and ξ_h is a normally distributed error term. The matrix form of all H continuous variables can be written as:

$$y = \delta + dz^* + \xi \quad \cdots\cdots \text{Eq. (2.2.11)}$$

Similarly, the matrix form for the measurement equation for G $(g = 1, 2, \ldots, G)$ ordinal variables with j_g categories can be written as:

$$y^* = \tilde{\delta} + \tilde{d} z^* + \tilde{\xi}, \quad \psi_{low} < y^* < \psi_{up} \quad \cdots\cdots \text{Eq. (2.2.12)}$$

where ψ_{low} and ψ_{up} indicate a $(G \times 1)$ vector of lower threshold and upper threshold, respectively, with $\psi_{g,0} < \psi_{g,1} < \psi_{g,2} \cdots < \psi_{g,N_g - 1} < \psi_{N_g}$; $\psi_{g,0} = -\infty$, $\psi_{g,1} = 0$ and $\psi_{g,J} = +\infty$. Define Σ_{y^*} as the correlation matrix of $\tilde{\xi} = (\tilde{\xi}_1, \tilde{\xi}_2, \ldots, \tilde{\xi}_G)$. Further, define $\breve{y} = \left(y', \left[y^* \right]' \right)'$, $\breve{\delta} = (\delta', \tilde{\delta}')'$, $\breve{d} = (d', \tilde{d}')'$, and $\breve{\xi} = (\xi', \tilde{\xi}')'$. Subsequently, combine the Eq. (2.2.11) and Eq. (2.2.12) into the following matrix form:

$$\breve{y} = \breve{\delta} + \breve{d}z^* + \breve{\xi}, \text{ with } E(\breve{y}) = \begin{bmatrix} \delta + dz^* \\ \tilde{\delta} + \tilde{d}z^* \end{bmatrix},$$

$$\text{and } Var(\breve{\xi}) = \breve{\Sigma} = \begin{bmatrix} \Sigma_y & \Sigma_{yy^*} \\ \Sigma_{yy^*}' & \Sigma_{y^*} \end{bmatrix} \quad \cdots\cdots \text{Eq. (2.2.13)}$$

In the choice model, consider the index of alternatives is i $(i = 1,2,\ldots,I)$ and the utility for alternative i is defined as:

$$U_i = \beta' x_i + \gamma_i'(\varphi_i z^*) + \varepsilon_i \quad \cdots\cdots \text{Eq. (2.2.14)}$$

where x_i is a $(D \times 1)$ column vector of explanatory variables, and β is a $(D \times 1)$ vector of corresponding coefficients. γ_i and φ_i are parameters to capture the effect of latent variables, and ε_i is a normally distributed error term. The matrix form of Eq. (2.2.14) can be written as:

$$U = x\beta + \lambda z^* + \varepsilon \text{ , with } \lambda = \gamma\varphi \quad \cdots\cdots \text{Eq. (2.2.15)}$$

In terms of utility, since only the difference matters, define the difference in utilities with respect to the chosen alternative as $u_{im}^* = U_i - U_m$ $(i \neq m)$, where U_m indicates the chosen alternative. Further, define the covariance matrix as $\breve{\Lambda}$, with elements of the difference of error terms as $\varsigma_i = \varepsilon_i - \varepsilon_1$ $(i \neq 1)$. Moreover, define the matrix Λ which is constructed from $\breve{\Lambda}$ by adding a row and a column of 0s in the first row and the first column. The reduced form of the ICLV model is defined by stacking Eq. (2.2.13) and the Eq. (2.2.15) together. Define $YU = \left[\breve{y}', U'\right]'$ and $YU \sim MVN_{H+G+I}\left(B, \Omega\right)$ where:

$$B = \begin{bmatrix} \breve{\delta} + \breve{d}\alpha w \\ x\beta + \lambda\alpha w \end{bmatrix} \text{ and } \Omega = \begin{bmatrix} \breve{d}\Gamma\breve{d}' + \breve{\Sigma} & \breve{d}\Gamma\lambda' \\ \lambda\Gamma\breve{d}' & \lambda\Gamma\lambda' + \Lambda \end{bmatrix} \cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. (2.2.16)}$$

This ICLV model can be easily estimated by the maximum approximation composite marginal likelihood (MACML) approach (Bhat, 2011) with no identification problem by ensuring that matrix Γ is a correlation matrix, and matrix $\breve{\Sigma}$ is a diagonal matrix with elements corresponding to ordinal variables equal to 1.

The above choice models attempt to incorporate latent psychological variables into a traditional choice model to gain a better understanding of individual decision-making and information processes, hence, making better predictions. However, including ICLV models, choice models with inclusion of latent psychological variables hardly ever consider the effect of social interaction which is one of the most important factors in shaping an individual's attitude and perceptions. As introduced in the previous chapter, an individual's "inherent preferences" are also the result of social interaction, including word-of-mouth and observed learning. Therefore, it is necessary to simultaneously consider the effect of social interaction on an individual's attitude and perception formation, and the effect of psychological factors on an individual's decision-making process.

2.2.2 Choice Models with Spatial/Social Dependence

Most econometric studies consider social interaction by focusing on the proportion of adopted individuals (refer to Figure 7). Since the major focus of individual behavior in microeconomics is an individual's purchasing behavior (or adoption behavior), studies on this behavior focus more on individual characteristics and purchasing behavior, not the individual's decision-making process but the outcome. Therefore, traditional choice models that include social interaction tend to consider the behavior of other individuals (e.g., the adoption behavior of other individuals). An individual's preference can be influenced by observing the consumption behaviors of others through gaining a sense of belongingness and socialization effects (Janssen & Jager, 2001).

An agent-based model is a powerful tool for modeling social interaction. The most common way of modeling social interaction in the agent-based models is by dividing the utility into a weighted sum. A weighted utility function includes individual preference and social influence (Delre, Jager, Bijmolt, & Janssen, 2010; McCoy & Lyons, 2014).

Roozmand et al. (2011) considered a utility function combined with social status, social responsibility, price, cultural effects, and an individual's personality. Broekhuizen, Delre, and Torres (2011) proposed a weighted sum of personal utility and social influence, with two types of social influence: past behavior of others and preferences of individual peers. Other studies have considered strong and weak ties between individuals without specifying a utility function (Goldenberg, Libai, & Muller, 2001; Tran, 2012). Kim and Hur (2013) suggested a utility function which is a weighted sum of inclination to a product, number of adopted peers, and number of adopted opinion leaders. Xiong, Payne, and Kinsella (2016) divided peer effects into information, experience, and externality effects.

An agent-based model generally focuses on theoretical development (Gilbert, 1997) and aims to understand social interaction in an abstract sense. Theoretical agent-based models lack the capacity to represent the real-world situation; empirical agent-based models highly rely on individual-level survey data (Zhang & Vorobeychik, 2017).

Brock and Durlauf (2001) generalized logistic models of individual choice which incorporate an additive term into the utility function to reflect social interaction. They assume that the utility an individual has received from a behavior directly depends on the choice of others. Hartmann (2010) developed a model that can be used to estimate social interactions and analyze their implications. The model extends a typical discrete choice model to include the decisions of a customer's peers. The model considers decisions as the equilibrium outcome of a coordination game.

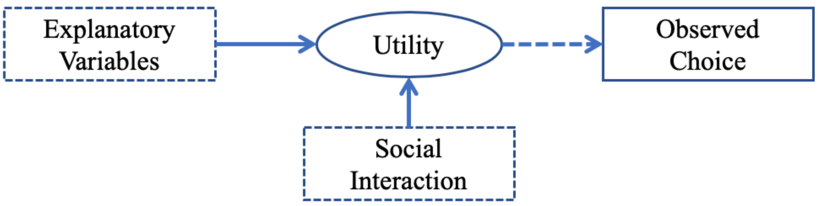

Figure 7. General Choice Model with Social Interaction

Sidharthan and Bhat (2012) formulated a multi-period discrete choice model considering spatial dependence for a land-use choice situation that incorporated both heterogeneity in the decision makers and the spatial "spillover" effect. The utility of individual q at period t with alternative i is U_{qti} which is a $(K \times 1)$ column vector defined as:

$$U_{qti} = \delta \sum_{q'} w_{qq'} U_{q'ti} + \tilde{\alpha}_{qi} + \beta'_q x_{qti} + \tilde{\varepsilon}_{qti} \quad \cdots\cdots \text{Eq. (2.3.1)}$$

where $w_{qq'}$ is qq'th element of weight matrix W which is a $(Q \times Q)$ matrix that indicates a distance-based spatial weight between individual q and q'. The weight matrix W is defined as a row normalized matrix with $w_{qq} = 0$ and $\sum_{q'} w_{qq'} = 1$ for $\forall q$. The spatial lag autoregressive parameter δ is bounded between 0 and 1, $0 < \delta < 1$. To define the heterogeneity of individuals, β_q is defined as $\beta_q = b + \breve{\beta}_q$ and $\breve{\beta}_q \sim MVN_K\left(0, \tilde{\Omega}\right)$, and $\tilde{\alpha}_{qi}$ is defined as $\tilde{\alpha}_{qi} = \tilde{a}_i + \breve{\alpha}_{qi}$. The matrix form of Eq. (2.3.1) is rearranged as:

$$U_{qti} = \left[S\left\{ \left(\mathbf{1}_{QT} \otimes \tilde{\mathbf{A}} \right) + \mathbf{xb} \right\} \right]_{d_{qti}} + \left[S\left\{ \breve{\alpha} + \tilde{\mathbf{x}}\breve{\beta} + \mathbf{C}\tilde{\eta} \right\} \right]_{d_{qti}} \quad \cdots\cdots \text{Eq. (2.3.2)}$$

where $S = \left[IDEN_{QTI} - \left\{ \left(\delta W \otimes IDEN_T \right) \otimes IDEN_I \right\} \right]^{-1}$ is a $(QTI \times QTI)$ matrix and $\mathbf{C} = \left[IDEN_{QTI} - IDEN_Q \otimes \left(\rho \mathbf{R} \otimes IDEN_I \right) \right]^{-1}$ is also a $(QTI \times QTI)$ matrix. $IDEN_E$ represents an identity matrix of size E, $\mathbf{1}_E$ represents a column vector of size E with all elements equal to 1, $[\cdot]_E$ represents the Eth element of the column vector $[\cdot]$. The model is estimated using the maximum approximation composite marginal likelihood approach, with the CML function written as:

$$L_{CML}(\theta) = \prod_{q=1}^{Q}\prod_{q'=q}^{Q}\prod_{t=1}^{T}\prod_{t'=t}^{T}\text{Prob}\left(C_{qt} = m_{qt}, C_{q't'} = m_{q't'}\right) \quad \text{with } q \neq q' \text{ when } t = t' \qquad \text{Eq. (2.3.3)}$$

The above marginal likelihood function examines the probability of individual q choosing alternative m at period t, and individual q' choosing alternative m at period t'. This model considers spatial dependence through an additive term which is a weighted sum of other individuals' utility and closeness of other individuals. The closeness is defined as the spatial distance between individuals which is appropriate for the land-use choice situation.

Castro, Paleti, and Bhat (2013) proposed a model that accommodates unobserved heterogeneity and spatial dependencies to analyze the severity of an injury. The dependent variable of interest is the number of crashes and the injury severity level which is an ordinal variable. Considering spatial dependence, the latent underlying injury risk propensity for crash q $(q = 1,2,\ldots,Q)$ with injury severity level k $(k = 1,2,\ldots,K)$ is defined as:

$$y_q^* = \delta\sum_{q'=1}^{Q} w_{qq'} y_{q'}^* + \beta_q' x_q + \varepsilon_q, \quad y_q = k \quad \text{if} \quad \psi_{q,k-1} < y_q^* < \psi_{q,k} \qquad \text{Eq. (2.3.4)}$$

The latent variable y_q^* is connected to observed injury level y_q, with thresholds $\psi_{q,k}$, where $-\infty < \psi_{q,1} < \psi_{q,2} < \cdots < \psi_{q,K-1} < \infty \ \ \forall q$, $\psi_{q,0} = -\infty$ and $\psi_{q,K} = \infty$. $w_{qq'}$ is the qq'th element of weight matrix W which is identical to the weight matrix in the study by Sidharthan and Bhat (2012). To allow for heterogeneity among observations,

β_q is defined as $\beta_q = b + \tilde{\beta}_q$, and $\tilde{\beta}_q$ follows a multivariate normal distribution. The matrix form for Eq. (2.3.4) is

$$y^* = S\left(xb + \tilde{x}\tilde{\beta} + \varepsilon\right) \quad \text{Eq. (2.3.5)}$$

where $S = \left[IDEN_Q - \delta W\right]^{-1}$ is a $\left(Q \times Q\right)$ matrix. The vector y^* follows, and the multivariate normal distribution with mean vector is $B_{LAG} = Sxb$, and covariance matrix is $\Sigma_{LAG} = S\left[\tilde{x}\left(IDEN_Q \otimes \Omega\right)\tilde{x}' + IDEN_Q\right]S'$. This study applied a similar approach to accommodate spatial dependence with an ordinal dependent variable.

Bhat (2015b) extended the traditional panel discrete choice model with inclusion of a spatial/social drift effect to consider an endogenous group formation phenomenon which is widely discussed in social interaction within economics. Consider the most basic utility function of individual q of alternative i at period t:

$$U_{qti} = \tilde{a}_i + b'x_{qti} + \tilde{\varepsilon}_{qti} \quad \text{Eq. (2.3.6)}$$

To incorporate social/spatial interaction, the utility function is defined as:

$$U_{qti} = \delta\sum_{q'} w_{qq'}U_{q'ti} + \tilde{\alpha}_{qi} + \beta_q' x_{qti} + \rho\tilde{\varepsilon}_{q,t-1,i} + \tilde{\eta}_{qti} \quad \text{Eq. (2.3.7)}$$

where $\tilde{\alpha}_{qi} = \tilde{a}_i + \breve{\alpha}_{qi}$ represents time-invariant individual specific unobserved preferences across individuals; $\beta_q = b + \tilde{\beta}_q$ represents response sensitivity across individuals; $\tilde{\varepsilon}_{qti} = \rho\tilde{\varepsilon}_{q,t-1,i} + \tilde{\eta}_{qti}$ represents cross alternative choice occasion-specific covariance and cross-time fading unobserved preference. To incorporate self-selection effects further define $\breve{\alpha}_{qi} = \theta \sum_{q'} w_{qq'} \breve{\alpha}_{q'i} + \tilde{\tau}_{qi}$, where θ $(0 < \theta < 1)$ is the self-selection parameter for capturing generic preferences. To incorporate a spatial structure into the unobserved preference, define $\beta_{qk} = b_k + \tilde{\beta}_{qk}$, where b_k is the general effect of the k th explanatory variable. Subsequently, define $\tilde{\beta}_{qk} = \lambda_k \sum_{q'} w_{qq'} \tilde{\beta}_{q'} k + \tilde{\gamma}_{qk}$, where λ_k $(0 < \lambda_k < 1)$ represents the self-selection effect of capturing unobserved sensitivities of the attributes. This model is also estimated using the maximum approximation composite marginal likelihood approach which is consistent with the study by Sidharthan and Bhat (2012). This study considered multiple types of spatial/social interaction within a single choice situation. One of the most basic assumptions in economics is that an individual's utility q, U_q, cannot be observed by others (including researchers and other individuals). Therefore, it is justified to doubt the reasonability of including other individuals' utility, $U_{q'}$, into individual q 's utility function U_q.

To improve the way of incorporating spatial/social interaction, Bhat, Pinjari, Dubey, and Hamdi (2016) combined a Generalized Heterogeneous Data Model (GHDM), which is capable of dealing with multiple types of dependent variables, with spatial/social dependency through latent constructs. The framework is shown in Figure 8. The solid arrows indicate the causal effect represented by structural equations, and the dashed arrows indicate underlying relations.

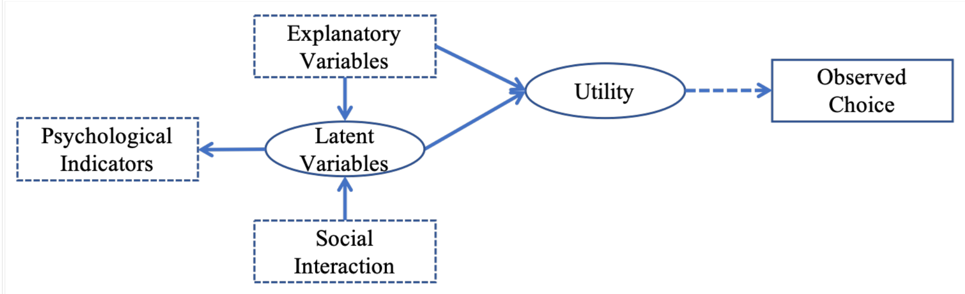

Figure 8. GHDM with Social Interaction

This model can be observed as an extension of the ICLV model mentioned in the previous section. There are two components in this GHDM model: (1) a latent variable structural equation model (SEM) system, and (2) a latent variable measurement equation model (MEM) system. There are two major improvements compared to previous ICLV models: (1) the inclusion of spatial/social interaction in the latent variable, and (2) the ability to capture endogeneity between multiple types of dependent variables. In the latent variable SEM system, consider the lth latent psychological variable of individual q, z_{ql}^*, as a linear function of exogeneous variables and spatial auto-correlation:

$$z_{ql}^* = \alpha_l' s_q + \eta_{ql} + \delta_l \sum_{q'=1}^{Q} w_{qq'} z_{q'l}^* \quad \cdots\cdots \text{Eq. (2.3.8)}$$

The matrix form of Eq. (2.3.8) can be written as:

$$z^* = S\tilde{s}\alpha + S\eta \quad \cdots\cdots \text{Eq. (2.3.9)}$$

where $S = \left[IDEN_{QL} - \tilde{\delta}.*\left(W \otimes IDEN_L\right)\right]^{-1}$ is a $\left(QL \times QL\right)$ matrix. z^* follows a multivariate normal distribution $z^* \sim MVN_{QL}\left(B, \Theta\right)$, where $B = S\tilde{s}\alpha$ and $\Theta = S\left[IDEN_Q \otimes \Gamma \right] S'$ (Γ is the correlation matrix of η_{ql}).

In the latent variable MEM system comprising H continuous outcomes, N ordinal outcomes, C count outcomes, and G nominal outcomes, define $E = H + N + C$, $\vec{G} = \sum_{g=1}^{G} I_g$, and $\tilde{G} = \sum_{g=1}^{G}\left(I_g - 1\right)$, where I_g indicates the number of alternatives of gth nominal outcome. The matrix form of MEM for individual q can be written as:

$$\begin{aligned} &\left(yU\right)_q = \vec{b}x_q + \vec{c}z_q^* + \xi_q \ , \\ &\text{with } \ Var\left(\xi_q\right) = \vec{\Sigma} \quad \left[\left(E + \vec{G}\right) \times \left(E + \vec{G}\right) \ \text{matrix}\right] \end{aligned} \quad \cdots\cdots\cdots\cdots\text{Eq. (2.3.10)}$$

The matrix form of MEM for all Q individuals can be defined as:

$$yU = \vec{b}x + \vec{c}z^* + \xi \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\text{Eq. (2.3.11)}$$

where $yU = \left[\left(yU\right)'_1, \left(yU\right)'_2, \ldots, \left(yU\right)'_Q\right]'$ is a $Q\left(E + \vec{G}\right) \times 1$ vector. Substitute Eq. (2.3.9) into Eq. (2.3.11). The reduced form of the whole model can be written as:

$$
\begin{aligned}
yU &= \vec{b}x + \vec{c}\left[S\vec{s}\alpha + S\eta\right] + \xi \\
&= \vec{b}x + \vec{c}\left[B + S\eta\right] + \xi \\
&= \left(\vec{b}x + \vec{c}B\right) + \left(\vec{c}S\eta + \xi\right)
\end{aligned}
\qquad \cdots\cdots \text{Eq. (2.3.12)}
$$

Subsequently, yU follows a multivariate normal distribution $yU \sim MVN_{Q(E+\tilde{G})}\left[\left(\vec{b}x + \vec{c}B\right), \left(\vec{c}\Theta\vec{c}' + IDEN_Q \otimes \vec{\Sigma}\right)\right]$. This model also can be estimated using the MACML approach by maximizing the marginal likelihood function $L_{CML}(\theta) = \prod_{q=1}^{Q-1}\prod_{q'=q+1}^{Q} L_{CML,qq'}(\theta)$, where $L_{CML,qq'}(\theta)$ can be defined as:

$$
\begin{aligned}
L_{CML,qq'}(\theta) = &\left(\prod_{h=1}^{2H}\omega_{\tilde{\Omega}_{qq',y}}\right)^{-1}\varphi_{2H}\left(\left[\omega_{\tilde{\Omega}_{qq',y}}\right]^{-1}\left[y_{qq'} - \tilde{B}_{qq',y}\right];\tilde{\Omega}_{qq',y}\right) \\
&\times\left(\prod_{v=1}^{2N+2C-1}\prod_{v'=v+1}^{2N+2C}\left[\begin{array}{l}\Phi_2\left(\mu_{v,up},\mu_{v',up},\rho_{vv'}\right) - \Phi_2\left(\mu_{v,up},\mu_{v',low},\rho_{vv'}\right) \\ -\Phi_2\left(\mu_{v,low},\mu_{v',up},\rho_{vv'}\right) - \Phi_2\left(\mu_{v,low},\mu_{v',low},\rho_{vv'}\right)\end{array}\right]\right) \\
&\times\left(\prod_{v=1}^{2N+2C}\prod_{g=1}^{2G}\begin{array}{l}\Phi_{I_g}\left[\omega^{-1}_{\hat{\Omega}_{qq',vg}}H_{vg}\left\{\vec{\psi}_{qq',up} - \vec{B}_{qq',\tilde{u}}\right\};\hat{\Omega}^{*}_{qq',vg}\right] \\ -\Phi_{I_g}\left[\omega^{-1}_{\hat{\Omega}_{qq',vg}}H_{vg}\left\{\vec{\psi}_{qq',low} - \vec{B}_{qq',\tilde{u}}\right\};\hat{\Omega}^{*}_{qq',vg}\right]\end{array}\right) \\
&\times\left(\prod_{g=1}^{2G-1}\prod_{g'=g+1}^{2G}\Phi_{I_g+I_{g'}-2}\left[\omega^{-1}_{\tilde{\Omega}_{qq',gg'}}Y_{gg'}\left\{-\vec{B}_{qq',\tilde{u}}\right\};\vec{\Omega}^{*}_{qq',gg'}\right]\right)
\end{aligned}
\qquad \cdot\text{Eq. (2.3.13)}
$$

The notation of Eq. (2.3.13) is specified in the online supplement of Bhat et al. (2016). This model accommodates spatial/social interaction through the latent psychological variable, and assume that the latent psychological variables have direct effects on endogenous outcome. The latent psychological variables are modeled as independent explanatory variables of the endogenous outcomes.

Vinayak et al. (2018) extended the concept of proximity-based dyadic interactions by introducing the idea of attitudes, habits, and lifestyle preferences as new dimensions and measures of proximity. This methodology is applied in this study to account for both interdependencies among decision makers in the spatial-attitudinal space and dynamics of self-selection due to inherent attitudes, preferences, and habits affecting the frequency with which individuals use car-sharing and ride-sourcing mobility services. The latent constructs reflecting attitudes, habits, and preferences are based on observed psychometric indicators and/or other variables describing observed behavior (e.g., smartphone ownership), and scores of these latent constructs are estimated using a GHDM.

Most economic choice models that have considered social interaction treat the social interaction term as an additive term to the expected utility function and do not consider the effect of social interaction on an individual's preference which is the most important element in the expected utility function. As discussed in the previous chapter, latent psychological variables should have an effect on dependent variables through interaction with independent variables, and not as independent variables themselves.

2.2.3 Models of Mixed Data

Individual behaviors are not just correlated with others' behavior but are also correlated with behaviors of the him/herself. Ignoring the dependency between the outcomes and estimating the outcomes separately may lead to inefficient and inconsistent estimation for each outcome. General Location Model (GLOM) assumes an arbitrary marginal distribution for the discrete outcomes, and a conditional normal distribution for the continuous outcome. However, GLOM cannot accommodate ordinal dependent variables as well as dependency between outcomes. Conditional Grouped Continuous Model (CGCM) assumes a latent variable for binary and ordinal outcomes, and a multivariate normal distribution for the continuous outcomes. The joint distribution is derived using a marginal distribution of the continuous outcomes and conditional distribution of the latent variables. However, since nominal outcomes cannot be partitioned into a single latent variable by thresholds, CGCM cannot incorporate such outcomes. General Mixed Data Model (GMDM) is an extension of CGCM with the incorporation of nominal outcomes. GMDM applies GLOM for the joint distribution of the nominal and continuous outcomes, and CGCM for the joint distribution of the ordinal and continuous outcomes. The dimension of GMDM explodes as both the number of nominal outcomes and the number of categories of each nominal outcome increase.

Factor Analysis, which is widely accepted by psychology studies, is able to deal with dependency among mixed outcomes by considering the outcomes as a function of unobserved psychological constructs. The latent constructs are defined as functions of exogenous variables, and the dependency between latent constructs is defined in measurement equations. The ICLV model extends the above model by including nominal outcomes. The ICLV model assumes a normal distribution for latent constructs, logistic distribution for the ordinal outcomes, and type-I extreme value error term for nominal outcomes. The integral of the likelihood function of the ICLV model is difficult to evaluate using traditional simulation techniques. Bhat (2015a) proposed a GHDM which is an SEM-like model. A GHDM is cable of jointly simulating mixed types of dependent variables. It is an extended version of CGCM which uses a latent continuous variable to represent all non-continuous outcomes.

If a decision maker's dependent outcomes are co-determined because of common underlying unobserved factors or psychological constructs (attitudes, values, lifestyles, etc.), it is very likely that social dependence will exist not just across one of those outcomes but across all the outcomes. There may be common underlying unobserved factors (attitudes, values, and lifestyle factors) among decision makers that simultaneously impact multiple dependent outcomes. Ignoring the dependency and considering each dimension separately may cause inefficient estimation of covariate effects for each outcome because such an approach fails to borrow information on other outcomes (Teixeira–Pinto & Harezlak, 2013).

In the case of non-continuous outcomes, accommodating social dependence, in general, leads to multidimensional integration of the order of the number of decision makers for ordered-response outcomes, and of the order of the number of decision makers times the number of alternatives minus one for unordered response outcomes. Typical simulation-based methods, such as frequentist recursive importance sampling (RIS) estimation and the Bayesian Markov Chain Monte Carlo (MCMC)-based estimator, are impractical if not infeasible when dealing with moderate to large estimation sample sizes (Bhat, 2011). Bhat and colleagues have suggested a composite marginal likelihood (CML) inference approach to estimate spatial binary/ordered response probit/count models. The maximum approximate composite marginal likelihood (MACML) approach is easy to implement, requires no simulation, and only involves univariate and bivariate cumulative normal distribution function evaluations.

Almost all previous spatial/social dependency model studies, (regardless of the estimation technique used) have focused on a single dependent outcome rather than multiple and mixed dependent outcomes for each decision maker. Bhat et al. (2016) incorporated spatial dependence into multiple mixed dependent outcomes including continuous and count outcomes, and ordered and unordered responses, pointing out that no previous study in the econometric literature has undertaken a spatial dependence analysis in the context of a relatively large mixed multidimensional model system. The spatial dependencies introduced in latent constructs permeate into all the endogenous outcomes influenced by the latent constructs. This approach obviates the need to incorporate spatial dependencies separately for each and every endogenous variable. This model considered the social interaction effect as an additive term of endogenous outcomes. Vinayak et al. (2018) defined psychological factors as proximity only, and did not consider direct social interaction effects in the multiple endogenous choice model.

2.2.3.1 Maximum Approximation Composite Marginal Likelihood

The dimensionality of the likelihood function of GHDM is also extremely high to literally evaluate it using traditional simulation techniques. As an alternative, the dimensionality of the integral in the CML function is independent from the number of latent variables, the number of nominal outcomes, and the number of categories of each nominal outcome. Moreover, with the CML, it is easy to derive a covariance structure with general inverse of a sandwich information matrix. Sidharthan and Bhat (2012) introduced the maximum approximate approach to simplify the estimation procedure in the CML function of GHDM.

There are two components in the MACML approach: the approximation method and the CML approach. The former is used to evaluate the multivariate standard normal cumulative distribution (MVNCD) function. The latter is an estimation method.

In the multinomial probit model, considering a choice situation with multiple alternatives, the probability of an individual choosing an alternative is a multivariate normal cumulative distribution function (MVNCD) with the dimension of the number of alternatives minus 1. Usually, the probability is approximated by maximum simulated likelihood (MSL) inference approach with the Geweke-Hajivassiliou-Keane (GHK) simulator or the Genz-Bretz (GB) simulator. Bayesian simulation using Markov Chain Monte Carlo (MCMC) techniques is also widely applied (McCulloch, Polson, & Rossi, 2000; Train, 2009). When the dimension of the integration increases, the MSL and Bayesian techniques require extensive simulation and encounter convergence assessment problems (Bhat, 2011).

An analytic approximation method is much more accurate and able to deal with high dimension integrations, which was first proposed by Solow (1990), and developed by Joe (1995). Bhat (2011) proposed an analytic approximation method to evaluate the MVNCD function which decomposes the function into a product of conditional probabilities. An analytic approximation only involves univariate and bivariate cumulative normal distribution function evaluation, independent from the number of categories within each nominal outcome. This approximation approach decomposes a joint probability into a bivariate marginal probability and univariate conditional probability. Subsequently, the conditional probability is approximated in a linear regression sense.

Instead of evaluating the whole MVNCD function, the CML maximizes a surrogate likelihood function which evaluates a set of observed marginal events. The events in the CML function are defined as pairwise observation across all or a subset of the outcomes.

2.3 Limitations of Previous Research and Research Motivation

Opinion dynamics studies comprehensively investigate the effect of social interaction on an individual's attitude and opinion formation. However, most of these studies do not consider the effect of changes in attitudes and opinions on an individual's behaviors. Although CODA models deal with individual behaviors, such behaviors in these models face no constraints and other exogeneous factors, which is different from general economic adoption behaviors.

Researchers who notice the importance of individual psychological characteristics in an individual's behavior attempt to incorporate such factors into choice models. Economic choice models that consider an individual's attitude and opinions, including ICLV models, usually do not pay attention to the effect of social interactions on individual characteristics such as individual attitude and opinions. On the other hand, choice models that consider the effect of social interaction only focus on the choice occasions of innovations with strong network externality effects. Moreover, previous studies that considered social interaction effects tended to model the social interaction term as an independent variable. To the best of the author's knowledge, there are only a few choice models that have considered the effect of social interaction on individual characteristics (such as attitudes and opinions), and no study has attempted to incorporate such individual characteristics as covariates of a choice model.

This study considers the social interaction effect on individual characteristics, and incorporates such characteristics into the choice model as covariates. The individual characteristics influence the final individual choices by interacting with explanatory variables, and are not considered as independent additive terms. The proposed model can capture the effect of social interaction in a more logical framework. The use of a social lag structure allows choice behavior of a decision maker to be influenced by that of his or her peers in the attitudinal space. The proposed model can assist in developing estimates of market adoption of emerging technologies as it captures the social interaction effects engendered by multiple sources. Policy strategies can be better informed via the various inter-dependency effects captured by the model. Agencies interested in greater adoption rates could identify attributes that are more affected by social interaction effects, and conduct marketing activities more efficiently.

Chapter 3. Model Specification

There are two components in the proposed model: (1) the latent psychological variable structural equation model, which is identical with the model proposed by Bhat et al. (2016), (2) the latent variable measurement equation model. The latent psychological variables (the first component) have direct effects on the endogenous outcome variables (the second component) based on psychological theories or related empirical studies. The latent psychological variables are defined as linear combination of exogenous observed variables that can reflect the individuals' attitude and opinions. In the latent variable measurement equation model, the endogenous outcome variables which represent individuals' choices are defined as linear combination of exogenous variables, latent psychological constructs, and other endogenous outcomes. The framework of the proposed model is shown in Figure 9. The solid arrows indicate the causal effect represented by structural equations and the dashed arrows indicate underlying relation represented by measurement equations. The two components are estimated jointly by maximum analytic composite marginal likelihood function.

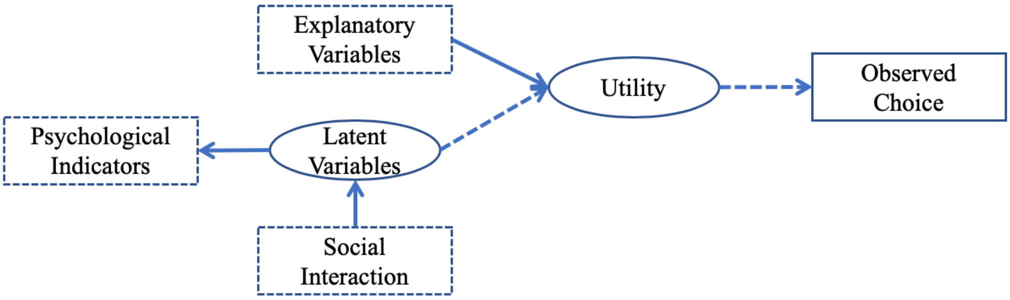

Figure 9. The Framework of The Proposed Model

The latent psychological variable structural equation model is introduced in section 3.1. Notice that since the model specification of latent psychological variable is identical with previous study, the specific equation definition is introduced in Appendix 1. In section 3.2, the case of single dependent variable is considered first, then the case of multiple dependent variables is introduced. Then estimation methodologies of both single dependent variable and multiple dependent variables are specified. The performance of the proposed model is verified in section 3.4 through simulation studies.

3.1 Latent Psychological Variable Structural Equation Model

This part is identical with the model proposed by Bhat et al. (2016) (See Appendix 1 for model specification). The l th latent psychological variable for individual q is noted as z_{ql}^{*} which can be defined as:

$$z_{ql}^{*} = \alpha_l' s_q + \delta_l \sum_{q'=1}^{Q} w_{qq'} z_{q'l}^{*} + \eta_{ql} \quad \text{Eq. (3.1.1)}$$

z^* follows an MVN distribution: $z^* \sim MVN_{QL}\left(B, \Xi\right)$, where $B = S\tilde{s}\alpha$ and $\Xi = S\left[IDEN_Q \otimes \Gamma\right]S'$. Note that $IDEN_E$ represents an identity matrix of size E and $\otimes$ indicates the Kronecker product.

3.2 Latent Variable Measurement Equation Model

3.2.1 Single Dependent Variable

Consider the dependent variable is a nominal outcome with $i = 1,2,\ldots,I$ alternatives. Define the observed outcome of individual q is m_q. The utility of individual q by choosing alternative i, U_{qi}, can be defined as:

$$U_{qi} = \left(\gamma_{qi}' + \left(d_i z_q^{*}\right)'\right) x_q + \varsigma_{qi} \quad \text{Eq. (3.2.1)}$$

where x_q is a $\left(A \times 1\right)$ vector of explanatory variables. γ_{qi} is a $\left(A \times 1\right)$ vector of corresponding coefficients of x_q and d_i is a $\left(A \times L\right)$ matrix captures the effect of

x_q through latent psychological variable with social interaction. $\varsigma_q = \left(\varsigma_{q1}, \varsigma_{q2}, \ldots, \varsigma_{qI}\right)'$ is a $\left(I \times 1\right)$ vector with $\varsigma_q \sim MVN_I\left(\mathbf{0}, \Lambda\right)$ and ς_q are independent across individuals. Since only the difference of utility of alternatives matters, define $\widehat{\Lambda}$ is an error term difference matrix respect to the first alternative with elements equals to: $\widehat{\varsigma}_q = \left(\widehat{\varsigma}_{q2}, \widehat{\varsigma}_{q3}, \ldots, \widehat{\varsigma}_{qI}\right)$, where $\widehat{\varsigma}_{qi} = \varsigma_{qi} - \varsigma_{q1} \quad \forall i \neq 1$. Further define the 1st element of the 1st row of $\widehat{\Lambda}$ as "1" to unify the scale. To derive the matrix form of Eq. (3.2.1) following terms are defined:

$U_q = \left(U_{q1}, U_{q2}, \ldots, U_{qI}\right)'$ a $\left(I \times 1\right)$ vector,

$\gamma_q = \left(\gamma_{q1}', \gamma_{q2}', \ldots, \gamma_{qI}'\right)'$ a $\left(I \times A\right)$ matrix,

$dz_q^* = \left(\left(d_1 z_q^*\right)', \left(d_2 z_q^*\right)', \ldots, \left(d_I z_q^*\right)'\right)'$ a $\left(I \times A\right)$ matrix,

where $z_q^* = \mathbf{1}_I \otimes z_q^*$ a $\left(L \times I\right)$ vector.

Then, the matrix form of individual q is:

$$U_q = \left(\gamma + dz_q^*\right) x_q + \varsigma_q \quad \cdots\cdots \text{Eq. (3.2.2)}$$

where $\varsigma_q \sim MVN_I\left(\mathbf{0}_I, \Lambda\right)$, and $\Lambda = \begin{bmatrix} 0 & 0 \\ 0 & \widehat{\Lambda} \end{bmatrix}$.

3.2.1.1 The Reduced Form

To derive the reduced form of the single dependent variable, define:

$U = \left(U_1', U_2', \ldots, U_Q'\right)'$ a $\left(QI \times 1\right)$ vector,

$\vec{\gamma} = IDEN_Q \otimes \gamma$ a $\left(QI \times QA\right)$ matrix,

$\vec{d}z^* = \left(IDEN_Q \otimes d\right) z^*$ a $\left(QI \times QA\right)$ matrix,

$x = \left(x_1', x_2', \ldots, x_Q'\right)'$ a $\left(QA \times 1\right)$ vector,

$\varsigma = \left(\varsigma_1', \varsigma_2', \ldots, \varsigma_Q'\right)'$ a $\left(QI \times 1\right)$ vector.

Define $\tilde{x} = \begin{bmatrix} x_1 & 0 & 0 & 0 \\ 0 & x_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & x_Q \end{bmatrix}$ a $\left(QA \times QI\right)$ matrix.

The reduced form for all individual q can be written as:

$$U = \left(\vec{\gamma} + \vec{d}z^*\right) x + \varsigma \quad \cdots\cdots \text{Eq. (3.2.3)}$$

Finally, substitute the z^* into the Eq. (3.2.3):

$$\begin{aligned} U &= \left(\vec{\gamma} + \vec{d}z^*\right) x + \varsigma \\ &= \left(\vec{\gamma} + \vec{d}\left\{S\tilde{s}\alpha + S\eta\right\}\right) x + \varsigma \quad \cdots\cdots \text{Eq. (3.2.4)} \\ &= \left(\vec{\gamma} + \vec{d}\left\{S\tilde{s}\alpha\right\}\right) x + \left(\left\{S\eta\right\} x + \varsigma\right) \end{aligned}$$

Then, $U \sim MVN_{QI}\left[B,\Omega\right]$, where $B=\left(\tilde{\gamma}+\vec{d}\left\{S\tilde{s}\alpha\right\}\right)x$ and $\Omega=\left(\vec{d}\tilde{x}\right)\cdot S\left(IDEN_Q\otimes\Gamma\right)S'\cdot\left(\vec{d}\tilde{x}\right)'+IDEN_Q\otimes\Lambda$.

3.2.2 Multiple Dependent Variables

The following section considers mixed types of dependent variables, including H continuous variables, N ordinal variables and G nominal variables. The multiple dependent variables are endogenously correlated with other dependent variables and effected by latent psychological variable where social interactions take place.

3.2.2.1 Continuous Outcome

There are H continuous outcomes with index $h=1,2,\ldots,H$. The h th continuous variable of individual q, y_{qh} can be defined as:

$$\breve{y}_{qh}=\left(\breve{\gamma}_h{}'+\left(\breve{d}_h z_q^*\right)'\right)x_q+\breve{\varepsilon}_{qh} \quad \text{Eq. (3.2.5)}$$

where x_q is a $\left(A\times1\right)$ vector of explanatory variables including observed exogenous variables as well as other endogenous dependent variables. $\breve{\gamma}_h$ is a $\left(A\times1\right)$ vector of corresponding coefficients of x_q captures pure effect of explanatory variables. $\breve{d}_h$ is a $(A \times L)$ matrix captures the effect of x_q through latent psychological variable with social interaction. $\breve{\varepsilon}_{qh}$ is a normally distributed random error term. To derive the matrix form of Eq. (3.2.5) following terms are defined:

$\breve{y}_q = \left(\breve{y}_{q1}, \breve{y}_{q2}, \ldots, \breve{y}_{qH}\right)'$ a $(H \times 1)$ vector,

$\breve{\gamma} = \left(\breve{\gamma}_1{}', \breve{\gamma}_2{}', \ldots, \breve{\gamma}_H{}'\right)'$ a $(H \times A)$ matrix,

$\breve{d}z_q^* = \left(\left(\breve{d}_1 z_q^*\right)', \left(\breve{d}_2 z_q^*\right)', \ldots, \left(\breve{d}_H z_q^*\right)'\right)'$ a $(H \times A)$ matrix,

where $z_q^* = \mathbf{1}_H \otimes z_q^*$ a $(L \times H)$ vector and $\mathbf{1}_H$ represents a $(H \times 1)$ column vector with all element equals to 1.

$\breve{\varepsilon}_q = \left(\breve{\varepsilon}_{q1}, \breve{\varepsilon}_{q2}, \ldots, \breve{\varepsilon}_{qH}\right)'$ a $(H \times 1)$ vector.

The matrix form of H continuous variables for individual q is:

$$\breve{y}_q = \left(\breve{\gamma} + \breve{d}z_q^*\right) x_q + \breve{\varepsilon}_q \quad \cdots\cdots\cdots \text{Eq. (3.2.6)}$$

Assume the error term $\breve{\varepsilon}_q$ follows a MVN distribution: $\breve{\varepsilon}_q \sim MVN_H\left(\mathbf{0}_H, \Sigma\right)$, the off-diagonal elements equal to 0 for identification purpose, and there is no correlation between $\breve{\varepsilon}_q$.

3.2.2.2 Ordinal Outcome

There are N ordinal outcomes with index $n = 1,2,\ldots,N$. $j_n = 1,2,\ldots,J_n$ represents the number of categories of n th ordinal variable, and $J_n \geq 2 \quad \forall n$. Define the observed outcome of individual q for ordinal variable n is a_{qn}. The latent variable of individual q for ordinal variable n, $\tilde{y}^*_{qn}$ can be defined as:

$$\tilde{y}^*_{qn} = \left(\tilde{\gamma}_n{}' + \left(\tilde{d}_n z^*_q \right)' \right) x_q + \tilde{\varepsilon}_{qn} \quad \cdots\cdots \text{Eq. (3.2.7)}$$

where $\tilde{\psi}_{q,n,a_{qn-1}} < \tilde{y}^*_{qn} < \tilde{\psi}_{q,n,a_{qn}}$. For each ordinal outcome: $\tilde{\psi}_{q,n,0} < \tilde{\psi}_{q,n,1} < \cdots < \tilde{\psi}_{q,n,J_n-1} < \tilde{\psi}_{q,n,J_n}$ and $\tilde{\psi}_{q,n,0} = -\infty$, $\tilde{\psi}_{q,n,1} = 0$, and $\tilde{\psi}_{q,n,J_n} = +\infty$. x_q is a $(A \times 1)$ vector of explanatory variables same as defined in continuous outcomes. $\tilde{\gamma}_n$ is a $(A \times 1)$ vector of corresponding coefficients of x_q and $\tilde{d}_h$ is a $(A \times L)$ matrix captures the effect of x_q through latent psychological variable with social interaction. $\tilde{\varepsilon}_{qh}$ is a normally distributed random error term. To derive the matrix form of Eq. (3.2.7) following terms are defined:

$\tilde{y}^*_q = \left(\tilde{y}^*_{q1}, \tilde{y}^*_{q2}, \ldots, \tilde{y}^*_{qN} \right)'$ a $(N \times 1)$ vector,

$\tilde{\gamma} = \left(\tilde{\gamma}_1{}', \tilde{\gamma}_2{}', \ldots, \tilde{\gamma}_H{}' \right)'$ a $(N \times A)$ matrix,

$\tilde{d}z_q^* = \left(\left(\tilde{d}_1 z_q^*\right)', \left(\tilde{d}_2 z_q^*\right)', \ldots, \left(\tilde{d}_N z_q^*\right)' \right)'$ a $(N \times A)$ matrix,

where $z_q^* = \mathbf{1}_N \otimes z_q^*$ a $(L \times N)$ vector,

$\tilde{\varepsilon}_q = \left(\tilde{\varepsilon}_{q1}, \tilde{\varepsilon}_{q2}, \ldots, \tilde{\varepsilon}_{qN}\right)'$ a $(N \times 1)$ vector.

$\tilde{\psi}_{q,low} = \left(\tilde{\psi}_{q,n,a_{qn-1}} \left(n = 1,2,\ldots,N\right)\right)$ a $(N \times 1)$ vector

$\tilde{\psi}_{q,up} = \left(\tilde{\psi}_{q,n,a_{qn}} \left(n = 1,2,\ldots,N\right)\right)$ a $(N \times 1)$ vector

The matrix form of N ordinal variables for individual q is:

$$\tilde{y}_q^* = \left(\tilde{\gamma} + \tilde{d}z_q^*\right) x_q + \tilde{\varepsilon}_q \quad \cdots\cdots \text{Eq. (3.2.8)}$$

where $\tilde{\psi}_{q,low} < \tilde{y}_q^* < \tilde{\psi}_{q,up}$. Assume the error term $\tilde{\varepsilon}_q$ follows a MVN distribution: $\tilde{\varepsilon}_q \sim MVN_N\left(\mathbf{0}_N, IDEN_N\right)$, the off-diagonal elements equal to 0, the diagonal elements equal to 1 for identification purpose, and there is no correlation between $\tilde{\varepsilon}_q$.

3.2.2.3 Nominal Outcome

There are G nominal outcomes with index $g = 1,2,\ldots,G$. $i_g = 1,2,\ldots,I_g$ represents the number of alternatives of g th nominal variable. Define the observed outcome of individual q for nominal variable g is m_{qg}. The utility of individual q for nominal variable g by choosing alternative i_g, U_{qgi_g} can be defined as:

$$U_{qgi_g} = \left(\hat{\gamma}_{gi_g}{}' + \left(\hat{d}_{gi_g} z_q^* \right)' \right) x_q + \varsigma_{qgi_g} \quad \cdots\cdots \text{Eq. (3.2.9)}$$

where x_q is a $(A \times 1)$ vector of explanatory variables same as defined previously. $\hat{\gamma}_{qgi_g}$ is a $(A \times 1)$ vector of corresponding coefficients of x_q and $\hat{d}_{gi_g}$ is a $(A \times L)$ matrix captures the effect of x_q through latent psychological variable with social interaction. $\varsigma_{qg} = \left(\varsigma_{qg1}, \varsigma_{qg2}, \ldots, \varsigma_{qgI_g} \right)'$ is a $(I_g \times 1)$ vector with $\varsigma_{qg} \sim MVN_{I_g}\left(\mathbf{0}, \Lambda_g \right)$ and ς_{qg} is independent across individuals. Since only the difference of utility of alternatives matters, define $\hat{\Lambda}_g$ is a error term difference matrix respect to the first alternative with elements equals to: $\hat{\varsigma}_{qg} = \left(\hat{\varsigma}_{qg2}, \hat{\varsigma}_{qg3}, \ldots, \hat{\varsigma}_{qgI_g} \right)$, where $\hat{\varsigma}_{qgi} = \varsigma_{qgi} - \varsigma_{qg1}$ $\forall i \neq 1$. Further define the 1st element of the 1st row of $\hat{\Lambda}_g \left(g = 1,2,\ldots,G \right)$ as "1" to unify the scale. To derive the matrix form of Eq. (3.2.9) following terms are defined:

$U_{qg} = \left(U_{qg1}, U_{qg2}, \ldots, U_{qgI_g} \right)'$ a $(I_g \times 1)$ vector,

$\hat{\gamma}_{qg} = \left(\hat{\gamma}_{qg1}{}', \hat{\gamma}_{qg2}{}', \ldots, \hat{\gamma}_{qgI_g}{}' \right)'$ a $(I_g \times A)$ matrix,

$\hat{d}_g z_q^* = \left(\left(\hat{d}_{g1} z_q^* \right)', \left(\hat{d}_{g2} z_q^* \right)', \ldots, \left(\hat{d}_{gI_g} z_q^* \right)' \right)'$ a $(I_g \times A)$ matrix,

where $z_q^* = \mathbf{1}_{I_g} \otimes z_q^*$ a $(L \times I_g)$ vector.

Further define: $\ddot{G}=\sum_{g=1}^{G} I_g$, $\tilde{G}=\sum_{g=1}^{G}\left(I_g-1\right)$,

$U_q=\left(U_{q1}{}',U_{q2}{}',\ldots,U_{qG}{}'\right)'$ a $\left(\ddot{G}\times 1\right)$ vector,

$\widehat{\gamma}=\left(\widehat{\gamma}_1{}',\widehat{\gamma}_2{}',\ldots,\widehat{\gamma}_G{}'\right)'$ a $\left(\ddot{G}\times A\right)$ matrix,

$\widehat{dz}_q^*=\left(\left(\widehat{d}_1 z_q^*\right)',\left(\widehat{d}_2, z_q^*\right)',\ldots,\left(\widehat{d}_G z_q^*\right)'\right)'$ a $\left(\ddot{G}\times A\right)$ matrix.

Then, the matrix form of individual q for G nominal variables is:

$$U_q=\left(\widehat{\gamma}+\widehat{dz}_q^*\right)x_q+\varsigma_q \quad \cdots\cdots\cdots \text{Eq. (3.2.10)}$$

where $\varsigma_q \sim MVN_{\ddot{G}}\left(\mathbf{0}_{\ddot{G}},\Lambda\right)$, and

$\Lambda=\begin{bmatrix}\Lambda_1 & & & \\ & \Lambda_2 & & \\ & & \ddots & \\ & & & \Lambda_G\end{bmatrix}$ a $\left(\ddot{G}\times\ddot{G}\right)$ matrix, $\Lambda_g=\begin{bmatrix}0 & 0\\ 0 & \widehat{\Lambda}_g\end{bmatrix}$.

3.2.2.4 The Reduced Form of Latent Measurement Equation Model

Let $E=H+N$ and define following notations:

$\left(yU\right)_q=\left(\breve{y}',\tilde{y}^{*\prime},U_q{}'\right)'$ a $\left(\left(E+\ddot{G}\right)\times 1\right)$ vector,

$\gamma = \left(\breve{\gamma}', \tilde{\gamma}', \hat{\gamma}'\right)'$ a $\left(\left(E+\vec{G}\right)\times A\right)$ matrix,

$dz_q^* = \left(\left(\breve{dz}_q^*\right)', \left(\tilde{dz}_q^*\right)', \left(\hat{dz}_q^*\right)'\right)'$ a $\left(\left(E+\vec{G}\right)\times A\right)$ matrix,

$\varepsilon_q = \left(\breve{\varepsilon}_q', \tilde{\varepsilon}_q', \hat{\varsigma}_q'\right)'$ a $\left(\left(E+\vec{G}\right)\times 1\right)$ vector.

Then the equation for all latent measurement variables of individual q is:

$$\left(yU\right)_q = \left(\gamma + dz_q^*\right)x_q + \varepsilon_q \quad \cdots\cdots \text{Eq. (3.2.11)}$$

where the error term ε_q can be written as:

$$\mathrm{Var}\left(\varepsilon_q\right) = \begin{bmatrix} \breve{\Sigma} & & \\ & IDEN_N & \\ & & \Lambda \end{bmatrix} \text{ a } \left(\left(E+\vec{G}\right)\times\left(E+\vec{G}\right)\right) \text{ matrix.}$$

The reduced form of all outcomes for all individual q can be written as:

$$yU = \left(\vec{\gamma} + \vec{dz}^*\right)x + \varepsilon \quad \cdots\cdots \text{Eq. (3.2.12)}$$

where $yU = \left(\left(yU\right)_1', \left(yU\right)_2', \ldots, \left(yU\right)_Q'\right)'$ a $\left(Q\left(E+\vec{G}\right)\times 1\right)$ vector,

$\vec{\gamma} = IDEN_Q \otimes \gamma$ a $\left(Q\left(E+\vec{G}\right)\times QA\right)$ matrix,

$\vec{dz}^* = \left(IDEN_Q \otimes d\right)z^*$ a $\left(Q\left(E+\vec{G}\right)\times QA\right)$ matrix,

$x = \left(x_1', x_2', \ldots, x_Q'\right)'$ a $\left(QA\times 1\right)$ vector,

$\varepsilon = \left(\varepsilon_1{}', \varepsilon_2{}', \ldots, \varepsilon_Q{}'\right)'$ a $\left(Q\left(E + \vec{G}\right) \times 1\right)$ vector.

Define $\tilde{x} = \begin{bmatrix} x_1 & 0 & 0 & 0 \\ 0 & x_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & x_Q \end{bmatrix}$ a $\left(QA \times Q\left(E + \vec{G}\right)\right)$ matrix.

Finally, substitute the z^* into the Eq. (3.2.12):

$$\begin{aligned} yU &= \left(\vec{\gamma} + \vec{d}z^*\right)x + \varepsilon \\ &= \left(\vec{\gamma} + \vec{d}\left\{S\tilde{s}\alpha + S\eta\right\}\right)x + \varepsilon \\ &= \left(\vec{\gamma} + \vec{d}\left\{S\tilde{s}\alpha\right\}\right)x + \left(\left\{S\eta\right\}x + \varepsilon\right) \end{aligned} \qquad \text{Eq. (3.2.13)}$$

Then, $yU \sim MVN_{Q\left(E+\vec{G}\right)}\left[B, \Omega\right]$, where $B = \left(\vec{\gamma} + \vec{d}\left\{S\tilde{s}\alpha\right\}\right)x$ and $\Omega = \left(\vec{d}\tilde{x}\right) \cdot S\left(IDEN_Q \otimes \Gamma\right)S' \cdot \left(\vec{d}\tilde{x}\right)' + IDEN_Q \otimes \Sigma$.

3.3 Estimation Methodology

Stack all parameters into vector $\lambda = \left[Vech(\alpha), Vech(\Sigma), Vech(\gamma), Vech(\delta), \psi\right]$, where $Vech(\cdot)$ means vectorize the elements of the matrix/vector. The likelihood function can be written as (See Appendix 2 for specification):

$$L(\lambda) = f_{QH}\left(y \middle| \tilde{B}_y, \tilde{\Omega}_y\right) \times \Pr\left[\vec{\psi}_{low} < \tilde{u} < \vec{\psi}_{up}\right]$$
$$= f_{QH}\left(y \middle| \tilde{B}_y, \tilde{\Omega}_y\right) \times \int_r f_{Q(N+\tilde{G})}\left(r \middle| \vec{B}_{\tilde{u}}, \vec{\Omega}_{\tilde{u}}\right) dr \quad \text{Eq. (3.3.1)}$$

The first component of the above likelihood function is a multivariate density function of dimension QH (with mean $\tilde{B}_y$ and covariance matrix $\tilde{\Omega}_y$) and the second component is a integral to evaluate the conditional likelihood of all dis-continuous outcomes with dimension $Q\left(N+\tilde{G}\right)$. As mentioned in the Chapter 2, it is almost impossible to evaluate such high dimension integral in traditional estimation technique. The maximum approximate composite marginal likelihood (MACML) approach is feasible for the proposed model. In the composite marginal likelihood (CML) approach, the likelihood function is divided into a product of low dimensional marginal densities.

For the proposed model, the CML can be defined as a product of pairwise marginal densities across all pairs of individuals:

$$L_{CML}(\lambda) = \prod_{q=1}^{Q-1} \prod_{q'=q+1}^{Q} L_{CML,qq'}(\lambda) \quad \text{Eq. (3.3.2)}$$

The MACML of the proposed model with single dependent variable can be defined as:

$$L_{MACML,qq'}(\lambda) = \prod_{q'=1}^{Q-1} \prod_{q'=q}^{Q} \Pr\left(C_q = m_q, C_{q'} = m_{q'}\right) \text{ with } q \neq q' \quad \text{Eq. (3.3.3)}$$

where C_q represents the choice of individual q.

The MACML of the proposed model with multiple dependent variables can be defined as (See Appendix 3 for specification):

$$
\begin{aligned}
L_{MACML,qq'}(\lambda) = & \left(\prod_{h=1}^{H} \omega_{\tilde{\Omega}_{qq',y}}\right)^{-1} \phi_H\left(\left[\omega_{\tilde{\Omega}_{qq',y}}\right]^{-1}\left[y_{qq'} - \tilde{B}_{qq',y}\right]; \tilde{\Omega}^{*}_{qq',y}\right) \\
& \times \left(\prod_{v=1}^{N-1}\prod_{v'=v+1}^{N}\left[\begin{array}{l} \Phi_2\left(\mu_{v,up},\mu_{v',up},\rho_{vv'}\right) - \Phi_2\left(\mu_{v,up},\mu_{v',low},\rho_{vv'}\right) \\ -\Phi_2\left(\mu_{v,low},\mu_{v',up},\rho_{vv'}\right) + \Phi_2\left(\mu_{v,low},\mu_{v',low},\rho_{vv'}\right) \end{array}\right]\right) \\
& \times \left(\prod_{v=1}^{N}\prod_{g=1}^{G}\left[\begin{array}{l} \Phi_{I_g}\left[\omega^{-1}_{\hat{\Omega}_{qq',vg}} H_{vg}\left\{\vec{\psi}_{qq',up} - \vec{B}_{qq',\tilde{u}}\right\}; \hat{\Omega}^{*}_{qq',vg}\right] \\ -\Phi_{I_g}\left[\omega^{-1}_{\hat{\Omega}_{qq',vg}} H_{vg}\left\{\vec{\psi}_{qq',low} - \vec{B}_{qq',\tilde{u}}\right\}; \hat{\Omega}^{*}_{qq',vg}\right] \end{array}\right]\right)
\end{aligned}
\quad \text{Eq. (3.3.4)}
$$

In the MACML of individual pair qq', the first component indicates the marginal likelihood of the continuous outcomes, the second component indicates the likelihood of pairs of outcomes between ordinal outcomes, the third component indicates the likelihood of pairs of outcomes between ordinal outcomes and nominal outcomes. (See Appendix 4 for estimation approach for covariance matrix)

3.4 Simulation Study

3.4.1 Simulation Design

There are total 1,000 individuals in the simulation study with 3 latent psychological variables. In case of multiple dependent variable, 1 continuous outcome, 2 ordinal outcome and 1 nominal outcome with 3 alternatives are considered. In case of single dependent variable, only the nominal outcome is considered with identical true value settings.

Assume there are 3 latent psychological variables with 2 explanatory variable each and no explanatory variable effect on more than one latent psychological variable. The elements in α matrix and L_{Γ} matrix are need to be estimated. Arrange the elements of α into vector $Vech(\alpha) = [\alpha_1 = 0.4, \alpha_2 = 0.4, \alpha_3 = 0.3, \alpha_4 = 0.3, \alpha_5 = 0.3, \alpha_6 = 0.5]$ and off diagonal elements of L_{Γ} as $l_{\Gamma}(l_{\Gamma 1} = 0.3,\ l_{\Gamma 2} = 0.3,\ l_{\Gamma 3} = 0.22)$. Since matrix Γ is a correlation matrix only the off diagonal elements of Cholesky decomposed matrix need to be estimated.

$$\begin{bmatrix} z_1^* \\ z_2^* \\ z_3^* \end{bmatrix} = \begin{bmatrix} 0.4 & 0.4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.3 & 0.3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.3 & 0.5 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix} \quad \cdots\cdots\cdots\cdots\cdots \text{Eq. (3.4.1)}$$

and

$$
\Gamma = \begin{bmatrix} 1 & 0.3 & 0.3 \\ 0.3 & 1 & 0.3 \\ 0.3 & 0.3 & 1 \end{bmatrix}
= L_{\Gamma} L_{\Gamma}{}' = \begin{bmatrix} 1 & 0.3 & 0.3 \\ 0 & 0.95 & 0.22 \\ 0 & 0 & 0.93 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0.3 & 0.95 & 0 \\ 0.3 & 0.22 & 0.93 \end{bmatrix} \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. (3.4.2)}
$$

Next, assume there are 1 continuous outcome and 2 ordinal outcomes with 1 explanatory variable each for simplification. The elements in vector $\gamma = \left(\breve{\gamma}_1 = 0.5, \tilde{\gamma}_1 = 0.5, \tilde{\gamma}_2 = 0.5\right)$ and non-zero elements of vectorized matrix $d = \left(\breve{d}_{11} = 0.3, \breve{d}_{13} = 0.5, \tilde{d}_{11} = 0.5, \tilde{d}_{21} = 0.2, \tilde{d}_{22} = 0.5, \tilde{d}_{23} = 0.2\right)$ are need to be estimated. The first element of Σ matrix which corresponding to the continuous outcome $\breve{\varepsilon}_1 = 1$ is the parameter need to be estimated. The error terms of ordinal outcomes are fixed to 1 and only the upper thresholds $\psi = \left(\psi_1 = 4, \psi_2 = 4\right)$ are estimated.

$$
\begin{bmatrix} \breve{y}_1 \\ \tilde{y}_1^* \\ \tilde{y}_2^* \end{bmatrix} = \left[\begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \end{bmatrix} + \begin{bmatrix} 0.3 & 0.0 & 0.5 \\ 0.5 & 0.0 & 0.0 \\ 0.2 & 0.5 & 0.2 \end{bmatrix} \begin{bmatrix} z_1^* \\ z_2^* \\ z_3^* \end{bmatrix} \right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} \breve{\varepsilon}_1 \\ \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \end{bmatrix} \qquad \cdots\cdots\cdots \text{Eq. (3.4.3)}
$$

Finally, assume there are 1 nominal outcome with 3 alternatives and 2 alternative specific variables. The elements in vector $\gamma = \left(\hat{\gamma}_1 = 0.5, \hat{\gamma}_2 = -1.0, \hat{\gamma}_3 = -1.0, \hat{\gamma}_4 = -0.8\right)$ and non-zero elements of vectorized matrix $d = \left(\hat{d}_{22} = 0.3, \hat{d}_{31} = 0.3, \hat{d}_{33} = 0.3\right)$ are need to be estimated. Also, $l_\Lambda = \left(l_{\Lambda 1} = 0.8, l_{\Lambda 2} = 1.0 \right)$, the off diagonal elements in the Cholesky decomposed Λ matrix are the parameters need estimate.

$$\begin{bmatrix} U_{11} \\ U_{12} \\ U_{13} \end{bmatrix} = \left[\begin{bmatrix} 0.5 \\ -1.0 \\ -1.0 \\ -0.8 \end{bmatrix} + \begin{bmatrix} 0.0 & 0.0 & 0.0 \\ 0.0 & 0.3 & 0.0 \\ 0.3 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.3 \end{bmatrix} \begin{bmatrix} z_1^* \\ z_2^* \\ z_3^* \end{bmatrix} \right] \times \begin{bmatrix} 0 & 0 & x_{11} & x_{21} \\ 1 & 0 & x_{12} & x_{22} \\ 0 & 1 & x_{13} & x_{23} \end{bmatrix} + \begin{bmatrix} \varsigma_1 \\ \varsigma_2 \\ \varsigma_3 \end{bmatrix} \quad \text{Eq. (3.4.4)}$$

and

$$\Lambda = \begin{bmatrix} 0.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.8 \\ 0.0 & 0.8 & 1.64 \end{bmatrix} = L_\Lambda L_\Lambda{}' = \begin{bmatrix} 0.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 0.0 & 0.8 & 1.0 \end{bmatrix} \begin{bmatrix} 0.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.8 \\ 0.0 & 0.0 & 1.0 \end{bmatrix} \quad \text{Eq. (3.4.5)}$$

The value of exogeneous variables $\left(s, x \right)$ are drawn from independent uniform distribution. The indicator of ordinal outcome and chosen alternative is simulated with true values.

3.4.2 Simulation Results

The simulation study is done with sample size of 1,000 and with 21 parameters: α_1, α_2, α_3, α_4, α_5, α_6, $l_{\Gamma 1}$, $l_{\Gamma 2}$, $l_{\Gamma 3}$, $\hat{d}_{22}$, $\hat{d}_{31}$, $\hat{d}_{33}$, $l_{\Lambda 1}$, $l_{\Lambda 2}$, δ_1, δ_2, δ_3, $\hat{\gamma}_1$, $\hat{\gamma}_2$, $\hat{\gamma}_3$, $\hat{\gamma}_4$ for the case of single dependent variable. In case of multiple dependent variables there are totally 33 parameters: α_1, α_2, α_3, α_4, α_5, α_6, $l_{\Gamma 1}$, $l_{\Gamma 2}$, $l_{\Gamma 3}$, $\breve{d}_{11}$, $\breve{d}_{13}$, $\tilde{d}_{11}$, $\tilde{d}_{21}$, $\tilde{d}_{22}$, $\tilde{d}_{23}$, $\hat{d}_{22}$, $\hat{d}_{31}$, $\hat{d}_{33}$, $\breve{\varepsilon}_1$, $l_{\Lambda 1}$, $l_{\Lambda 2}$, ψ_1, ψ_2, δ_1, δ_2, δ_3, $\breve{\gamma}_1$, $\tilde{\gamma}_1$, $\tilde{\gamma}_2$, $\hat{\gamma}_1$, $\hat{\gamma}_2$, $\hat{\gamma}_3$, $\hat{\gamma}_4$. In order to verify how can the model covers the true value, I conduct 3 simulation studies with different initial values for both single dependent variable case and multiple dependent variables case.

The true value of the parameter and the estimates corresponding to each initial value are indicated in the Table 1 for the single dependent variable case and Table 2 for the multiple dependent variable case. From the simulation result we can confirm that the model can covers the true value from the initial value well with limited sample size.

The errors of the estimates are generally around 1% to 5% some are over 10%. Notice that the results are from single realization simulations. There is some randomness within the simulation process. Therefore, it is literally impossible to 100% recover the true value. Moreover, the simulations conducted in this section set the initial values quite far away from the true value. The model can recover the true value in 1% to 5% error if the initial values are not too far from the true values. Unfortunately, with consideration of social dependence, if the initial value is too far from the true value (for example -1 true value with 0.1 initial value) the model then needs extensive simulation to recover the true value.

Furthermore, the simulations are done with only 1,000 synthesize individuals. Increase the simulation sample size can also increase the performance of the model.

Table 1. Simulation Results for the Single Dependent Variable

Parameter	true value	Simulation 1			Simulation 2			Simulation 3		
		Init.	**Est.**	**Err.**	**Init.**	**Est.**	**Err.**	**Init.**	**Est.**	**Err.**
α_1	0.4	0.1	0.361	0.039	0.2	0.411	0.011	0.1	0.394	0.006
α_2	0.4	0.1	0.426	0.026	0.2	0.386	0.014	0.1	0.347	0.053
α_3	0.3	0.1	0.331	0.031	0.2	0.297	0.003	0.1	0.326	0.026
α_4	0.3	0.1	0.247	0.053	0.2	0.267	0.033	0.1	0.317	0.017
α_5	0.3	0.1	0.271	0.029	0.2	0.331	0.031	0.1	0.261	0.039
α_6	0.5	0.1	0.594	0.094	0.2	0.491	0.009	0.1	0.527	0.027
$l_{\Gamma 1}$	0.3	0.1	0.209	0.091	0.1	0.273	0.027	0.1	0.291	0.009
$l_{\Gamma 2}$	0.3	0.1	0.340	0.040	0.1	0.323	0.023	0.1	0.314	0.014
$l_{\Gamma 3}$	0.3	0.1	0.308	0.008	0.1	0.281	0.019	0.1	0.323	0.023
$\widehat{d}_{22}$	0.5	0.1	0.551	0.051	0.1	0.457	0.043	0.2	0.476	0.024
$\widehat{d}_{31}$	0.5	0.1	0.489	0.011	0.1	0.536	0.036	0.2	0.459	0.041
$\widehat{d}_{33}$	0.5	0.1	0.476	0.024	0.1	0.519	0.019	0.2	0.531	0.031

(Table continue)

Para meter	true value	Simulation 1			Simulation 2			Simulation 3		
		Init.	Est.	Err.	Init.	Est.	Err.	Init.	Est.	Err.
$l_{\Lambda 1}$	0.8	0.1	0.971	0.171	0.5	1.034	0.234	0.5	0.991	0.191
$l_{\Lambda 2}$	1.64	0.1	1.214	0.426	1.25	1.915	0.275	1.25	1.694	0.054
δ_1	0.2	0.1	0.196	0.004	0.1	0.223	0.023	0.4	0.176	0.024
δ_2	0.3	0.1	0.331	0.031	0.1	0.341	0.041	0.4	0.355	0.055
δ_3	0.2	0.1	0.261	0.061	0.1	0.197	0.003	0.4	0.189	0.011
$\hat{\gamma}_1$	0.5	0.1	0.516	0.016	0.1	0.561	0.061	0.2	0.543	0.043
$\hat{\gamma}_2$	-1	0.1	-0.921	0.079	0.1	-0.981	0.019	0.2	-1.421	0.421
$\hat{\gamma}_3$	-1	0.1	-1.239	0.239	0.1	-0.914	0.086	0.2	-1.294	0.294
$\hat{\gamma}_4$	-0.8	0.1	-0.694	0.106	0.1	-0.931	0.131	0.2	-0.987	0.187

Table 2. Simulation Results for the Multiple Dependent Variables

Para meter	true value	Simulation 1			Simulation 2			Simulation 3		
		Init.	**Est.**	**Err.**	**Init.**	**Est.**	**Err.**	**Init.**	**Est.**	**Err.**
α_1	0.4	0.1	0.445	0.045	0.2	0.428	0.028	0.1	0.407	0.007
α_2	0.4	0.1	0.419	0.019	0.2	0.38	0.02	0.1	0.387	0.013
α_3	0.3	0.1	0.372	0.072	0.2	0.267	0.033	0.1	0.299	0.001
α_4	0.3	0.1	0.299	0.001	0.2	0.287	0.013	0.1	0.241	0.059
α_5	0.3	0.1	0.248	0.052	0.2	0.262	0.038	0.1	0.293	0.007
α_6	0.5	0.1	0.663	0.163	0.2	0.562	0.062	0.1	0.471	0.029
$l_{\Gamma 1}$	0.3	0.1	0.323	0.023	0.1	0.324	0.024	0.1	0.399	0.099
$l_{\Gamma 2}$	0.3	0.1	0.333	0.033	0.1	0.331	0.031	0.1	0.393	0.093
$l_{\Gamma 3}$	0.3	0.1	0.342	0.042	0.1	0.339	0.039	0.1	0.384	0.084
$\breve{d}_{11}$	0.3	0.1	0.332	0.032	0.1	0.295	0.005	0.2	0.265	0.035
$\breve{d}_{13}$	0.5	0.1	0.574	0.074	0.1	0.51	0.01	0.2	0.486	0.014
$\tilde{d}_{11}$	0.5	0.1	0.551	0.051	0.1	0.513	0.013	0.2	0.499	0.001
$\tilde{d}_{21}$	0.2	0.1	0.143	0.057	0.1	0.156	0.044	0.2	0.197	0.003

(Table continue)

Para meter	true value	Simulation 1			Simulation 2			Simulation 3		
		Init.	**Est.**	**Err.**	**Init.**	**Est.**	**Err.**	**Init.**	**Est.**	**Err.**
$\tilde{d}_{22}$	0.5	0.1	0.507	0.007	0.1	0.512	0.012	0.2	0.509	0.009
$\tilde{d}_{23}$	0.2	0.1	0.135	0.065	0.1	0.229	0.029	0.2	0.186	0.014
$\widehat{d}_{22}$	0.5	0.1	0.606	0.106	0.1	0.481	0.019	0.2	0.493	0.007
$\widehat{d}_{31}$	0.5	0.1	0.58	0.08	0.1	0.527	0.027	0.2	0.421	0.079
$\widehat{d}_{33}$	0.5	0.1	0.504	0.004	0.1	0.466	0.034	0.2	0.421	0.079
$\breve{\varepsilon}_1$	1	1	1.638	0.638	1	1.561	0.561	1	1.591	0.591
$l_{\Lambda 1}$	0.8	0.1	0.624	0.176	0.5	0.773	0.027	0.5	0.761	0.039
$l_{\Lambda 2}$	1.64	0.1	0.962	0.678	1.25	0.985	0.655	1.25	1.103	0.537
ψ_1	4	0	4.582	0.582	1	4.55	0.55	1	4.141	0.141
ψ_2	4	1	4.56	0.56	2	4.413	0.413	2	4.291	0.291
δ_1	0.2	0.1	0.182	0.018	0.1	0.176	0.024	0.4	0.179	0.021
δ_2	0.3	0.1	0.281	0.019	0.1	0.277	0.023	0.4	0.276	0.024
δ_3	0.2	0.1	0.177	0.023	0.1	0.265	0.065	0.4	0.177	0.023

(Table continue)

Para meter	**true value**	**Simulation 1**			**Simulation 2**			**Simulation 3**		
		Init.	**Est.**	**Err.**	**Init.**	**Est.**	**Err.**	**Init.**	**Est.**	**Err.**
$\breve{\gamma}_1$	0.5	0.1	0.547	0.047	0.1	0.562	0.062	0.2	0.504	0.004
$\tilde{\gamma}_1$	0.5	0.1	0.551	0.051	0.1	0.529	0.029	0.2	0.534	0.034
$\tilde{\gamma}_2$	0.5	0.1	0.51	0.01	0.1	0.485	0.015	0.2	0.536	0.036
$\hat{\gamma}_1$	0.5	0.1	0.636	0.136	0.1	0.537	0.037	0.2	0.481	0.019
$\hat{\gamma}_2$	-1	0.1	-0.995	0.005	0.1	-0.906	0.094	0.2	-0.905	0.095
$\hat{\gamma}_3$	-1	0.1	-0.939	0.061	0.1	-0.967	0.033	0.2	-0.934	0.066
$\hat{\gamma}_4$	-0.8	0.1	-0.891	0.091	0.1	-0.819	0.019	0.2	-0.914	0.114

Chapter 4. Empirical Study

4.1 Empirical Study Background and Specification

From Advanced Mobile Phone System (AMPS) so called 1G, to Global System for Mobile Communications (GSM) and CDMA known as 2G and advanced 3G, to Long Term Evolution (LTE) be widely adopted nowadays, with the development of technology, telecommunication service has become one of the most important elements within our daily life. April 2019 South Korea have launched the very first fifth generation (5G) of telecommunication service. Diffusion of 5G would accelerate: 1) hyper real-time processing service, 2) augmented reality/virtual reality service, 3) hyper-connectivity telecommunication service. Moreover, 5G service not just effect on B2C (Business to Consumer) market but also effect on B2B (Business to Business) market such as IoT. 5G service is one of the core infrastructures of the fourth industrial revolution (Kim, 2017). These new services might act as drive engine of economic growth of following decades (Jeong, Hong, & Ji, 2020). In order to develop the future services, the diffusion of 5G service is necessary. The adoption of 5G for individuals are correlated with many other factors such as mobile phone usage and telecommunication expenditure. On the other hand, the adoption behaviors in the early diffusion stage are highly affected by individual characteristics. Therefore, the proposed model is suitable for 5G adoption related situation.

The proposed Generalized Heterogeneous Data Model (GHDM) with Social interaction is capable of estimating multiple endogenous outcome simultaneously with consideration of social interaction. This is very useful in complicated choice situations especially in situations that associated with multiple choices. Under the situation of 5G introduction, the empirical study is focused on 5G adoption behavior of individuals. The framework of the empirical study is shown in Figure 10. The empirical study assumes 2 latent psychological variables effect on 4 endogenously correlated heterogenous outcomes. More specifically there are 1 continuous outcome, 2 ordinal outcomes and 1 nominal outcome.

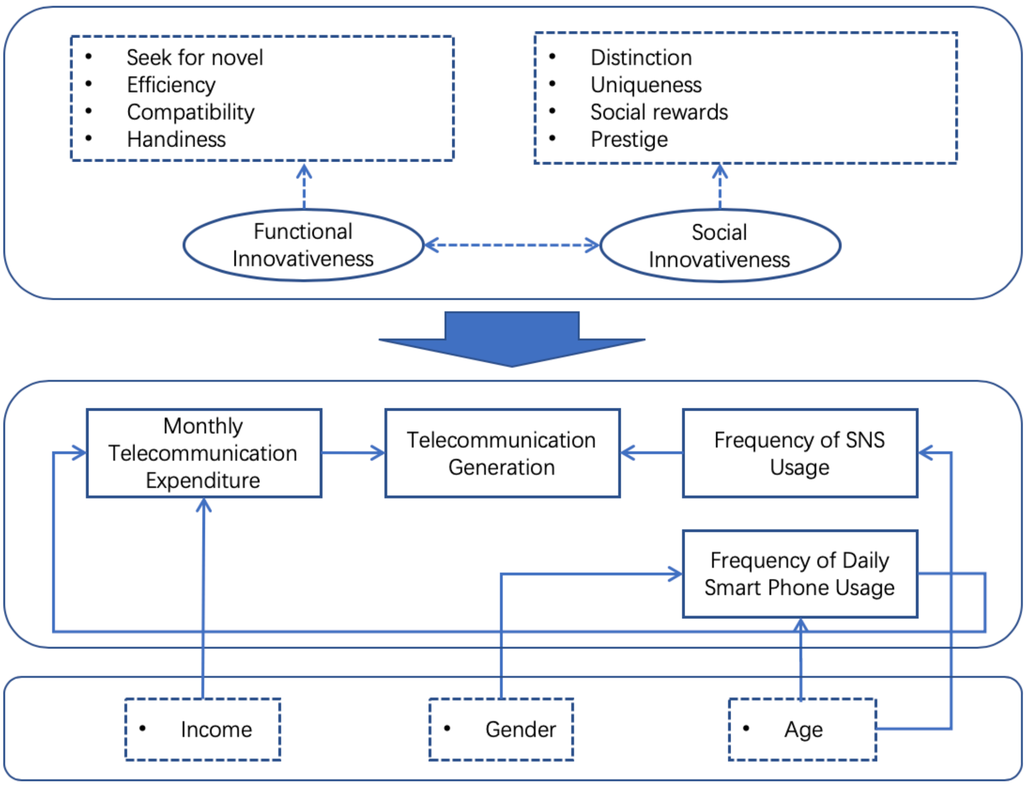

Figure 10. Empirical Study Framework

The two latent psychological variables are functional innovativeness and social innovativeness respectively which are depicted in the oval on the top. Each latent psychological variable is represented by four exogeneous variables. The specific definitions and reliabilities of the latent psychological variables are discussed in section 4.1.1. The continuous outcome of the measurement equation is monthly telecommunication expenditure. The ordinal outcomes are the frequency of SNS usage and daily smart phone usage with 3 levels. The nominal outcome is telecommunication generation choice with three alternatives: 3G, 4G and 5G. The explanatory variables for the measurement equation model are income, gender and age. The specific relationship between outcomes are discussed in section 4.1.2. The solid arrows indicate the causal effects and the dashed arrows indicates the latent effects between latent psychological variables. The latent psychological variables are assumed to affecting on all outcomes in the measurement equation model.

4.1.1 Latent Psychological Variables

Individual innovativeness has been proved that has essential effect on innovation adoption in consumer choice studies and innovation diffusion studies. Especially high-tech related adoption behaviors are highly affected by individual innovativeness. There are a lot of studies investigated about the definition and measurement of individual innovativeness. Vandecasteele and Geuens (2010) divide individual innovativeness into: functional innovativeness, social innovativeness, hedonic innovativeness, and cognitive innovativeness. This segmentation offers better understand of motivation of individual behaviors.

Functional innovativeness is motivated by the usefulness, compatibility, efficiency and comfort of innovation. Social innovativeness is motivated by uniqueness, distinction, social rewards and prestige of innovation adoption. Functional innovativeness and social innovativeness are measured by 4 questions with 5 Likert scale each. Since the latent psychological variables are measured in positive Likert scale the parameter of social interaction δ are naturally bounded as positive values. The two latent psychological variables used in the empirical studies are specified as follows.

$$z^*_{1,\text{func.ino.}} = \begin{bmatrix} \alpha_{11} \\ \alpha_{12} \\ \alpha_{13} \\ \alpha_{14} \end{bmatrix}' \begin{bmatrix} \text{novel} \\ \text{efficiency} \\ \text{compatibility} \\ \text{handiness} \end{bmatrix} + \delta_1 W z_1^* + \eta_1 \quad \cdots\cdots \text{Eq. (4.1.1)}$$

$$z^*_{2,\text{soc.ino.}} = \begin{bmatrix} \alpha_{21} \\ \alpha_{22} \\ \alpha_{23} \\ \alpha_{24} \end{bmatrix}' \begin{bmatrix} \text{distinction} \\ \text{uniqueness} \\ \text{social reward} \\ \text{prestige} \end{bmatrix} + \delta_2 W z_2^* + \eta_2 \quad \cdots\cdots \text{Eq. (4.1.2)}$$

$$\Gamma = \begin{bmatrix} 1 & l_{\Gamma 1}' \\ l_{\Gamma 1} & 1 \end{bmatrix} \quad \cdots\cdots \text{Eq. (4.1.3)}$$

where α, δ and l_Γ are the parameters need to be estimated. Note that z^* only act as a vehicle that transfers the corresponding specifications into the measurement equation model as shown in Chapter 3. Therefore, z^* are not parameters that need to be estimated. Moreover, the matrix Γ is defined as a correlation matrix in order to ensure the positive definite of the whole covariance matrix.

4.1.1.1 Social Interaction

Individuals are affected by others who have similar aspect with him or her. The higher the similarity is the more likely to have interaction either through word-of-mouth or observed learning. Social connections are usually modeled through the observed behavior with a sensitivity to social interactions (Tucker, 2008) or the geographic and/or demographic proximities of individuals (Bell & Song, 2007; Nam, Manchanda, & Chintagunta, 2010; Yang & Allenby, 2003).

In the empirical study, basic socio demographic indicators and product purchase behavior related value and life-style indicators are used to generate the weight matrix which examines the extent of influence of other individuals. The basic socio demographic indicators are age, gender, income level, and education level. All socio demographic indicators except gender are coded in ordinal categories: 6 categories within age, 10 categories within income level, and 5 categories within education level. There are totally 10 indicators in the purchase behavior related value and life-style indicators with 5 Likert scale each.

To construct the weight matrix, first record the similarity level between all individual pairs by the number of identical indicators. For example, individual i is a 32-year-old female with 3 million won monthly income and has a bachelor's degree. individual j is a 35-year-old male with 3.5 million won monthly income and has an undergraduate degree. Then individual i and j have 2 social demographic indicators in the same categories (age and income). Assume that individual i and j have 3 purchase behavior related value and life-style indicators in the same categories. Then the similarity between individual i and j is 5. Record all individuals' similarities with individual i and normalized the similarities with the sum equal to 1. Then the weight matrix W, a $(Q \times Q)$ matrix, is the row normalized weight matrix. Element of W, w_{ij} is the element in the i th row and j th column and indicates the weight of individual j on i. Note that $w_{ii} = 0 \ \forall i$.

4.1.2 Endogenous Outcomes

Telecommunication generation selection is directly correlated with mobile phone usage behaviors. Therefore, telecommunication generation choice not just effected by attributes of each technology generation but also effected by other correlated choices. Monthly telecommunication expenditure can largely represent individual's mobile phone usage. Moreover, the frequency of mobile phone usage directly correlated with telecommunication expenditure. The frequency of SNS usage some determines the exposure level of information about 5G service.

In the empirical study, telecommunication generation is assumed to be correlated with monthly telecommunication expenditure and frequency of SNS usage. The monthly telecommunication expenditure is correlated with frequency of mobile phone usage.

$$\breve{y}_{1,\text{epdt.}} = \left(\begin{bmatrix} \breve{\gamma}_{\breve{y}1,1} \\ \breve{\gamma}_{\breve{y}1,2} \\ \breve{\gamma}_{\breve{y}1,3} \end{bmatrix}' + \left(\begin{bmatrix} 0 & 0 \\ \breve{d}_{\breve{y}1,11} & \breve{d}_{\breve{y}1,12} \\ \breve{d}_{\breve{y}1,21} & \breve{d}_{\breve{y}1,22} \end{bmatrix} \begin{bmatrix} z_1^* \\ z_2^* \end{bmatrix} \right)' \right) \quad \cdots\cdots\cdots\cdots \text{Eq. (4.1.4)}$$

$$\times \begin{bmatrix} 1 \\ \text{income} \\ pho.freq. \end{bmatrix} + \breve{\varepsilon}_1$$

$$\tilde{y}_{1,\text{pho.freq.}} = \left(\begin{bmatrix} \tilde{\gamma}_{\tilde{y}2,1} \\ \tilde{\gamma}_{\tilde{y}2,2} \end{bmatrix}' + \left(\begin{bmatrix} 0 & 0 \\ \tilde{d}_{\tilde{y}1,11} & \tilde{d}_{\tilde{y}1,11} \end{bmatrix} \begin{bmatrix} z_1^* \\ z_2^* \end{bmatrix} \right)' \right) \quad \cdots\cdots\cdots\cdots \text{Eq. (4.1.5)}$$

$$\times \begin{bmatrix} 1 \\ \text{age} \end{bmatrix} + \tilde{\varepsilon}_1$$

$$\tilde{y}_{2,\text{SNS.freq.}} = \left(\begin{bmatrix} \tilde{\gamma}_{\tilde{y}2,1} \\ \tilde{\gamma}_{\tilde{y}2,2} \\ \tilde{\gamma}_{\tilde{y}2,3} \end{bmatrix}' + \left(\begin{bmatrix} 0 & 0 \\ \tilde{d}_{\tilde{y}2,11} & \tilde{d}_{\tilde{y}2,12} \\ \tilde{d}_{\tilde{y}2,21} & \tilde{d}_{\tilde{y}2,22} \end{bmatrix} \begin{bmatrix} z_1^* \\ z_2^* \end{bmatrix} \right)' \right) \quad \cdots\cdots\cdots\cdots \text{Eq. (4.1.6)}$$

$$\times \begin{bmatrix} 1 \\ \text{gender} \\ \text{age} \end{bmatrix} + \tilde{\varepsilon}_2$$

$$
\begin{bmatrix} U_{1,3G} \\ U_{1,4G} \\ U_{1,5G} \end{bmatrix} = \left(\begin{bmatrix} \widehat{\gamma}_{U1,1} \\ \widehat{\gamma}_{U1,2} \\ \widehat{\gamma}_{U1,3} \\ \widehat{\gamma}_{U1,4} \\ \widehat{\gamma}_{U1,5} \\ \widehat{\gamma}_{U1,6} \\ \widehat{\gamma}_{U1,7} \end{bmatrix}' + \left(\begin{bmatrix} \widehat{d}_{U1,11} & \widehat{d}_{U1,12} \\ \widehat{d}_{U1,21} & \widehat{d}_{U1,22} \\ \widehat{d}_{U1,31} & \widehat{d}_{U1,32} \\ 0 & 0 \\ \widehat{d}_{U1,41} & \widehat{d}_{U1,42} \\ 0 & 0 \\ \widehat{d}_{U1,51} & \widehat{d}_{U1,52} \end{bmatrix} \begin{bmatrix} z_1^* \\ z_2^* \end{bmatrix} \right)' \right)
$$

$$
\times \begin{bmatrix} \mathrm{cvr}_{3G} & \mathrm{spd}_{3G} & \mathrm{rto}_{3G} & 0 & 0 & 0 & 0 \\ \mathrm{cvr}_{4G} & \mathrm{spd}_{4G} & \mathrm{rto}_{4G} & 1 & epdt & 0 & 0 \\ \mathrm{cvr}_{5G} & \mathrm{spd}_{5G} & \mathrm{rto}_{5G} & 0 & 0 & 1 & SNS.freq. \end{bmatrix} \cdots\cdots\cdots \text{Eq. (4.1.7)}
$$

$$
+ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & l_{\Lambda 1}' \\ 0 & l_{\Lambda 1} & l_{\Lambda 2} \end{bmatrix}
$$

where γ, d, ε and l_{Λ} are parameters need to be estimated. Note that the error terms of ordinal outcome are normalized to 1 and only upper thresholds ψ are estimable. Moreover, since only the difference of utility matters there are only two parameters need to be estimation in the error term of the nominal outcome.

4.2 Data Description

The data used in empirical study is from Korean media panel survey conducted by Korea Information Society Development Institute (KISDI). In order to analyze the media usage patterns and changes of individuals and households, Korean media panel survey follows approximately 10,000 individuals annually since 2010. The media panel data provides detailed information about social demographics and is useful for gain knowledge about the media ecosystem and changes in individual and family media related usage behaviors and useful both for cross-sectional and time-series studies (Lee, Wong, Oh, & Chang, 2019). The individual survey consists of two parts. The first part is about media ownership, services subscription, expenditure, and usage behaviors. The second part is called media diary which consists the individuals' behaviors associated with all kinds of media in 15 minutes unit for 3 days.

The empirical study uses the penal data from 2015 to 2019 with individuals who answered all years and who owned at least one mobile phone. There are totally 1637 individuals answered from 2010 to 2019. After excluding cases that does not consist a telecommunication generation choice 5316 observations in the final dataset. The survey questionnaire used in the empirical study is shown in Appendix 7.

The functional innovativeness and social innovativeness are surveyed in 2019. Consider that the internal characteristic of individual does not change in a relatively short time it is okay to assume the latent psychological indicators does not change across the time under consideration. The statistic of functional innovativeness and social innovativeness are shown in Table 3 and Table 4 respectively.

Table 3. Statistics of Functional Innovativeness

	Q1-1		Q1-2		Q1-3		Q1-4	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**
1	976	14.94	1082	16.56	1053	16.14	1133	17.34
2	2440	37.34	2218	33.95	2073	31.77	2103	32.19
3	2387	36.53	2083	31.88	2056	31.51	2388	36.55
4	691	10.58	1095	16.76	1263	19.36	826	12.64
5	40	0.61	56	0.86	80	1.23	84	1.29

Note: Cronbach Alpha Coefficient = 0.884

Table 4. Statistics of Social Innovativeness

	Q1-1		Q1-2		Q1-3		Q1-4	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**
1	815	12.47	1021	15.63	1081	16.54	1053	16.12
2	1934	29.60	2131	32.61	1811	27.72	2103	32.19
3	2389	36.56	2342	35.84	2472	37.83	2268	34.71
4	1244	19.04	921	14.10	1065	16.30	1044	15.98
5	152	2.33	119	1.82	105	1.61	66	1.01

Note: Cronbach Alpha Coefficient = 0.884

From above table we can notice that majority functional innovativeness and social innovativeness levels are concentrated in the middle. Averagely 16.25% and 15.44% of individuals show very low level of functional and social innovativeness respectively. Averagely 1% and 1.68% of individuals show very high level of functional and social innovativeness respectively. These statistics are coincidence with distribution of innovativeness of population. Moreover, the Cronbach alpha coefficient of both latent psychological variables are 0.884, which indicate high reliability of utilize the indicators as the latent psychological variables.

The phone ownership ratio is 96.5% in 2019 and 91.7% ownership is a smart phone. Compare to 89.2% in 2011, the penetration of mobile phone increases steadily. Table 5 and Figure 11 shows the change of the ratio of each telecommunication generation. The 2G is already in a downward situation in 2010 and the number of 3G normal phone users has reached the peak at the same time. The development of 3G smart-phone is somehow suppressed by 4G smart-phones. The launch of LTE-A in 2014 has made a huge development and has not reach the peak yet. The number of PDA users are very limited. Therefore, it is reasonable to exclude the PDA phone as an alternative in telecommunication generation choice. Similar with the 2G phone users. Hence the only considerable alternatives in telecommunication generation choice are restricted to 3G (3G normal phone and 3G smart-phone), 4G (LTE smart phone and LTE-A smart phone) and 5G (5G smart-phone).

Table 5. Telecommunication Generation Ratio of Major Using Phone

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
2G N.Pho	32.38	20.45	12.35	8.63	5.89	4.54	3.33	2.67	1.95	1.76
3G N.Pho	56.81	55.12	33.46	19.38	14.52	3.18	2.43	1.95	1.49	1.35
PDA	0.73	0.38	0.27	0.09	0.13	0.47	0.08	-	-	-
3G s.pho	10.03	24.03	42.49	34.68	23.56	13.64	6.59	4.43	3.02	2.28
LTE S.pho	-	-	11.43	37.20	44.28	57.58	45.40	41.33	35.68	27.64
LTE-A S.Pho	-	-	-	-	11.63	28.78	48.00	54.24	61.30	67.01
5G S.Pho	-	-	-	-	-	-	-	-	-	3.07

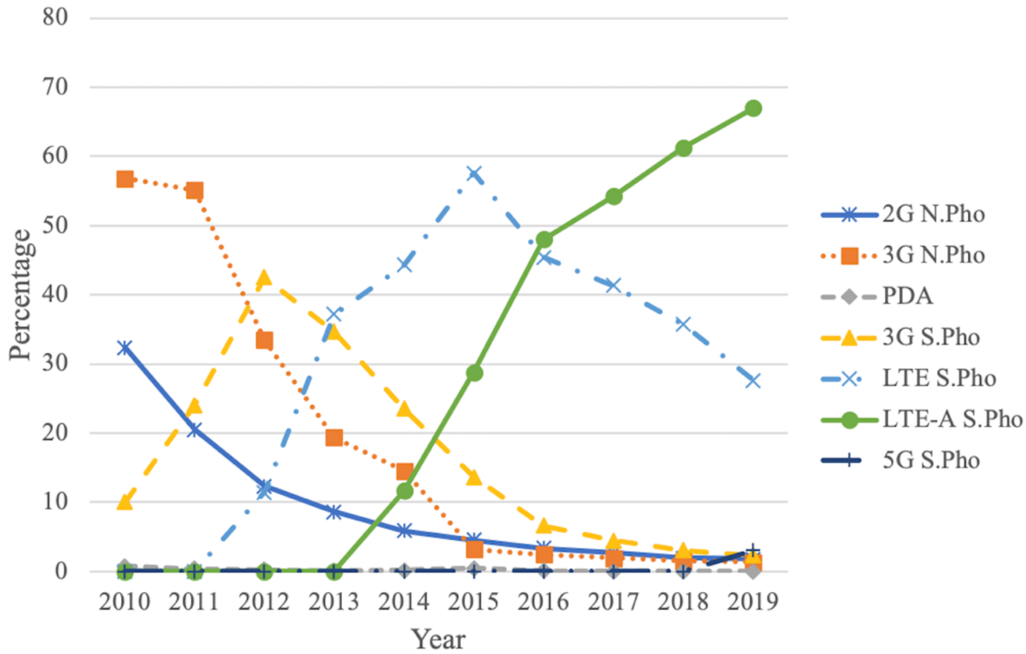

Figure 11. Diffusion of Telecommunication Generations

As shown in Figure 12, the usage of SNS surged form 16.8% in 2011 to 47.7% in 2019, which indicates the importance of social network in individual's daily life.

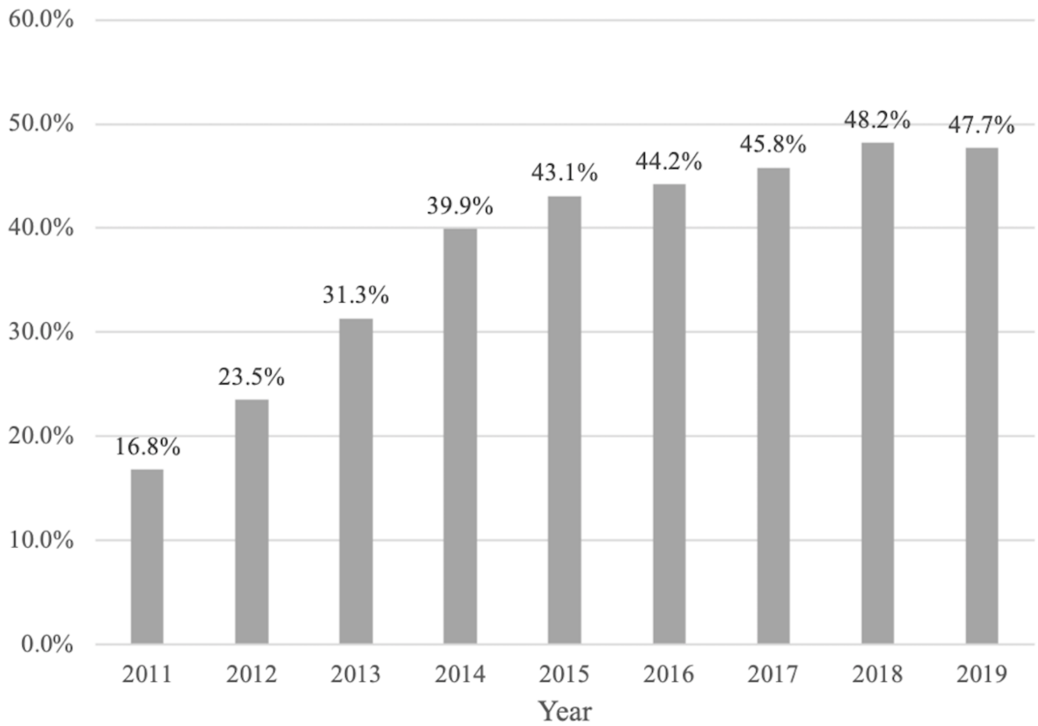

Figure 12. Change of SNS Usage

We consider 3 alternative specific variables which are coverage, speed and ratio. The National Information Society Agency (NIA) evaluate the quality of telecommunication service annually since 2012. The coverage is measured by the proportion of succeed send of certain information which is shown in Table 6. The speed is measured by the download speed as presented in Table 7. The ratio is the number of adopted individuals which is reported by the Ministry of Science and ICT monthly as shown in Table 8.

Table 6. Coverage of Telecommunication Generation (%)

	3G	4G	5G
2015	97.74%	99.72%	0
2016	97.15%	99.62%	0
2017	99.99%	99.21%	0
2018	98.96%	99.43%	0
2019	98.43%	99.18%	30.00%

Table 7. Speed of Telecommunication Generation (MB/bps)

	3G	4G	5G
2015	4.75	117.51	0
2016	5.59	120.09	0
2017	5.24	133.43	0
2018	6.08	150.68	0
2019	5.5	158.53	243.8

Table 8. Number of Users of Telecommunication Generation (100,000)

	3G	4G	5G
2015	12.54	41.69	0.00
2016	11.44	46.31	0.00
2017	10.66	50.44	0.00
2018	9.55	55.13	0.00
2019	7.52	55.69	4.67

The coverage of 3G improves slightly from 2015 to 2019 and the coverage of 4G remains at level of 99%. There is no official coverage statistics of 5G. 30% coverage level is speculated by the open coverage map of each telecommunication service providers. The download speed of 3G and 4G have huge differences and so is 5G.

However, consider the theoretical download speed of 5G is 20 times faster than 4G, the current download speed of 5G cannot meet the expectations. The number of 3G users decrease continuously and the number of 4G users increase gradually. The 5G service is launched in 2019. Therefore, the number of 5G users are very limited.

4.3 Estimation Results

In In order to compare the proposed model with the typical choice models, the author conduct as ordinal least square for continuous outcome, two ordered logistic regressions for ordinal outcomes, and a conditional multinomial logit model for nominal outcome. Moreover, normal GHDM without latent variables and unsocial GHDM model is also perform in purpose of comparison. The unsocial GHDM model treat the latent psychological variables as $z_{ql}^{*} = \alpha_l' s_q + \eta_{ql}$ which do not consist the term of social interaction. The social GHDM model is the proposed model of this study.

The following contents are constructed with three parts. The first part discusses the estimation results of the structural equation model for the latent psychological variables. The second part discusses the estimation results of the measurement equation models for the endogenous outcomes. The third part conducts a comparison between proposed model and other models.

4.3.1 Structural Equation Model for Latent Psychological Variables

The concepts of functional innovativeness and social innovativeness are well introduced in Vandecasteele and Geuens (2010). The estimation results which is shown in Table 9 indicates that the psychological indicator variables are suitable as the proxies of the latent psychological variables.

Table 9. Estimation Result of Latent Psychological Variables

	Unsocial GHDM		Social GHDM	
	Coeff.	T. Stat.	Coeff.	T.Stat
Functional Innovativeness				
Novel	0.343	19.174	0.304	10.134
Efficiency	0.062	3.891	0.102	2.941
Compatibility	0.082	5.374	0.056	2.118
Handiness	0.082	6.022	0.034	3.094
Social Innovativeness				
Distinction	0.202	13.046	0.131	6.094
Uniqueness	0.044	2.623	0.032	2.624
Social Reward	0.224	11.844	0.153	6.915
Prestige	0.131	7.453	0.057	1.974
Correlation	0.262	5.881	0.242	5.314
Social Dependence of Funt. Inno.			0.308	2.469
Social Dependence of Soci. Inno.			0.589	2.891

Seek for novel is the most powerful instinct for the functional innovativeness of individuals. Novelty usually companions with uncertainty. Innovative individuals are good at dealing with uncertainties than others (Rogers, 2010). Therefore, seek for novel is an important characteristic for individuals with high function innovativeness. On the other hand, the needs for distinction and social reward are the important motivations of adopting the innovations. The correlation between functional innovativeness and social innovativeness are smaller than 0.3 and statistically significant. This indicates the necessary of distinction of the innovativeness as Vandecasteele and Geuens (2010) pointed that the same observed behavior may due to different motivations and it is important to capture the differences in the behavioral motivations to understand individual behaviors.

The social interaction parameter estimates for functional innovativeness and social innovativeness are 0.308 and 0.589, respectively, and statistically significant. This indicates that both latent psychological variables are social dependent. The formation of individual's functional and social innovativeness levels depend on other individuals who are close to him/her in perspective of social-demographic and product purchase related value and life-styles. Compare to the functional innovativeness, the social innovativeness is more likely effected by peers. This is intuitive because social innovativeness has strong motivation associated with social behaviors, such as showing uniqueness and getting social rewards. On the other hand, functional innovativeness is motivated by the functional performance of the innovation. The social interactions enhance the need for distinction and social rewards. Therefore, with peers who have relatively high level of social innovativeness, the individual would also like to improve his/her social innovativeness level in order to maintain the uniqueness and so on. In order to improve a group of people's social innovativeness level it would be a good idea to focus on a small group of individuals who consist traits of opinion leader and more sensitive to information from external communication channel. On the other hand, to improve the functional innovativeness level of a certain group the same strategy may not as effective as it has been for the social innovativeness improvement.

4.3.2 Effect of Latent Psychological Variables on Endogenous Outcomes

In this empirical study, there are totally four of three types of outcomes are considered and are endogenously correlated. There is one continuous outcome: the monthly expenditure of telecommunication service, two ordinal outcomes: the frequency of daily smart phone usage and the frequency of SNS usage, one nominal outcome: the choice of telecommunication generation. The frequency of daily smart phone usage has a direct effect on the monthly expenditure of telecommunication service. The exogeneous factor: the age of individuals both affect on frequency of daily smart phone usage and frequency of SNS usage. The income level has and direct effect on monthly telecommunication expenditure. The choice of telecommunication generation are directly affected by monthly telecommunication service expenditure and frequency of SNS usage. The latent psychological variables simultaneously effect on all outcomes. Since the major concern of this empirical study is the individuals' choice of telecommunication generation service, the estimation results are divided into two parts. The first part discusses about the estimation results of the continuous outcome and the ordinal outcomes. The second part covers the estimation results of the nominal outcome.

Table 10. Estimation Results of Continuous Outcome and Ordinal Outcomes

	Traditional Regressions		GHDM without Latent Variables		Unsocial GHDM						Social GHDM					
					Pure Coeff.		Funt. Inno.		Soci. Inno.		Pure Coeff.		Funt. Inno.		Soci. Inno.	
	Coeff.	T. Stat.	Coeff.	T. Stat.	Coeff.	T. Stat.	Coeff.	T. Stat.	Coeff.	T. Stat.	Coeff.	T. Stat.	Coeff.	T.Stat	Coeff.	T. Stat.
Monthly Telecommunication Expenditure.																
Income	2.12	21.38	1.654	3.951	0.936	3.943					0.932	1.994				
							1.140	8.297	1.633	20.941			1.144	5.948	1.628	1.697
Pho. Freq.	3.31	12.25	2.617	5.543	2.408	4.741					2.411	2.973				
							5.058	16.398	5.654	24.237			5.078	15.943	5.67	1.992
Phone Usage Frequency																
Age	-0.34	-23.9	-1.596	-7.693	-0.170	-9.534					-0.058	-4.844				
							-0.268	-30.65	-0.059	-5.892			-0.141	-5.648	-0.013	-2.103
SNS Usage Frequency																
Gender	0.78	6.95	-0.694	-5.614	-0.499	-2.307					-0.496	-1.894				
							-0.055	-0.481	0.453	4.953			-0.023	-0.973	0.48	0.914
Age	-0.17	-5.73	-0.339	-7.481	-1.020	-9.983					-1.008	-9.736				
							-1.153	-16.97	0.232	4.083			-1.095	-1.321	0.286	4.004

Table 11. Estimation Result of Nominal Outcomes (Telecommunication Generation Choice)

	Conditional Multinomial Logit Model		GHDM without Latent Variables		Unsocial GHDM						Social GHDM					
					Pure Coeff.		Funt. Inno.		Soci. Inno.		Pure Coeff.		Funt. Inno.		Soci. Inno.	
	Coeff.	T. Stat.	Coeff.	T. Stat.	Coeff.	T. Stat.	Coeff.	T. Stat.	Coeff.	T. Stat.	Coeff.	T. Stat.	Coeff.	T.Stat	Coeff.	T. Stat.
Coverage	-0.179	-8.136	-0.259	-10.85	-0.169	-7.673					-0.169	-2.184				
							0.168	7.623	0.183	7.054			0.191	4.849	0.206	1.982
Speed	0.032	1.231	0.194	4.841	0.115	4.435					0.113	2.194				
							0.709	7.961	-0.289	-3.799			0.725	3.947	-0.272	-1.947
Ratio	0.51	23.182	1.009	59.641	1.395	63.391					1.39	10.954				
							0.106	2.155	0.267	2.379			0.125	1.984	0.29	1.893
Expd.			0.477	5.915	0.219	4.459					0.218	1.294				
(3G)	-0.029	-14.50														
(5G)	0.014	1.273														
							0.658	6.645	0.580	11.845			0.68	1.774	0.603	8.648
SNS.Freq.			0.193	1.907	0.168	1.716					0.17	0.694				
(3G)	0.021	0.188														
(5G)	0.053	0.115														
							0.233	2.082	0.216	2.199			0.255	1.367	0.679	3.471

4.3.2.1 Estimation Results of Continuous Outcome and Ordinal Outcomes

The parameter estimates of the loadings of the latent psychological variables on the monthly telecommunication expenditure and frequency of SNS usage and daily smart phone usage are shown in the Table 10. The first column indicates the results of OLS regression with the dependent variable is monthly telecommunication expenditure and the ordered logistic regression with the dependent variables are frequency of daily smart phone usage and SNS usage. The second column represents the GHDM model without consideration of latent psychological variables. The third column shows the result of GHDM model with latent psychological variables (no social interaction within the latent psychological variables). The last column provides the result of the proposed model, the GHDM model with social interaction affected latent psychological variables considered simultaneously. “Pure coefficient” indicates the pure effect of the explanatory variable on the dependent variable.

Generally, the pure coefficients of explanatory variables decrease with consideration of latent psychological variables. For example, the coefficient of age on the frequency of daily smart phone usage is -0.34 in the ordered logistic regression. The coefficient is -1.596 in no latent GHDM model. The pure effect of age on the frequency of daily smart phone usage decreased to 0.17 in unsocial GHDM and 0.058 in social GHDM. Though the results are still statistically significant, the pure effect of age has been explained by other factors.

Mobile phone usage frequency has more effect on monthly telecommunication expenditure than income, which is identical across all models. The smart phone usage frequency directly affects on the amount of data usage which is related with usage of rate system in the most cases. Individuals with high level of social innovativeness are more sensitive to the frequency of daily smart phone usage. High social innovativeness individuals may be more enthusiastic to up to date to surrounding environments. This propensity may drive the more frequently use of smart phone. Compare to high social innovativeness individuals, high functional innovativeness individuals are less sensitive to income and daily smart phone usage frequency. May be because high functional innovativeness individuals more rely on other communication channel such as work place network or TV based mass media.

Younger individuals are more likely to have higher mobile phone usage frequency and higher SNS usage frequency, which is indicated from the ordered logistic regression. Individuals with higher functional innovativeness are less sensitive to age. Clearly, individual's functional innovativeness level are not related with age. Sometime, age have an negative effect on individual's functional innovativeness. Because individuals with relatively high level of functional innovativeness are relatively face less budget constraints while facing an adoption decision. And income level is somehow related with age. Moreover, individuals with higher functional innovativeness are less sensitive to gender which is intuitive. However, individuals with higher social innovativeness are less sensitive to gender.

4.3.2.2 Estimation Result of Nominal Outcomes

The parameter estimates of the loadings of the latent psychological variables on the choice of telecommunication generation are shown in the Table 11. In telecommunication generation choice the number of adopted individuals largely effect on individuals' choice and increase of monthly telecommunication expenditure leads to lower probability of selecting 3G compare to 4G and increase of phone usage frequency leads to higher probability of selecting 5G compare to 4G. Note that in the GHDM models the base category is coded as 3G. In the conditional multinomial logit model, speed is not statistically significant. The functional innovativeness has strong positive effect on preference of speed, but the social innovativeness somehow mutualized the effect. Besides, the social innovativeness has stronger effect on preference of adopted number of individuals than functional innovativeness. Latent psychological variables enhanced the effect of telecommunication expenditure and the effect of SNS usage frequency on telecommunication generation choice. From the conditional multinomial logit model, we can only confirm the effect of telecommunication expenditure and the effect of SNS usage frequency. From the GHDM with social interaction we know that part of the effect of the explanatory variables come from social interaction.

4.3.3 Comparison of the GHDM models

The adjusted composite likelihood ratio test (ADCLRT) statistic can be used to confirm the statistical fit of the model with maximum approximation composite marginal likelihood (MACML) approach (Bhat, 2011). The log-CML value is -7318.13 for the unsocial GHDM and -6085.54 for the social GHDM. The ADCLRT statistic has a chi-square asymptotic distribution. The calculated ADCLRT statistic is 115.94, which is higher than the critical chi-square value with two degrees of freedom. It is the evidence of the importance of considering the social interaction. The composite likelihood information criterion (CLIC) (Varin & Vidoni, 2005) can be used to compare the no latent GHDM model and unsocial GHDM model. Model with higher CLIC value is preferred. The CLIC for the no latent GHDM model is -5967.18 and -3467.53 for the unsocial GHDM model. Therefore, unsocial GHDM is preferred over no latent GHDM model.

Bhat et al. (2016) suggest that average treatment effects (ATEs) can be computed to compare the different models. To further compare the no latent GHDM model, unsocial GHDM model, and the proposed social GHDM model, compute the ATEs in a simulation process to obtain the expected difference in the alternatives for an individual change his/her daily smart phone usage frequency.

For each model, synthesize an individual with random draw of yU with parameter estimates of the corresponding model. Calculate the value of dependent variables. After that set the value of daily smart phone usage frequency as 1. Calculate the synthetic individual's utility for each alternative (3G, 4G, and 5G) and record the expected share. Then set the value of daily smart phone usage frequency as 3 and record the expected share same as the previous step. Then, compute the difference in expected share obtained from above two steps. Repeat this procedure for 500 times to generate 500 synthetic individuals. Then compute the mean and standard error of the 500 draws. The results of ATEs for choice of telecommunication generation for the no latent GHDM, unsocial GHDM, and social GHDM are shown in Table 12. The t-test significant level of ATE of unsocial GHDM is according to GHDM without latent variables. The significant level of ATE of social GHDM is according to unsocial GHDM.

Table 12. Average Treatment Effects (ATEs)

	GHDM without Latent Variables		Unsocial GHDM		Social GHDM	
	ATE	Std.Err	ATE	Std.Err	ATE	Std.Err
3G	-0.067	0.034	-0.039**	0.056	-0.023	0.041
4G	0.043	0.021	0.016	0.011	0.040	0.034
5G	0.164	0.074	0.057**	0.061	0.034*	0.081

Note: * indicates 90% significant level; ** indicates 95% significant level

Assume that there are 100 random chosen individuals change their daily smartphone usage from low level to high level, the 3G share would decrease 6.7% averagely in no latent GHDM model. However, in the unsocial GHDM model the decrease is only 3.9% and 2.3% in social GHDM model. Without consideration of latent psychological variables, the effect of daily smartphone usage frequency is exaggerated from 3.9% to 6.7%. Without consideration of social interaction within latent psychological variables, the effect of daily smartphone usage frequency is exaggerated from 2.3 % to 3.9%. Similar results in 5G adoption situation. Improve in daily smartphone usage frequency has only 3.4% increase in share in social GHDM model. However, the share is estimated to16.4% in no latent GHDM model.

Studies of forecasting 5G adoption without consider the individual behavioral characteristics (such as Lim and Kim (2017) and Jahng and Park (2020)) tend to over-estimate the number of potential adopters and the diffusion speed. Ignore the effect of social environment on individual's behaviors may leads to misunderstanding. Smart phone adoption behavior is highly effected by individual's belief related factors (Kim & Kim, 2011). Jeong et al. (2020) prove that individual 5G adoption intention and continuous usage intention is positively correlated with individual innovativeness. Moreover, smartphone usage behavior is highly correlated with individual innovativeness (Al-Obthani & Ameen, 2019; Dayour, Park, & Kimbu, 2019; Tussyadiah, 2016). Therefore, it is necessary to consider individual innovativeness while studying on individual telecommunication generation choice.

As mentioned in section 2.1, the change of individual's characteristics such as attitudes and opinions highly depend on interactions with other individuals. Hence, from the proposed social GHDM model we can not only get better understand of individuals' choice behavior and the endogenously correlated relationships but also disentangle the variable effect by latent psychological variables which are social dependent.

Chapter 5. Conclusion

5.1 Concluding Remarks and Contribution

This study proposed a new Generalized Heterogenous Data Model (GHDM) with social interaction. The abstract concept of the proposed model is shown in Figure 13. The proposed model embeds social interaction in the latent psychological variables which generally indicate an individual's attitude and propensity. Subsequently, the latent psychological variable has an effect on the explanatory variables, and is not considered as an additive term. The proposed model has mainly two contributions to choice modeling.

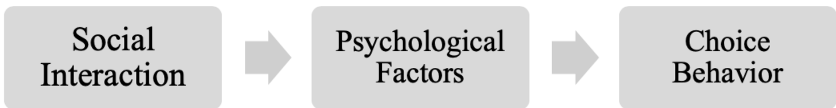

Figure 13. Abstract Concept of the Propose Model

First, the proposed model can incorporate word-of-mouth, which is one type of social interaction, into the choice model. Social interactions are considered in the choice of adoption of innovations with strong network or spill-over effects. Studies on social interaction in economics have focused on observed learning, which is one type of social interaction. However, word-of-mouth is hardly ever incorporated into choice models. The major reason is that during a word-of-mouth interaction, the individuals acquire information about the innovation and adjust their attitude toward the innovation. Hence, the word-of-mouth interaction deals with an individual's level of information, attitude, value, and propensity, and does not have a direct effect on an individual's behavior. An individual's choices are affected by both characteristics of the alternatives and the characteristics of the individual. Some characteristics of the individual, especially psychological characteristics, cannot be observed by the researchers, such as an individual's attitude and opinions. There are many studies that make an effort to include such unobserved individual characteristics into choice models, including the ICLV model. However, only a few studies have considered the effect of social interaction on unobserved individual characteristics. Ignoring the effect of social interaction and the unobserved individual characteristics may lead to inconsistent estimation and biased estimates. The proposed model incorporates word-of-mouth social interaction into the latent psychological variables which generally reflect individuals' attitudes and opinions, and further incorporates the latent psychological variables into the choice model. By doing so, the proposed model is able to capture the effect of word-of-mouth social interaction on individual choice behaviors.

Second, the proposed model can better reflect heterogeneity across individuals. Previous studies have incorporated social interaction as an additive term in the utility function, which indicates that the adoption behavior of other individuals itself increases the expected utility. In case the innovation has strong network externality or spill-over effects, the additive term of social interaction is reasonable. However, in case the innovation has limited network externality or spill-over effects, the information on the adoption behavior of others intuitively influences an individual's attitude and expectations, not the utility. Of course, there is persuasion theory and the pure-exposure theory, which argue that the simple repeated exposure of an object can increase a positive attitude toward the object, which cannot explain most choice behaviors. When an individual is informed about an innovation (either through external or internal communication channels), he or she forms an attitude toward the innovation according to the information he or she has. Only when the attitude toward the innovation exceeds a certain threshold and other constraints are met that the individual would decide to adopt the innovation. A change in attitude requires information which largely depends on social interaction. To the best of the author's knowledge, there is no economic choice model that has modeled social interaction in this way.

Furthermore, the proposed model also considers multiple endogenously correlated outcomes which are mixed types of dependent variables, including continuous, ordinal, and nominal outcomes. This model specification leads to evaluation of a high dimensional multivariate distribution function which may be literally impossible to estimate using full likelihood function estimation approaches. We applied a composite marginal likelihood function with the maximum approximation approach which is known as the MACML approach. This approach decomposes the likelihood function into a surrogated marginal likelihood function, and only univariate and bivariate cumulative normal distributions need to be evaluated.

We first conducted a simulation study to confirm whether the model can recover true values from the initial values and we found that the model performs well even with a limited sample size. Subsequently, we conducted an empirical study with the proposed model and established that by using a GHDM with social interaction, we can better understand an individual's behavior. The estimation result shows that ignoring social interaction and unobserved individual characteristics leads to biased estimates. The proposed model not only provides more information on an individual's decision process but also provides more precise estimates of the parameters. Utilizing the additional information can help policy makers create more efficient propaganda, and assist a marketing director in developing more effective marketing strategies.

5.2 Limitations and Future Studies

The empirical study conducted in this work is limited in reflecting the complexity of the telecommunication industry in Korea. The only exogeneous variables considered are gender, income, and age. Other factors such as subsidies and marketing activities are not considered in the empirical study due to the limitation of the data set.

This study incorporates individual psychological characteristics into a covariate of explanatory variables. This process somehow increases the non-linearity of the model. The number of parameters increases exponentially with an increase in the number of explanatory and latent psychological variables. Therefore, the proposed model requires an abundance of previous research to restrict the effect of latent psychological variables on explanatory variables. The interpretation of the absolute value of the estimated parameters, such as the social interaction parameter δ, is somehow restricted. Up to the current model, only the relative amount of the social interaction parameter is meaningful in interpretation.

In the proposed model, the weight matrix in the latent psychological variables is defined externally. Future studies need to focus on the construction and estimation of the weight matrix, simultaneously.

The effect of social interaction on a latent psychological variable is assumed to be identical across individuals in the current proposed framework. It is clear that individuals with different characteristics have different sensitivity toward social interaction. To better model the heterogeneity among individuals, one needs to assume an error term in the social interaction parameter.

Bibliography

Ajzen, I. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Akerlof, G. A., & Kranton, R. E. (2000). Economics and Identity. *The Quarterly Journal of Economics, 115*(3), 715-753. doi:10.1162/003355300554881

Al-Obthani, F., & Ameen, A. (2019). Influence of Overall Quality and Innovativeness on Actual Usage of Smart Government: An Empirical Study on the UAE Public Sector. *International Journal on Emerging Technologies, 10*(1), 141-146.

Ashok, K., Dillon, W. R., & Yuan, S. (2002). Extending Discrete Choice Models to Incorporate Attitudinal and Other Latent Variables. *Journal of Marketing Research, 39*(1), 31-46. doi:10.1509/jmkr.39.1.31.18937

Bandura, A. (1986). *Social foundations of thought and action: a social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Banerjee, A. V. (1992). A Simple Model of Herd Behavior. *The Quarterly Journal of Economics, 107*(3), 797-817. doi:10.2307/2118364

Bell, D. R., & Song, S. (2007). Neighborhood Effects and Trial on the Internet: Evidence from Online Grocery Retailing. *Quantitative Marketing and Economics, 5*(4), 361-400. doi:10.1007/s11129-007-9025-5

Ben-Akiva, M., Walker, J., Bernardino, A. T., Gopinath, D. A., Morikawa, T., &

Polydoropoulou, A. (2002). *Integration of choice and latent variable models*. Pergamon: Elsevier Science.

Bernheim, B. D. (1994). A Theory of Conformity. *Journal of Political Economy, 102*(5), 841-877. doi:10.1086/261957

Bhat, C. R. (1997). An Endogenous Segmentation Mode Choice Model with an Application to Intercity Travel. *Transportation Science, 31*(1), 34-48. doi:10.1287/trsc.31.1.34

Bhat, C. R. (1998). Accommodating Variations in Responsiveness to Level-of-Service Measures in Travel Mode Choice Modeling. *Transportation Research Part A: Policy and Practice, 32*(7), 495-507. doi:10.1016/S0965-8564(98)00011-1

Bhat, C. R. (2011). The Maximum Approximate Composite Marginal Likelihood (MACML) Estimation of Multinomial Probit-based Unordered Response Choice Models. *Transportation Research Part B: Methodological, 45*(7), 923-939. doi:10.1016/j.trb.2011.04.005

Bhat, C. R. (2014). The Composite Marginal Likelihood (CML) Inference Approach with Applications to Discrete and Mixed Dependent Variable Models. *Foundations and Trends in Econometrics, 7*(7), 1-117. doi:10.1561/0800000022

Bhat, C. R. (2015a). A New Generalized Heterogeneous Data Model (GHDM) to Jointly Model Mixed Types of Dependent Variables. *Transportation Research Part B: Methodological, 79*, 50-77. doi:10.1016/j.trb.2015.05.017

Bhat, C. R. (2015b). A New Spatial (Social) Interaction Discrete Choice Model

Accommodating for Unobserved Effects Due to Endogenous Network Formation. *Transportation, 42*(5), 879-914. doi:10.1007/s11116-015-9651-9

Bhat, C. R., & Dubey, S. K. (2014). A New Estimation Approach to Integrate Latent Psychological Constructs in Choice Modeling. *Transportation Research Part B: Methodological, 67*, 68-85. doi:10.1016/j.trb.2014.04.011

Bhat, C. R., Pinjari, A. R., Dubey, S. K., & Hamdi, A. S. (2016). On Accommodating Spatial Interactions in a Generalized Heterogeneous Data Model (GHDM) of Mixed Types of Dependent Variables. *Transportation Research Part B: Methodological, 94*, 240-263. doi:10.1016/j.trb.2016.09.002

Bhat, C. R., Schofer, J. L., Koppelman, F. S., & Bautch, R. C. (1993). Driver Recruitability for Advanced Traveler Information System Experiments. *Transportation Research Part C: Emerging Technologies, 1*(4), 265-274. doi:10.1016/0968-090X(93)90001-V

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy, 100*(5), 992-1026. doi:10.1086/261849

Bisin, A., & Verdier, T. (2001). The Economics of Cultural Transmission and the Dynamics of Preferences. *Journal of Economic Theory, 97*(2), 298-319. doi:10.1006/ /jeth.2000.2678

Bolduc, D., & Alvarez-Daziano, R. (2010). On estimation of hybrid choice models. In S. Hess & A. Daly (Eds.), *Choice modelling: the state-of-the-art and the state-of-*

practice, Proceedings from the inaugural international choice modelling conference (pp. 259–287). England: Emerald.

Brock, W. A., & Durlauf, S. N. (2001). Discrete Choice with Social Interactions. *The Review of Economic Studies, 68*(2), 235-260. doi:10.1111/1467-937x.00168

Broekhuizen, T. L. J., Delre, S. A., & Torres, A. (2011). Simulating the Cinema Market: How Cross-Cultural Differences in Social Influence Explain Box Office Distributions. *Journal of Product Innovation Management, 28*(2), 204-217. doi:10.1111/j.1540-5885.2011.00792.x

Carletti, T., Fanelli, D., Grolli, S., & Guarino, A. (2006). How to Make an Efficient Propaganda. *Europhysics Letters, 74*(2), 222-228. doi:10.1209/epl/i2005-10536-9

Castellano, C., Fortunato, S., & Loreto, V. (2009). Statistical Physics of Social Dynamics. *Reviews of Modern Physics, 81*(2), 591-646. doi:10.1103/RevModPhys.81.591

Castro, J., Lu, J., Zhang, G., Dong, Y., & Martínez, L. (2018). Opinion Dynamics-Based Group Recommender Systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 48*(12), 2394-2406. doi:10.1109/TSMC.2017.2695158

Castro, M., Paleti, R., & Bhat, C. R. (2013). A Spatial Generalized Ordered Response Model to Examine Highway Crash Injury Severity. *Accident Analysis & Prevention, 52*, 188-203. doi:10.1016/j.aap.2012.12.009

Currarini, S., & Mengel, F. (2016). Identity, Homophily, and In-Group Bias. *European Economic Review, 90*, 40-55. doi:10.1016/j.euroecorev.2016.02.015

Dayour, F., Park, S., & Kimbu, A. N. (2019). Backpackers' Perceived Risks Towards

Smartphone Usage and Risk Reduction Strategies: A Mixed Methods Study. *Tourism Management, 72*, 52-68. doi:10.1016/j.tourman.2018.11.003

Deffuant, G., Amblard, F., Weisbuch, G., & Faure, T. (2002). How Can Extremism Prevail? A Study Based on the Relative Agreement Interaction Model. *Journal of Artificial Societies and Social Simulation, 5*(4).

Deffuant, G., Neau, D., Amblard, F., & Weisbuch, G. (2000). Mixing Beliefs Among Interacting Agents. *Advances in Complex Systems, 03*(01n04), 87-98. doi:10.1142/S0219525900000078

DeGroot, M. H. (1974). Reaching a Consensus. *Journal of the American Statistical Association, 69*(345), 118-121.

Delre, S. A., Jager, W., Bijmolt, T. H. A., & Janssen, M. A. (2010). Will It Spread or Not? The Effects of Social Influences and Network Topology on Innovation Diffusion. *Journal of Product Innovation Management, 27*(2), 267-282. doi:10.1111/j.1540-5885.2010.00714.x

Dittmer, J. C. (2001). Consensus Formation Under Bounded Confidence. *Nonlinear Analysis-Theory Methods and Applications, 47*(7), 4615-4622.

Duesenberry, J. S. (1949). *Income, saving and the theory of consumer behavior*. Cambridge: Harvard University Press.

Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations, 7*(2), 117-140.

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston: Peterson.

Flavell, J. H. (1963). *The developmental psychology of Jean Piaget*. New York: Reinhold.

Gilbert, N. (1997). A Simulation of the Structure of Academic Science. *Sociological Research Online, 2*(2), 91-105. doi:10.5153/sro.85

Godambe, V. P. (1960). An Optimum Property of Regular Maximum Likelihood Estimation. *Annals of Mathematical Statistics, 31*(4), 1208-1211. doi:10.1214/aoms/1177705693

Goldenberg, J., Libai, B., & Muller, E. (2001). Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth. *Marketing Letters, 12*(3), 211-223. doi:10.1023/A:1011122126881

Gramzow, R. H., Gaertner, L., & Sedikides, C. (2001). Memory for In-Group and Out-Group Information in a Minimal Group Context: The Self as an Informational Base. *Journal of Personality and Social Psychology, 80*(2), 188-205. doi:10.1037/0022-3514.80.2.188

Guhl, D., Baumgartner, B., Kneib, T., & Steiner, W. J. (2018). Estimating Time-Varying Parameters in Brand Choice Models: A Semiparametric Approach. *International Journal of Research in Marketing, 35*(3), 394-414. doi:10.1016/j.ijresmar.2018.03.003

Harris, K. M., & Keane, M. P. (1998). A Model of Health Plan Choice:: Inferring Preferences and Perceptions from a Combination of Revealed Preference and Attitudinal Data. *Journal of Econometrics, 89*(1), 131-157. doi:10.1016/S0304-4076(98)00058-X

Hartmann, W. R. (2010). Demand Estimation with Social Interactions and the Implications for Targeted Marketing. *Marketing Science, 29*(4), 585-601. doi:10.1287/mksc.1100.0559

Hartmann, W. R., Manchanda, P., Nair, H., Bothner, M., Dodds, P., Godes, D., . . . Tucker, C. (2008). Modeling Social Interactions: Identification, Empirical Methods and Policy Implications. *Marketing Letters, 19*(3), 287-304. doi:10.1007/s11002-008-9048-z

Heagerty, P. J., & Lumley, T. (2000). Window Subsampling of Estimating Functions with Application to Regression Models. *Journal of the American Statistical Association, 95*(449), 197-211. doi:10.1080/01621459.2000.10473914

Hegselmann, R., & Krause, U. (2002). Opinion Dynamics and Bounded Confidence Models, Analysis and Simulation. *Journal of Artificial Societies and Social Simulation, 5*(3), 1-2.

Hegselmann, R., & Krause, U. (2006). Truth and Cognitive Division of Labor: First Steps towards a Computer Aided Social Epistemology. *Journal of Artificial Societies and Social Simulation, 9*(3), 10.

Hogg, M. A. (2000). Subjective Uncertainty Reduction through Self-categorization: A Motivational Theory of Social Identity Processes. *European Review of Social Psychology, 11*(1), 223-255. doi:10.1080/14792772043000040

Innes, C. R. (2014). *Quantifying the effect of open-mindedness on opinion dynamics and advertising optimization.* (Master), Simon Fraser University, Jahng, J. H., & Park, S. K. (2020). Simulation-Based Prediction for 5G Mobile Adoption. *ICT Express, 6*(2), 109-112. doi:0.1016/j.icte.2019.10.002

Janssen, M. A., & Jager, W. (2001). Fashions, Habits and Changing Preferences: Simulation of Psychological Factors Affecting Market Dynamics. *Journal of Economic Psychology, 22*(6), 745-772. doi:10.1016/S0167-4870(01)00063-0

Jeong, M. S., Hong, D. S., & Ji, Y. G. (2020). A Study on the Factors Affecting the Usage Intentions of 5G Mobile Communication Service, *25*(1), 135-176. doi:10.7838/JSEBS.2020.25.1.135

Joe, H. (1995). Approximations to Multivariate Normal Rectangle Probabilities Based on Conditional Expectations. *Journal of the American Statistical Association, 90*(431), 957-964. doi:10.1080/01621459.1995.10476596

Kelman, H. C. (2017). Processes of opinion change. In *Attitude change* (pp. 205-233): Routledge.

Kim, D. W. (2017). Core Infrastructure in the Fourth Industrial Revolution, 5G. *KISDI Premium Report, 17*(6), 1-16.

Kim, H. J., & Kim, D. Y. (2011). Perceptions and Usages of Smartphone Users in the Different Phases of Adoption. *Korean Journal of Journalism & Communications Studies, 55*(4), 382-405.

Kim, J., & Hur, W. (2013). Diffusion of Competing Innovations in Influence Networks. *Journal of Economic Interaction and Coordination, 8*(1), 109-124. doi:10.1007/s11403-012-0106-5

Korea Information Society Development Insitutude (KISDI). *Media Panel Data (2015-2019)*. Retrieved from: https://www.kisdi.re.kr/kisdi/jsp/fp/kr/trend/KK_32710.jsp

Koppelman, F. S., & Hauser, J. R. (1978). Destination Choice Behavior for Non-Grocery-Shopping Trips. *Transportation Research Record,* 673, 157-165.

Krause, U. (2000). A Discrete Nonlinear and Non-autonomous Model of Consensus Formation. *Communications in Difference Equations, 2000*, 227-236. doi:10.1201/b16999-21

Kurz, S., & Rambau, J. (2011). On the Hegselmann-Krause Conjecture in Opinion Dynamics. *Journal of Difference Equations and Applications, 17*(6), 859-876. doi:10.1080/10236190903443129

Lee, H., Wong, S. F., Oh, J., & Chang, Y. (2019). Information Privacy Concerns and Demographic Characteristics: Data from a Korean Media Panel Survey. *Government Information Quarterly, 36*(2), 294-303. doi:10.1016/j.giq.2019.01.002

Li, Q., Braunstein, L. A., Havlin, S., & Stanley, H. E. (2011). Strategy of Competition Between Two Groups Based on an Inflexible Contrarian Opinion Model. *Physical Review E, 84*(6), 066101. doi:10.1103/PhysRevE.84.066101

Liang, H., Dong, Y., & Li, C. (2016). Dynamics of Uncertain Opinion Formation: An Agent-Based Simulation. *Journal of Artificial Societies and Social Simulation, 19*(4), 1. doi:10.18564/jasss.3111

Lim, D.-E., & Kim, T. (2017). An Application of a Multi-Generation Diffusion Model to Forecast 5g Mobile Telecommunication Service Subscribers in South Korea.

International Journal of Pure and Applied Mathematics, 116(23), 809-817.

Lovett, M. J., & Staelin, R. (2016). The Role of Paid, Earned, and Owned Media in Building Entertainment Brands: Reminding, Informing, and Enhancing Enjoyment. *Marketing Science, 35*(1), 142-157. doi:10.1287/mksc.2015.0961

Lu, T., & Tang, N. (2019). Social Interactions in Asset Allocation Decisions: Evidence from 401(k) Pension Plan Investors. *Journal of Economic Behavior & Organization, 159*, 1-14. doi:10.1016/j.jebo.2019.01.009

Ma, L., Krishnan, R., & Montgomery, A. L. (2015). Latent Homophily or Social Influence? An Empirical Analysis of Purchase Within a Social Network. *Management Science, 61*(2), 454-473. doi:10.1287/mnsc.2014.1928

Madanat, S., Yang, C. D., & Yen, Y. M. (1995). Analysis of Stated Route Diversion Intentions under Advanced Traveler Information Systems Using Latent Variable Modeling. *Transportation Research Record,* (1485), 10-17.

Manski, C. F. (2000). Economic Analysis of Social Interactions. *Journal of Economic Perspectives, 14*(3), 115-136. doi:10.1257/jep.14.3.115

Martins, A. C. R. (2008). Continuous Opinions and Discrete Actions in Opinion Dynamics Problems. *International Journal of Modern Physics C, 19*(4), 617-624. doi:10.1142/S0129183108012339

Martins, A. C. R. (2009). Bayesian Updating Rules in Continuous Opinion Dynamics Models. *Journal of Statistical Mechanics: Theory and Experiment, 2009*(2), P02017. doi:10.1088/1742-5468/2009/02/p02017

Martins, A. C. R. (2013). Trust in the CODA Model: Opinion Dynamics and the Reliability of Other Agents. *Physics Letters A, 377*(37), 2333-2339. doi:10.1016/j.physleta.2013.07.007

Martins, T. V., Pineda, M., & Toral, R. (2010). Mass Media and Repulsive Interactions in Continuous-Opinion Dynamics. *Europhysics Letters, 91*(4), 48003. doi:10.1209/0295-5075/91/48003

Maslow, A. (1943). A Theory of Human Motivation. *Psychological Review, 50*(4), 370-396.

McCoy, D., & Lyons, S. (2014). Consumer Preferences and the Influence of Networks in Electric Vehicle Diffusion: An Agent-Based Microsimulation in Ireland. *Energy Research & Social Science, 3*, 89-101. doi:10.1016/j.erss.2014.07.008

McCulloch, R. E., Polson, N. G., & Rossi, P. E. (2000). A Bayesian Analysis of the Multinomial Probit Model with Fully Identified Parameters. *Journal of Econometrics, 99*(1), 173-193. doi:10.1016/S0304-4076(00)00034-8

Müller, S., & Rau, H. A. (2019). Decisions Under Uncertainty in Social Contexts. *Games and Economic Behavior, 116*, 73-95. doi:10.1016/j.geb.2019.04.006

Nair, H. S., Manchanda, P., & Bhatia, T. (2010). Asymmetric Social Interactions in Physician Prescription Behavior: The Role of Opinion Leaders. *Journal of Marketing Research, 47*(5), 883-895. doi:10.1509/jmkr.47.5.883

Nam, S., Manchanda, P., & Chintagunta, P. K. (2010). The Effect of Signal Quality and Contiguous Word of Mouth on Customer Acquisition for a Video-on-Demand

Service. *Marketing Science, 29*(4), 690-700. doi:10.1287/mksc.1090.0550

Narayan, V., Rao, V. R., & Saunders, C. (2011). How Peer Influence Affects Attribute Preferences: A Bayesian Updating Mechanism. *Marketing Science, 30*(2), 368-384. doi:10.1287/mksc.1100.0618

Quattrociocchi, W., Caldarelli, G., & Scala, A. (2014). Opinion Dynamics on Interacting Networks: Media Competition and Social Influence. *Scientific Reports, 4*(1), 4938. doi:10.1038/srep04938

Quattrociocchi, W., Conte, R., & Lodi, E. (2011). Opinions Manipulation: Media, Power and Gossip. *Advances in Complex Systems, 14*(04), 567-586. doi:10.1142/S0219525911003165

Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. *The American Economic Review, 83*(5), 1281-1302.

Revelt, D., & Train, K. (1998). Incentives for Appliance Efficiency in a Competitive Energy Environment: Random Parameters Logit Models of Households' Choices. *Review of Economics and Statistics, 80*(4), 647-657.

Rogers, E. M. (2010). *Diffusion of innovations (4th ed.).* New York: The Free Press.

Roozmand, O., Ghasem-Aghaee, N., Hofstede, G. J., Nematbakhsh, M. A., Baraani, A., & Verwaart, T. (2011). Agent-Based Modeling of Consumer Decision Making Process Based on Power Distance and Personality. *Knowledge-Based Systems, 24*(7), 1075-1095. doi:10.1016/j.knosys.2011.05.001

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental Study of Inequality and

Unpredictability in an Artificial Cultural Market. *Science, 311*(5762), 854-856. doi:10.1126/science.1121066

Schulze, C. (2003). Advertising in the Sznajd Marketing Model. *International Journal of Modern Physics C, 14*(1), 95-98. doi:10.1142/S0129183103004255

Shalizi, C. R., & Thomas, A. C. (2011). Homophily and Contagion Are Generically Confounded in Observational Social Network Studies. *Sociological Methods & Research, 40*(2), 211-239. doi:10.1177/0049124111404820

Sidharthan, R., & Bhat, C. R. (2012). Incorporating Spatial Dynamics and Temporal Dependency in Land Use Change Models. *Geographical Analysis, 44*(4), 321-349. doi:10.1111/j.1538-4632.2012.00854.x

Simons, H. W. (2001). *Persuasion in society*. Thousand Oaks, Calif.: Sage Publications.

Sîrbu, A., Loreto, V., Servedio, V. D., & Tria, F. (2013). Cohesion, Consensus and Extreme Information in Opinion Dynamics. *Advances in Complex Systems, 16*(6), 1350035. doi:10.1142/S0219525913500355

Solow, A. R. (1990). A Method for Approximating Multivariate Normal Orthant Probabilities. *Journal of Statistical Computation and Simulation, 37*(3), 225-229. doi:10.1080/00949659008811306

Sznajd-Weron, K., & Weron, R. (2003). How Effective is Advertising in Duopoly Markets? *Physica A: Statistical Mechanics and its Applications, 324*(1), 437-444. doi:10.1016/S0378-4371(02)01904-0

Teixeira–Pinto, A., & Harezlak, J. (2013). Factorization and latent variable models for joint analysis of binary and continuous outcomes. In A. R. De Leon & K. C. Chough (Eds.), *Analysis of mixed data: methods & applications* (pp. 81-91). Boca Raton, FL: CRC Press/Taylor & Francis Group.

Temme, D., Paulssen, M., & Dannewald, T. (2008). Incorporating Latent Variables into Discrete Choice Models—A Simultaneous Estimation Approach Using SEM Software. *Business Research, 1*(2), 220-237. doi:10.1007/BF03343535

Thompson, T. L., & Mintzes, J. J. (2002). Cognitive Structure and the Affective Domain: On Knowing and Feeling in Biology. *International Journal of Science Education, 24*(6), 645-660. doi:10.1080/09500690110110115

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge: Cambridge university press.

Tran, M. (2012). Agent-Behaviour and Network Influence on Energy Innovation Diffusion. *Communications in Nonlinear Science and Numerical Simulation, 17*(9), 3682-3695. doi:10.1016/j.cnsns.2012.01.016

Tucker, C. (2008). Identifying Formal and Informal Influence in Technology Adoption with Network Externalities. *Management Science, 54*(12), 2024-2038. doi:10.1287/mnsc.1080.0897

Tussyadiah, I. P. (2016). The Influence of Innovativeness on On-Site Smartphone Use Among American Travelers: Implications for Context-Based Push Marketing. *Journal of Travel & Tourism Marketing, 33*(6), 806-823. doi:10.1080/10548408.2015.1068263

Vandecasteele, B., & Geuens, M. (2010). Motivated Consumer Innovativeness: Concept, Measurement, and Validation. *International Journal of Research in Marketing, 27*(4), 308-318. doi:10.1016/j.ijresmar.2010.08.004

Varin, C., & Vidoni, P. (2005). A Note on Composite Likelihood Inference and Model Selection. *Biometrika, 92*(3), 519-528. doi:10.1093/biomet/92.3.519

Varma, V. S., Morarescu, I.-C., Lasaulce, S., & Martin, S. (2017). *Opinion dynamics aware marketing strategies in duopolies.* Paper presented at the 2017 IEEE 56th Annual Conference on Decision and Control (CDC).

Vij, A., & Walker, J. L. (2014). Hybrid choice models: The identification problem. In S. Hess & A. Daly (Eds.), *Handbook of choice modelling*. Cheltenham,UK: Edward Elgar Publishing.

Vinayak, P., Dias, F. F., Astroza, S., Bhat, C. R., Pendyala, R. M., & Garikapati, V. M. (2018). Accounting for Multi-Dimensional Dependencies among Decision-Makers Within a Generalized Model Framework: An Application to Understanding Shared Mobility Service Usage Levels. *Transport Policy, 72*, 129-137. doi:10.1016/j.tranpol.2018.09.013

Wang, J., Aribarg, A., & Atchade, Y. F. (2013). Modeling Choice Interdependence in a Social Network. *Marketing Science, 32*(6), 977-997. doi:10.1287/mksc.2013.0811

Wang, L., & Mendel, J. M. (2016). Fuzzy Opinion Networks: A Mathematical Framework for the Evolution of Opinions and their Uncertainties Across Social Networks. *IEEE Transactions on Fuzzy Systems, 24*(4), 880-905.

doi:10.1109/TFUZZ.2015.2486816

Xiong, H., Payne, D., & Kinsella, S. (2016). Peer Effects in the Diffusion of Innovations: Theory and Simulation. *Journal of Behavioral and Experimental Economics, 63*, 1-13. doi:10.1016/j.socec.2016.04.017

Yang, S., & Allenby, G. M. (2003). Modeling Interdependent Consumer Preferences. *Journal of Marketing Research, 40*(3), 282-294. doi:10.1509/jmkr.40.3.282.19240

Zajonc, R. B. (1960). The Process of Cognitive Tuning in Communication. *The Journal of Abnormal and Social Psychology, 61*(2), 159-167. doi:10.1037/h0047987

Zhang, H., & Vorobeychik, Y. (2017). Empirically Grounded Agent-Based Models of Innovation Diffusion: A Critical Review. *Artificial Intelligence Review*, 1-35.

Zhang, J., & Liu, P. (2012). Rational Herding in Microloan Markets. *Management Science, 58*(5), 892-912. doi:10.1287/mnsc.1110.1459

Zhao, Y., & Joe, H. (2005). Composite Likelihood Estimation in Multivariate Data Analysis. *Canadian Journal of Statistics, 33*(3), 335-356. doi:10.1002/cjs.5540330303

Appendix 1: Specification of Latent Psychological Variables

The following specifications can be find in Bhat et al. (2016). There are L latent psychological variables with index $l = (1,2,\ldots,L)$ and Q individuals with index $q = (1,2,\ldots,Q)$. The lth latent psychological variable for individual q is noted as z^*_{ql} which can be defined as:

$$z^*_{ql} = \alpha'_l s_q + \delta_l \sum_{q'=1}^{Q} w_{qq'} z^*_{q'l} + \eta_{ql} \quad \cdots\cdots \text{Eq. (A1.1)}$$

where s_q is a $(F \times 1)$ vector of observed attitude and opinion related exogenous variables, α_l is a $(F \times 1)$ vector of corresponding coefficients, and η_{ql} is a standard normally distributed error term. δ_l is the social interaction parameter where $0 < \delta < 1$. The reason to bound the social interaction parameter between 0 and 1 is because individuals' attitudes are likely to be positively affected by social interaction. $w_{qq'}$ is the qq'th element of weight matrix W, which is a $(Q \times Q)$ row normalized matrix with $w_{qq} = 0$ and $\sum_{q' \neq q}^{Q} w_{qq'} = 1 \quad \forall q$. The weight matrix determines the closeness between individuals. This is a general term of social dependencies and can be defined in various ways with exogeneous variables as long as it is row normalized. In order to define the Eq. (A1.1) in a matrix form, following notations are defined:

$z_q^* = \left(z_{q1}^*, z_{q2}^*, \ldots, z_{qL}^*\right)'$ a $\left(L \times 1\right)$ vector, $z^* = \left(z_1^{*\prime}, z_2^{*\prime}, \ldots, z_Q^{*\prime}\right)'$ a $\left(QL \times 1\right)$ vector,

$\alpha = \left(\alpha_1', \alpha_2', \ldots, \alpha_L'\right)'$ a $\left(LF \times 1\right)$ vector,

$\tilde{s}_q = IDEN_L \otimes s_q'$ a $\left(QL \times LF\right)$ matrix, $\tilde{s} = \left(\tilde{s}_1', \tilde{s}_2', \ldots, \tilde{s}_Q'\right)'$ a $\left(QL \times LF\right)$ matrix,

$\eta_q = \left(\eta_{q1}, \eta_{q2}, \ldots, \eta_{qL}\right)'$ a $\left(L \times 1\right)$ vector, $\eta = \left(\eta_1', \eta_2', \ldots, \eta_Q'\right)'$ a $\left(QL \times 1\right)$ vector,

$\delta_q = \left(\delta_{q1}, \delta_{q2}, \ldots, \delta_{qL}\right)'$ a $\left(L \times 1\right)$ vector, $\delta = \left(\delta_1', \delta_2', \ldots, \delta_Q'\right)'$ a $\left(QL \times 1\right)$ vector,

where $IDEN_Q$ indicates an identity matrix with size Q, and $\otimes$ indicates the Kronecker product. Then the matrix form of Eq. (A1.1) for all individuals can be written as:

$$z^* = S\tilde{s}\alpha + S\eta \quad \cdots\cdots \text{Eq. (A1.2)}$$

where $S = \left[IDEN_{QL} - \delta.*\left(W \otimes IDEN_L\right)\right]^{-1}$, the notation "$.*$" represents element by element product. The social interaction effect is captured in S. Assume that η_q follows a standard multivariate normal (MVN) distribution: $\eta_q \sim MVN_L\left(\mathbf{0}_L, \Gamma\right)$, where 0_L

indicates an $(L \times 1)$ column vector with all elements equals to 0. Assume that there are correlations between latent psychological variables of an individual and no correlation between individuals, i.e., $\mathrm{Cov}(\eta_q, \eta_{q'}) = 0$, $\forall q \neq q'$. Then it is clear that z^* also follows an MVN distribution: $z^* \sim MVN_{QL}(B, \Xi)$, where $B = S\tilde{s}\alpha$ and $\Xi = S\left[IDEN_Q \otimes \Gamma\right]S'$.

Appendix 2: Specification for Likelihood Function

The following specifications can be find in Bhat et al. (2016). In case of the estimation of nominal variables, we consider the utility difference between chosen alternative and other non-chosen alternatives: $u_{qgi_g m_{qg}} = \left(U_{qgi_g} - U_{qgm_{qg}}\right)$, where $i_g \neq m_{qg}$ (note that m_{qg} indicates the chosen alternative of individual q for g th nominal variable). To rearrange the reduced form of the model, define following matrix and vectors:

$$u_{qg} = \left[\left(u_{qg1m_{qg}}, u_{qg2m_{qg}}, \ldots, u_{qgI_g m_{qg}}\right)' : i_g \neq m_g\right] \text{ a } \left(\left(I_g - 1\right) \times 1\right) \text{ vector,}$$

$$u_q = \left[\left(u_{q1}\right)', \left(u_{q2}\right)', \ldots, \left(u_{qG}\right)'\right]' \text{ a } \left(\tilde{G} \times 1\right) \text{ vector,}$$

$$\left(yu\right)_q = \left[\breve{y}_q{}', \tilde{y}_q^{*\prime}, u_q{}'\right]' \text{ a } \left(\left(E + \tilde{G}\right) \times 1\right) \text{ vector,}$$

$$yu = \left[\left(yu\right)_1{}', \left(yu\right)_2{}', \ldots, \left(yu\right)_Q{}'\right]' \text{ a } \left(Q\left(E + \tilde{G}\right) \times 1\right) \text{ vector.}$$

Define a selection matrix M of size $\left[Q\left(E+\tilde{G}\right) \times Q\left(E+\tilde{G}\right)\right]$ and transfer yU into yu (See Appendix 5). Then the distribution of yu can be defined as:

$$yu \sim MVN_{Q\left(E+\tilde{G}\right)}\left[\tilde{B}, \tilde{\Omega}\right],$$

where $\tilde{B} = M \cdot \left(\vec{\gamma} + \vec{d}\left\{S\tilde{s}\alpha\right\}\right)x$

and $\tilde{\Omega} = M \cdot \left[\left(\vec{d}\tilde{x}\right) \cdot S\left(IDEN_Q \otimes \Gamma\right)S' \cdot \left(\vec{d}\tilde{x}\right)' + IDEN_Q \otimes \Sigma\right] \cdot M'$.

Then, divide yu into two parts: continuous outcomes $\left(\breve{y}'\right)$ and non-continuous outcomes $\left(\tilde{y}^{*\prime}, u'\right)$. Redefined $yu = \left(y', \tilde{u}'\right)$, where $y = \breve{y}'$ and $\tilde{u} = \left(\tilde{y}^{*\prime}, u'\right)'$. Also define a selection matrix R of size $\left[Q\left(E+\tilde{G}\right) \times Q\left(E+\tilde{G}\right)\right]$ and divide $\tilde{B}$ and $\tilde{\Omega}$ into: $\tilde{B} = \begin{bmatrix} \tilde{B}_y \\ \tilde{B}_{\tilde{u}} \end{bmatrix}$ and $\tilde{\Omega} = \begin{bmatrix} \tilde{\Omega}_y & \tilde{\Omega}_{y\tilde{u}} \\ \tilde{\Omega}'_{y\tilde{u}} & \tilde{\Omega}_{\tilde{u}} \end{bmatrix}$ (see Appendix 5).

The likelihood function can be decomposed into the joint distribution of a product of marginal distribution and conditional distributions. The conditional distribution of $\tilde{u}|y \sim MVN\left(\vec{B}, \vec{\Omega}\right)$, where $\vec{B} = \tilde{B}_{\tilde{u}} + \tilde{\Omega}'_{y\tilde{u}}\tilde{\Omega}_y^{-1}\left(y - \tilde{B}_y\right)$ and $\vec{\Omega}_{\tilde{u}} = \tilde{\Omega}_{\tilde{u}} - \tilde{\Omega}'_{y\tilde{u}}\tilde{\Omega}_y^{-1}\tilde{\Omega}_{y\tilde{u}}$. Then define the threshold of discontinuous outcomes as: $\vec{\psi}_{low} = \left[\tilde{\psi}_{low}', \left(-\infty_{Q\tilde{G}}\right)'\right]'$ a $\left(Q\left(N+\tilde{G}\right) \times 1\right)$ vector and $\vec{\psi}_{up} = \left[\tilde{\psi}_{up}', \left(\mathbf{0}_{Q\tilde{G}}\right)'\right]'$ a $\left(Q\left(N+\tilde{G}\right) \times 1\right)$ vector, where $-\infty_{Q\tilde{G}}$ is a $\left(Q\tilde{G} \times 1\right)$ vector with all elements equal to negative infinities.

The likelihood function can be written as:

$$\begin{aligned} L(\lambda) &= f_{QH}\left(y \middle| \tilde{B}_y, \tilde{\Omega}_y\right) \times \Pr\left[\overleftrightarrow{\psi}_{low} < \tilde{u} < \overleftrightarrow{\psi}_{up}\right] \\ &= f_{QH}\left(y \middle| \tilde{B}_y, \tilde{\Omega}_y\right) \times \int_r f_{Q\left(N+\overleftrightarrow{G}\right)}\left(r \middle| \overleftrightarrow{B}_{\tilde{u}}, \overleftrightarrow{\Omega}_{\tilde{u}}\right) dr \end{aligned} \qquad \text{Eq. (A2.1)}$$

Appendix 3: Specification for Composite Marginal Likelihood Function

The following specifications can be find in online supplement of Bhat et al. (2016). In order to derive the individual pair (qq') CML function in Eq.(3.3.1), we need to select the necessary elements in $\tilde{B}$ and $\tilde{\Omega}$, and store them into $\tilde{B}_{qq'}$ and $\tilde{\Omega}_{qq'}$ (as well as $\breve{\psi}_{qq',low}$ and $\breve{\psi}_{qq',up}$ from $\breve{\psi}_{low}$ and $\breve{\psi}_{up}$ respectively). To do so, we need to define a selection matrix $D_{qq'}$ of size $\left[2\left(E+\tilde{G}\right)\times Q\left(E+\tilde{G}\right)\right]$, and another selection matrix $V_{qq'}$ of size $\left[2\left(N+\tilde{G}\right)\times Q\left(N+\tilde{G}\right)\right]$ (see Appendix 5). Then define following matrix and vectors:

$\tilde{B}_{qq'} = D_{qq'}\tilde{B}$ a $\left(2\left(E+\tilde{G}\right)\times 1\right)$ vector

$\tilde{\Omega}_{qq'} = D_{qq'}\tilde{\Omega}D_{qq'}{}'$ a $\left(2\left(E+\tilde{G}\right)\times 2\left(E+\tilde{G}\right)\right)$ matrix

$\breve{\psi}_{qq',low} = V_{qq'}\breve{\psi}_{low}$ a $\left(2\left(N+\tilde{G}\right)\times 1\right)$ vector

$\breve{\psi}_{qq',up} = V_{qq'}\breve{\psi}_{up}$ a $\left(2\left(N+\tilde{G}\right)\times 1\right)$ vector

Then in order to divide the $\tilde{B}_{qq'}$ and $\tilde{\Omega}_{qq'}$ into two parts (continuous outcomes and dis-continuous outcomes) we further define a selection matrix $R_{qq',y}$ of size

$\left[2H \times Q\left(E+\tilde{G}\right)\right]$ and another selection matrix $R_{qq',\tilde{u}}$ of size $\left[2\left(N+\tilde{G}\right) \times Q\left(E+\tilde{G}\right)\right]$ (see Appendix 5). Define following matrix and vectors:

$\tilde{B}_{qq',y} = R_{qq',y}\tilde{B}_{qq'}$ a $\left(2H \times 1\right)$ vector,

$\tilde{B}_{qq',\tilde{u}} = R_{qq',\tilde{u}}\tilde{B}_{qq'}$ a $\left(2\left(N+\tilde{G}\right) \times 1\right)$ vector,

$\tilde{\Omega}_{qq',y} = R_{qq',y}\tilde{\Omega}_{qq'}R'_{qq',y}$ a $\left(2H \times 2H\right)$ matrix,

$\tilde{\Omega}_{qq',\tilde{u}} = R_{qq',\tilde{u}}\tilde{\Omega}_{qq'}R'_{qq',\tilde{u}}$ a $\left(2\left(N+\tilde{G}\right) \times 2\left(N+\tilde{G}\right)\right)$ matrix.

Then $\tilde{B}_{qq'} = \begin{bmatrix} \tilde{B}_{qq',y} \\ \tilde{B}_{qq',\tilde{u}} \end{bmatrix}$ and $\tilde{\Omega}_{qq'} = \begin{bmatrix} \tilde{\Omega}_{qq',y} & \tilde{\Omega}_{qq',y\tilde{u}} \\ \tilde{\Omega}'_{qq',y\tilde{u}} & \tilde{\Omega}_{qq',\tilde{u}} \end{bmatrix}$. The conditional distribution of individual pair qq' is $\tilde{u}_{qq'} \left| y_{qq'} \sim MVN\left(\ddot{B}_{qq',\tilde{u}}, \ddot{\Omega}_{qq',\tilde{u}}\right)\right.$, where $\ddot{B}_{qq',\tilde{u}} = \tilde{B}_{qq',\tilde{u}} + \tilde{\Omega}'_{qq',y\tilde{u}}\tilde{\Omega}^{-1}_{qq',y}\left(y_{qq'} - \tilde{B}_{qq',y}\right)$ and $\ddot{\Omega}_{qq',\tilde{u}} = \tilde{\Omega}_{qq',\tilde{u}} - \tilde{\Omega}'_{qq',y\tilde{u}}\tilde{\Omega}^{-1}_{qq',y}\tilde{\Omega}_{qq',y\tilde{u}}$.

In order to match elements in individual pair CML of dis-continuous part, we further define a selection matrix $F_{qq'}$ of size $\left[2\left(N+\tilde{G}\right) \times 2\left(N+\tilde{G}\right)\right]$ (see Appendix 5) and define following matrix and vectors:

$\vec{B}_{qq',\tilde{u}} = F_{qq'}\ddot{B}_{qq',\tilde{u}}$, $\vec{\Omega}_{qq',\tilde{u}} = F_{qq'}\ddot{\Omega}_{qq',\tilde{u}}F'_{qq'}$,

$\vec{\psi}_{qq',low} = F_{qq'}\ddot{\psi}_{qq',low}$, $\vec{\psi}_{qq',up} = F_{qq'}\ddot{\psi}_{qq,up}$.

Replace the last $2\tilde{G}$ element in $\vec{\psi}_{qq',low}$ with 0: $\vec{\psi}_{qq',low}\left[2N+1:2\left(N+\tilde{G}\right)\right] = \mathbf{0}_{2\tilde{G}}$. Then define ω_Δ as the diagonal matrix of standard deviation of the matrix Δ, and ϕ_H

as the multivariate standard normal density function of dimension H, Φ_E as the multivariate standard normal cumulative distribution function of dimension E. Define a selection matrix H_{vg} of size $\left[I_g \times 2\left(N+\tilde{G}\right)\right]$ (see Appendix 5) and define:

$$\widehat{\Omega}_{qq',vg} = H_{vg}\overleftrightarrow{\Omega}_{qq',\tilde{u}}H'_{vg},$$

$$\mu_{v,up} = \frac{\left[\overleftrightarrow{\psi}_{qq',up}\right]_v - \left[\overleftrightarrow{B}_{qq',\tilde{u}}\right]_v}{\sqrt{\left[\overleftrightarrow{\Omega}_{qq',\tilde{u}}\right]_{vv}}}$$

$$\mu_{v,low} = \frac{\left[\overleftrightarrow{\psi}_{qq',low}\right]_v - \left[\overleftrightarrow{B}_{qq',\tilde{u}}\right]_v}{\sqrt{\left[\overleftrightarrow{\Omega}_{qq',\tilde{u}}\right]_{vv}}}$$

$$\rho_{vv'} = \frac{\left[\overleftrightarrow{\Omega}_{qq',\tilde{u}}\right]_{vv'}}{\sqrt{\left[\overleftrightarrow{\Omega}_{qq',\tilde{u}}\right]_{vv} \cdot \left[\overleftrightarrow{\Omega}_{qq',\tilde{u}}\right]_{v'v'}}}$$

where $\left[\cdot\right]_v$ represents the vth element of the vector, and $\left[\cdot\right]_{vv'}$ represents the vv'th element of the matrix. The MACML of the proposed model can be defined as:

$$L_{MACML}\left(\lambda\right) = \prod_{q=1}^{Q-1}\prod_{q'=q+1}^{Q} L_{MACML,qq'}\left(\lambda\right)$$

where

$$
\begin{aligned}
L_{MACML,qq'}(\lambda) = & \left(\prod_{h=1}^{H} \omega_{\tilde{\Omega}_{qq',y}} \right)^{-1} \phi_H \left(\left[\omega_{\tilde{\Omega}_{qq',y}} \right]^{-1} \left[y_{qq'} - \tilde{B}_{qq',y} \right]; \tilde{\Omega}^*_{qq',y} \right) \\
& \times \left(\prod_{v=1}^{N-1} \prod_{v'=v+1}^{N} \left[\begin{array}{l} \Phi_2\left(\mu_{v,up}, \mu_{v',up}, \rho_{vv'}\right) - \Phi_2\left(\mu_{v,up}, \mu_{v',low}, \rho_{vv'}\right) \\ -\Phi_2\left(\mu_{v,low}, \mu_{v',up}, \rho_{vv'}\right) + \Phi_2\left(\mu_{v,low}, \mu_{v',low}, \rho_{vv'}\right) \end{array} \right] \right) \\
& \times \left(\prod_{v=1}^{N} \prod_{g=1}^{G} \left[\begin{array}{l} \Phi_{I_g}\left[\omega^{-1}_{\widehat{\Omega}_{qq',vg}} H_{vg} \left\{ \vec{\ddot{\psi}}_{qq',up} - \vec{\ddot{B}}_{qq',\tilde{u}} \right\}; \widehat{\Omega}^*_{qq',vg} \right] \\ -\Phi_{I_g}\left[\omega^{-1}_{\widehat{\Omega}_{qq',vg}} H_{vg} \left\{ \vec{\ddot{\psi}}_{qq',low} - \vec{\ddot{B}}_{qq',\tilde{u}} \right\}; \widehat{\Omega}^*_{qq',vg} \right] \end{array} \right] \right)
\end{aligned} \quad \cdot \text{Eq. (A3.1)}
$$

Appendix 4: Specification for The Estimation Approach of Covariance Matrix

The MACML estimator in Eq. (3.3.4) is maximized by compute the pairwise loglikelihood for $Q(Q-1)/2$ individual pairs. The asymptotic covariance matrix of the parameter $V_{MACML}(\lambda)$ can be estimated by the inverse of Godambe (1960) sandwich information matrix.

$$V_{MACML}(\lambda)=\frac{\left[\hat{G}(\lambda)\right]^{-1}}{Q}=\frac{\left[\hat{H}^{-1}\right]\left[\hat{J}\right]\left[\hat{H}^{-1}\right]}{Q} \quad \cdots\cdots\cdots\cdots \text{Eq. (A4.1)}$$

The calculation of the Hessian matrix $(\hat{H})$ and the Jacobian matrix $(\hat{J})$ in the Eq. (A4.1) is introduced in Zhao and Joe (2005) and Sidharthan and Bhat (2012). According to windows sampling method (Heagerty & Lumley, 2000) Bhat (2011) suggest to use the information about the distances between individuals. The Jacobian matrix can be calculated as:

$$J=\frac{\tilde{W}}{D}\left[\sum_{d=1}^{D}\left[\frac{1}{N_d}\left(\left[S_{CML,d}(\theta)\right]\left[S_{CML,d}(\theta)\right]'\right)_{\hat{\theta}}\right]\right] \quad \cdots\cdots\cdots\cdots \text{Eq. (A4.2)}$$

where $\tilde{W}$ represents the total number of pairs in the likelihood function and D represents the observational units from the Q individuals. In order to ensure the covariance matrix be positive definite, the likelihood is estimated in terms of the Cholesky-decomposed elements.

Appendix 5: Specification of Selection Matrix

The following specifications can be find in online supplement of Bhat et al. (2016). Assume a situation with $Q=2$, $H=1$, $N=2$, $G=1$ and $I_1=3$.

Selection matrix of M:

First, create a matrix of size $\left[Q\left(E+\tilde{G}\right)\times Q\left(E+\ddot{G}\right)\right]$ with all element equals to 0. Every individual q occupies the blog of size $\left[\left(E+\tilde{G}\right)\times\left(E+\ddot{G}\right)\right]$ in $\left[\left(q-1\right)\cdot\left(E+\tilde{G}\right)+1\right]$ to $\left[\left(q-1\right)\cdot\left(E+\tilde{G}\right)+\left(E+\tilde{G}\right)\right]$ rows and $\left[\left(q-1\right)\cdot\left(E+\ddot{G}\right)+1\right]$ to $\left[\left(q-1\right)\cdot\left(E+\ddot{G}\right)+\left(E+\ddot{G}\right)\right]$ columns of M. Second, insert an identity matrix of size E into the first E rows and E columns of q 's blog. Third, insert an identity matrix of size $\left(I_1-1\right)$ (after insert a column of "–1" corresponding to the chosen alternative) in $E+1$ to $E+I_1-1$ rows and $E+1$ to $E+I_1$ columns of individual q 's blog.

For example, assume the first individual choose alternative 2 and the second individual choose alternative 3, then the selection matrix M is:

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

Selection matrix of R_y and $R_{\tilde{u}}$:

First, create a matrix R of size $\left[Q\left(E+\tilde{G}\right)\times Q\left(E+\tilde{G}\right)\right]$ with all element equals to 0. Second, for every individual q, insert an identity matrix of size H in $\left[\left(q-1\right)\cdot H+1\right]$ to $\left[\left(q-1\right)\cdot H+H\right]$ rows and $\left[\left(q-1\right)\cdot\left(E+\tilde{G}\right)+1\right]$ to $\left[\left(q-1\right)\cdot\left(E+\tilde{G}\right)+\left(E+\tilde{G}\right)\right]$ columns. Third, for every individual q, insert an identity matrix of size $\left(N+\tilde{G}\right)$ in $\left[QH+\left(q-1\right)\cdot\left(E+\tilde{G}-H\right)+1\right]$ to $\left[QH+\left(q-1\right)\cdot\left(E+\tilde{G}-H\right)+\left(E+\tilde{G}-H\right)\right]$ rows and $\left[\left(q-1\right)\cdot\left(E+\tilde{G}\right)+1\right]$ to $\left[\left(q-1\right)\cdot\left(E+\tilde{G}\right)+\left(E+\tilde{G}\right)\right]$ columns. Forth, divide R into two parts: $R_y = R\left[1:QH,1:Q\left(E+\tilde{G}\right)\right]$ and $R_{\tilde{u}} = R\left[QH+1:Q\left(E+\tilde{G}\right),1:Q\left(E+\tilde{G}\right)\right]$.

For example:

$$R = \left[\begin{array}{cccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{c} R_y \\ \hline R_{\tilde{u}} \end{array}\right]$$

Selection matrix of $D_{qq'}$:

First, create a matrix of size $\left[2\left(E+\tilde{G}\right)\times Q\left(E+\tilde{G}\right)\right]$ with all element equals to 0. Second, insert an identity matrix of size $\left(E+\tilde{G}\right)$ in first $\left(E+\tilde{G}\right)$ rows and $\left[\left(q-1\right)\cdot\left(E+\tilde{G}\right)+1\right]$ to $\left[\left(q-1\right)\cdot\left(E+\tilde{G}\right)+\left(E+\tilde{G}\right)\right]$ columns. Third, insert another identity matrix of size $\left(E+\tilde{G}\right)$ in $\left[E+\tilde{G}+1\right]$ to $\left[2\left(E+\tilde{G}\right)\right]$ rows and $\left[\left(q'-1\right)\cdot\left(E+\tilde{G}\right)+1\right]$ to $\left[\left(q'-1\right)\cdot\left(E+\tilde{G}\right)+\left(E+\tilde{G}\right)\right]$ columns.

For example, assume $Q=3$, $q=1$ and $q'=3$:

$$D_{qq'}=\begin{bmatrix} 1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&1&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&1&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&1&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&1 \end{bmatrix}$$

Selection matrix of $V_{qq'}$:

First, create a matrix of size $\left[2\left(N+\tilde{G}\right)\times Q\left(N+\tilde{G}\right)\right]$ with all element equals to 0. Second, insert an identity matrix of size $\left(N+\tilde{G}\right)$ in first $\left(N+\tilde{G}\right)$ rows and $\left[\left(q-1\right)\cdot\left(N+\tilde{G}\right)+1\right]$ to $\left[\left(q-1\right)\cdot\left(N+\tilde{G}\right)+\left(N+\tilde{G}\right)\right]$ columns. Third, insert another identity matrix of size $\left(N+\tilde{G}\right)$ in $\left[N+\tilde{G}+1\right]$ to $\left[2\left(N+\tilde{G}\right)\right]$ rows and $\left[\left(q'-1\right)\cdot\left(N+\tilde{G}\right)+1\right]$ to $\left[\left(q'-1\right)\cdot\left(N+\tilde{G}\right)+\left(N+\tilde{G}\right)\right]$ columns.

For example, assume $Q = 3$, $q = 1$ and $q' = 3$:

$$V_{qq'} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Selection matrix of $R_{qq',y}$ and $R_{qq',\tilde{u}}$

First, create a $R_{qq'}$ matrix of size $\left[2\left(E+\tilde{G}\right)\times Q\left(E+\tilde{G}\right)\right]$ with all element equals to 0. Second, for individual q insert an identity matrix of size H in first H rows and $\left[\left(q-1\right)\cdot\left(E+\tilde{G}\right)+1\right]$ to $\left[\left(q-1\right)\cdot\left(E+\tilde{G}\right)+H\right]$ columns. Third, insert another identity matrix of size $\left(N+\tilde{G}\right)$ in $\left[2H+1\right]$ to $\left[2H+\left(E+\tilde{G}-H\right)\right]$ rows and $\left[\left(q-1\right)\cdot\left(E+\tilde{G}\right)+1\right]$ to $\left[\left(q-1\right)\cdot\left(E+\tilde{G}\right)+\left(E+\tilde{G}\right)\right]$ columns. Fourth, for individual q', insert an identity matrix of size H in $\left[H+1\right]$ to $\left(2H\right)$ rows. Fifth, insert another identity matrix of size $\left(N+\tilde{G}\right)$ in $\left[2H+\left(N+\tilde{G}\right)+1\right]$ to $\left[2H+2\left(N+\tilde{G}\right)\right]$ rows and $\left[\left(q'-1\right)\cdot\left(E+\tilde{G}\right)+1\right]$ to $\left[\left(q'-1\right)\cdot\left(E+\tilde{G}\right)+\left(E+\tilde{G}\right)\right]$ columns. Then, divide $R_{qq'}$ into two parts: $R_{qq',y} = R_{qq'}\left[1:2H,1:Q\left(E+\tilde{G}\right)\right]$ and

$R_{qq',\tilde{u}} = R_{qq'}\left[2H+1:2\left(E+\tilde{G}\right),1:Q\left(E+\tilde{G}\right)\right]$

For example, assume $Q=3$, $q=1$ and $q'=3$:

$$R_{qq} = \left[\begin{array}{ccccccccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{c} R_{qq',y} \\ \hline R_{qq',\tilde{u}} \end{array}\right]$$

Selection matrix of $F_{qq'}$:

First, create a matrix of size $\left[2\left(N+\tilde{G}\right)\times 2\left(N+\tilde{G}\right)\right]$ with all element equals to 0. Second, insert an identity matrix of size N in first N rows and N columns. Third, insert another identity matrix of size N in $\left[N+1\right]$ to $\left[2N\right]$ rows and $\left[N+\tilde{G}+1\right]$ to $\left[\left(N+\tilde{G}\right)+N\right]$ columns. Fourth, insert an identity matrix of size $\tilde{G}$ in $\left[2N+1\right]$ to $\left[2N+\tilde{G}\right]$ rows and $\left[N+1\right]$ to $\left(N+\tilde{G}\right)$ columns. Fifth, insert another identity matrix of size $\tilde{G}$ in $\left[2N+\tilde{G}+1\right]$ to $\left[2\left(N+\tilde{G}\right)\right]$ rows and $\left[\left(N+\tilde{G}\right)+N+1\right]$ to

$\left[2\left(N+\tilde{G}\right)\right]$ columns.

For example:

$$F_{qq'} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Selection matrix of H_{vg} :

First, create a matrix of size $\left[I_g \times 2\left(N+\tilde{G}\right)\right]$ with all element equals to 0. Second, insert a "1" in the first rows and v th column. Third, insert another identity matrix of size $\left(I_g-1\right)$ in last $\left(I_g-1\right)$ rows and $\left[2N+\left[\sum_{j=1}^{g-1}\left(I_j-1\right)+1\right]^{th}\right]$ to $\left[2N+\left[\sum_{j=1}^{g}\left(I_j-1\right)\right]^{th}\right]$ columns (note that $\sum_{j=1}^{0}\left(I_j-1\right)=0$).

For example, assume $v=1$ and $g=1$:

$$H_{vg} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Assume $v = 2$ and $g = 2$:

$$H_{vg} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Appendix 6: Full Estimation Results

In the unsocial and social GHDM result, α indicates the coefficients of exogenous variables of the latent psychological variables. For example, α_{12} indicates the coefficient of 2nd exogeneous variable to the 1st latent psychological variable. δ indicates the social interaction and l_{Γ} indicates the correlation between the latent psychological variables. γ represents the pure effect of the explanatory variable and d represents the interactions between explanatory variables and latent psychological variables. For example, $\tilde{\gamma}_{\tilde{y}21}$ represents the coefficient of 1st explanatory variable of 2nd ordinal outcomes, $\widehat{d}_{U1,12}$ represents the interaction effect of the 2nd latent psychological variable and the 1st explanatory variable in the 1st nominal outcome. ψ indicates the upper threshold of the ordinal outcomes. ε and l_{Λ} represents the error term of continuous outcomes and nominal outcomes respectively.

Table 13. Estimation Result of Proposed Unsocial GHDM Model

Parameters	Est.	T.stat.	Parameters	Est.	T.stat.
α_{11}	0.343	19.174	$\breve{d}_{\breve{y}1,11}$	1.140	8.297
α_{12}	0.062	3.891	$\breve{d}_{\breve{y}1,12}$	1.633	20.941
α_{13}	0.082	5.374	$\breve{d}_{\breve{y}1,21}$	5.058	16.398
α_{14}	0.082	6.022	$\breve{d}_{\breve{y}1,22}$	5.654	24.237
α_{21}	0.202	13.046	$\tilde{d}_{\tilde{y}1,11}$	-0.268	-30.65
α_{22}	0.044	2.623	$\tilde{d}_{\tilde{y}1,12}$	-0.059	-5.892
α_{23}	0.224	11.844	$\tilde{d}_{\tilde{y}2,11}$	-0.055	-0.481
α_{24}	0.131	7.453	$\tilde{d}_{\tilde{y}2,12}$	0.453	4.953
$l_{\Gamma 1}$	0.262	5.881	$\tilde{d}_{\tilde{y}2,21}$	-1.153	-16.965
$\breve{\gamma}_{\breve{y}1,1}$	1.671	1.798	$\tilde{d}_{\tilde{y}2,22}$	0.232	4.083
$\breve{\gamma}_{\breve{y}1,2}$	0.936	3.943	$\widehat{d}_{U1,11}$	0.168	7.623
$\breve{\gamma}_{\breve{y}1,3}$	2.408	4.741	$\widehat{d}_{U1,12}$	0.183	7.054
$\tilde{\gamma}_{\tilde{y}1,1}$	2.860	10.518	$\widehat{d}_{U1,21}$	0.709	7.961
$\tilde{\gamma}_{\tilde{y}1,2}$	-0.170	-9.534	$\widehat{d}_{U1,22}$	-0.289	-3.799

$\tilde{\gamma}_{\tilde{y}2,1}$	-0.206	-0.972	$\widehat{d}_{U1,31}$	0.106	2.155
$\tilde{\gamma}_{\tilde{y}2,2}$	-0.499	-2.307	$\widehat{d}_{U1,32}$	0.267	2.379
$\tilde{\gamma}_{\tilde{y}2,3}$	-1.020	-9.983	$\widehat{d}_{U1,41}$	0.658	6.645
$\widehat{\gamma}_{U1,1}$	-0.169	-7.673	$\widehat{d}_{U1,42}$	0.580	11.845
$\widehat{\gamma}_{U1,2}$	0.115	4.435	$\widehat{d}_{U1,51}$	0.233	2.082
$\widehat{\gamma}_{U1,3}$	1.395	63.391	$\widehat{d}_{U1,52}$	0.216	2.199
$\widehat{\gamma}_{U1,4}$	-0.302	-27.445	$\breve{\varepsilon}_{\tilde{y}1}$	0.821	3.957
$\widehat{\gamma}_{U1,5}$	0.219	4.459	$\psi_{\tilde{y}1}$	0.685	42.831
$\widehat{\gamma}_{U1,6}$	-0.330	-2.945	$\psi_{\tilde{y}2}$	0.789	13.259
$\widehat{\gamma}_{U1,7}$	0.168	1.716			
$l_{\Lambda 1}$	0.599	4.831			
$l_{\Lambda 2}$	1.123	3.2			

Table 14. Estimation Result of Proposed Social GHDM Model

Parameters	Est.	T.stat.	Parameters	Est.	T.stat.
α_{11}	0.304	10.134	$\breve{d}_{\breve{y}1,11}$	1.144	5.948
α_{12}	0.102	2.941	$\breve{d}_{\breve{y}1,12}$	1.628	1.697
α_{13}	0.056	2.118	$\breve{d}_{\breve{y}1,21}$	5.078	15.943
α_{14}	0.034	3.094	$\breve{d}_{\breve{y}1,22}$	5.67	1.992
α_{21}	0.131	6.094	$\tilde{d}_{\tilde{y}1,11}$	-0.141	-5.648
α_{22}	0.032	2.624	$\tilde{d}_{\tilde{y}1,12}$	-0.013	-2.103
α_{23}	0.153	6.915	$\tilde{d}_{\tilde{y}2,11}$	-0.023	-0.973
α_{24}	0.057	1.974	$\tilde{d}_{\tilde{y}2,12}$	0.48	0.914
$l_{\Gamma 1}$	0.242	5.314	$\tilde{d}_{\tilde{y}2,21}$	-1.095	-1.321
$\breve{\gamma}_{\breve{y}1,1}$	1.677	1.622	$\tilde{d}_{\tilde{y}2,22}$	0.286	4.004
$\breve{\gamma}_{\breve{y}1,2}$	0.932	1.994	$\widehat{d}_{U1,11}$	0.191	4.849
$\breve{\gamma}_{\breve{y}1,3}$	2.411	2.973	$\widehat{d}_{U1,12}$	0.206	1.982
$\tilde{\gamma}_{\tilde{y}1,1}$	2.784	4.831	$\widehat{d}_{U1,21}$	0.725	3.947
$\tilde{\gamma}_{\tilde{y}1,2}$	-0.058	-4.844	$\widehat{d}_{U1,22}$	-0.272	-1.947

$\tilde{\gamma}_{\tilde{y}2,1}$	-0.206	-0.684	$\widehat{d}_{U1,31}$	0.125	1.984
$\tilde{\gamma}_{\tilde{y}2,2}$	-0.496	-1.894	$\widehat{d}_{U1,32}$	0.29	1.893
$\tilde{\gamma}_{\tilde{y}2,3}$	-1.008	-9.736	$\widehat{d}_{U1,41}$	0.68	1.774
$\widehat{\gamma}_{U1,1}$	-0.169	-2.184	$\widehat{d}_{U1,42}$	0.603	8.648
$\widehat{\gamma}_{U1,2}$	0.113	2.194	$\widehat{d}_{U1,51}$	0.255	1.367
$\widehat{\gamma}_{U1,3}$	1.39	10.954	$\widehat{d}_{U1,52}$	0.236	1.849
$\widehat{\gamma}_{U1,4}$	-0.301	-12.008	$\breve{\varepsilon}_{\tilde{y}1}$	0.679	1.234
$\widehat{\gamma}_{U1,5}$	0.218	1.294	$\psi_{\tilde{y}1}$	2.02	21.558
$\widehat{\gamma}_{U1,6}$	-0.331	0.764	$\psi_{\tilde{y}2}$	0.455	10.355
$\widehat{\gamma}_{U1,7}$	0.17	0.694	δ_1	0.308	2.469
$l_{\Lambda 1}$	0.601	0.879	δ_2	0.589	2.891
$l_{\Lambda 2}$	1.625	0.674			

Appendix 7: Media Panel Survey Questionnaire

Questionnaires about consumer innovativeness are consisted in 2019 KISDI personal media panel investigation.

IV. 소비자혁신성 (Consumer innovativeness)

문21) 다음은 평소 새로운 기능이나 속성이 추가된 신제품이나 서비스의 구매행태에 관한 질문들입니다.
다음 각각의 문장에 대하여 어느 정도 동의 하시는지를 응답해 주시기 바랍니다.

문 번호	항 목		동의 정도				
			전혀 그렇지 않다 ①	그렇지 않다 ②	보통 이다 ③	그렇다 ④	매우 그렇다 ⑤
문21-1)	기능적 혁신성	내가 현재 사용하는 제품에 없는 **새로운 기능이 추가된 신제품**이 나오면 바로 구매하는 편이다	①	②	③	④	⑤
		기존 제품에 비해 **시간을 절약해 주는 신제품**이 출시되면 바로 구매하는 편이다	①	②	③	④	⑤
		앞으로 출시될 신제품이 **현재 내가 사용하는 제품보다 편리**하다면 즉시 구매하는 편이다	①	②	③	④	⑤
		업무를 간소화 할 수 있는 제품이 출시되면 지체 없이 바로 구매하는 편이다	①	②	③	④	⑤
문21-2)	쾌락적 혁신성	전에 알지 못했던 **새롭고 신기한 제품을 발견**하면 즐겁고 재미있어진다	①	②	③	④	⑤
		새롭고 신기한 제품을 갖는 것은 늘 설레고 흥분된다	①	②	③	④	⑤
		신기한 제품을 사용하는 것은 재밌고 기쁘다	①	②	③	④	⑤
		혁신적인 제품을 사용하는 것은 일상을 활기차게 해주는 자극제이다	①	②	③	④	⑤
문21-3)	사회적 혁신성	**다른 사람들과 나를 구별해주는 제품**을 좋아 한다	①	②	③	④	⑤
		다른 사람들이 사용한 적이 없는 제품을 먼저 사용하여 보다 뛰어나고 싶다	①	②	③	④	⑤
		다른 사람들에게 깊은 인상을 주는 신제품 사용하는 것을 좋아 한다	①	②	③	④	⑤
		누구라도 부러워하고 호기심 가질 만한 눈에 띄는 제품을 먼저 구매 혹은 사용하고 싶다	①	②	③	④	⑤
문21-4)	인지적 혁신성	제품 사용에 **많은 지식을 필요로 하는 신제품**이 출시되면 즉시 구매하는 편이다	①	②	③	④	⑤
		나는 **논리적인 생각을 필요로 하는 신제품**을 종종 사용한다	①	②	③	④	⑤
		생각을 많이 하고 지적 호기심을 자극하는 신제품을 즉시 사용한다	①	②	③	④	⑤
		신제품이 나의 분석적인 생각을 충족해 준다면 대부분 구입한다	①	②	③	④	⑤

Questionnaires about consumer value and life-style about shopping behaviors are consisted in 2016 KISDI personal media panel investigation.

문20) [2003년 12월 31일 이전 출생자만 (만 13세 이상) 응답 / 2004년 1월 1일 이후 출생자부터 문21)로 이동]
다음은 상품이나 제품을 **구매하는 행태**에 대한 질문입니다. 각각의 사항에 대하여 응답해 주십시오.

문 번호	항 목	동의 정도				
		전혀 그렇지 않다 ①	그렇지 않다 ②	보통이다 ③	그렇다 ④	매우 그렇다 ⑤
문20-1)	최신 상품인지가 중요하다	①	②	③	④	⑤
문20-2)	남들이 잘 쓰지 않는 상품이 더 매력적이다	①	②	③	④	⑤
문20-3)	유행에 민감하다	①	②	③	④	⑤
문20-4)	주위 사람이나 인터넷에서의 평판이 구입에 있어서 중요하다	①	②	③	④	⑤
문20-5)	최소한의 기능으로 만족하며 가격이 저렴한 것이 중요하다	①	②	③	④	⑤
문20-6)	디자인이나 색상이 중요하다	①	②	③	④	⑤
문20-7)	유명 브랜드 여부가 중요하다	①	②	③	④	⑤
문20-8)	상품을 직접 눈으로 확인하고 가격을 흥정하기 위하여 관련 매장을 돌아다니 며 구매한다	①	②	③	④	⑤
문20-9)	내가 사용할 상품은 내가 직접 구입한다	①	②	③	④	⑤
문20-10)	상품에 대해서 알아보고 저렴하게 구입하기 위해서 인터넷을 활용한다	①	②	③	④	⑤

Questionnaires of monthly telecommunication expenditure from 2015 to 2019 are consisted in KISDI personal media panel investigation.

II. 방송통신 서비스 가입 및 지출 현황

문2) [휴대폰 사용자만 응답 / 휴대폰 비사용자는 문4)로 이동]
귀하께서 사용하시는 휴대폰의 월평균 통신요금(기기 할부금 제외)과 기기 할부금은 얼마입니까?

월평균 휴대폰 통신요금	기기 할부금
□□□ 만 □ 천 원 ※ 지난 3개월 동안 월평균 요금을 기준으로 콘텐츠·정보이용료, 소액결제 내역, 기기 할부금은 제외하고 응답해 주십시오. ※ 사용하는 휴대폰이 여러 대인 경우 휴대폰 요금을 합산해 주십시오.	□□□ 만 □ 천 원 ※ 매월 납부하는 기기 할부금을 적어 주십시오. ※ 할부금이 없을 경우에는 0으로 기록해 주십시오. ※ 사용하는 휴대폰이 여러 대인 경우 기기 할부금을 합산해 주십시오.

Questionnaires of telecommunication generation choice from 2015 to 2019 are consisted in KISDI personal media panel investigation.

I. 휴대폰 및 스마트 기기 이용현황

문1) 귀하께서 현재 사용 중인 휴대폰(PDA폰 제외)이 있습니까? 있다면 몇 대 사용하고 있는지도 응답해 주십시오.

① 있다 총 ( )대 ② 없다 ➔ 문2)로 이동

문1-1) 일반휴대폰	대 ➔ 문1-4)~문1-15) 응답	문1-2) 스마트폰	대 ➔ 문1-16)~문1-23) 응답	문1-3) 키즈폰	대 ➔ 문2)로 이동

A. 일반 휴대폰

문1-4) ~ 문1-15) 다음의 표에 현재 사용 중인 일반 휴대폰 중, 많이 이용하는 휴대폰 순서대로 최대 2대까지 기록해 주십시오.

※ 스마트폰을 제외한 일반 휴대폰에 대해서만 기록해 주시고, 사용 중인 스마트폰에 대해서는 문1-16) ~ 문1-23) 표에 기록해 주십시오.

문 번호	항 목	보기 설명	일반 휴대폰 1	일반 휴대폰 2
문1-4)	일반 휴대폰 구분	① 2G 일반 휴대폰 - 영상통화 불가능 ② 3G 일반 휴대폰 - 영상통화 가능	① ②	① ②
문1-5)	사진 촬영	① 사진 촬영 기능 있음 ② 사진 촬영 기능 없음	① ②	① ②
문1-6)	동영상 촬영	① 동영상 촬영 기능 있음 ② 동영상 촬영 기능 없음	① ②	① ②
문1-7)	MP3	① MP3 등 음원 파일 재생 기능 있음 ② MP3 등 음원 파일 재생 기능 없음	① ②	① ②
문1-8)	지상파 DMB	① 지상파 DMB 수신 기능 있음 ② 지상파 DMB 수신 기능 없음	① ②	① ②
문1-9)	와이파이(Wi-Fi) - 별도의 단말기 연결 없이, 무선 공유기(AP)가 설치된 특정 지역에서 무선 인터넷을 사용 가능	① 와이파이(Wi-Fi) 지원 가능 ② 와이파이(Wi-Fi) 지원 불가능	① ②	① ②
문1-10)	가입한 이동통신사	① SKT ② KT(구 KTF) ③ LG U⁺(구 LGT) ④ 알뜰폰 서비스(MVNO) 예) CJ헬로모바일, 홈플러스폰 등	① ② ③ ④	① ② ③ ④
문1-11)	가입한 요금제 - **23 페이지 통신사별 휴대폰 요금제** 표를 참고하여 응답해 주십시오. - 고지서상의 요금제를 확인하여 응답해 주십시오.		________	________
문1-12)	복지할인 여부 - 국가유공자, 장애인, 기초수급생활자를 대상으로 휴대폰 요금을 할인해 주는 제도	① 예 ② 아니오	① ②	① ②
문1-13)	최초 사용 연도 - 문1-4)에서 표기한 휴대폰을 사용하기 시작한 연도를 적어 주십시오.		년	년
문1-14)	예상 교체시기 - 문1-4)에서 표기한 휴대폰의 예상 교체시기에 대해 응답해 주십시오.	① 6개월 이내 ② 6개월 ~ 1년 이내 ③ 1년 ~ 2년 이내 ④ 2년 ~ 3년 이내 ⑤ 3년 이후	① ② ③ ④ ⑤	① ② ③ ④ ⑤
문1-15)	제조사	① 삼성 ② LG ③ 착한텔레콤(구 팬택 SKY) ④ 기타 국내 제조사( ) ⑤ 기타 해외 제조사( )	① ② ③ ④( ) ⑤( )	① ② ③ ④( ) ⑤( )

B. 스마트폰

문1-16) ~ 문1-23) 다음의 표에 현재 사용 중인 **스마트폰** 중, 많이 이용하는 스마트폰 순서대로 최대 2대까지 기록해 주십시오.

문 번호	항 목	보기 설명	스마트폰 1	스마트폰 2
문1-16)	스마트폰 구분	① 3G 스마트폰 ② LTE 스마트폰 ③ LTE-A 스마트폰 ④ 5G 스마트폰	① ② ③ ④	① ② ③ ④
문1-17)	지상파 DMB	① 지상파 DMB 수신 기능 있음 ② 지상파 DMB 수신 기능 없음	① ②	① ②
문1-18)	가입한 이동통신사	① SKT ② KT(구 KTF) ③ LG U⁺(구 LGT) ④ 알뜰폰 서비스(MVNO) 예) CJ헬로모바일, 홈플러스폰 등	① ② ③ ④	① ② ③ ④
문1-19)	가입한 요금제 - **23 페이지 통신사별 휴대폰 요금제** 표를 참고하여 응답해 주십시오. - 고지서상의 요금제를 확인하여 응답해 주십시오.		________	________
문1-19-1)	음성 무제한 서비스 가입 여부 - 가입한 요금제 기준으로 응답해 주십시오.	① 가입 ② 미가입	① ②	① ②
문1-19-2)	데이터 무제한 서비스 가입 여부 - 가입한 요금제 기준으로 응답해 주십시오.	① 가입 ② 미가입	① ②	① ②
문1-20)	복지할인 여부 - 국가유공자, 장애인, 기초수급생활자를 대상으로 휴대폰 요금을 할인해 주는 제도	① 예 ② 아니오	① ②	① ②
문1-21)	최초 사용 연도 - 문1-16)에서 표기한 휴대폰을 사용하기 시작한 연도를 적어 주십시오.		□□□□ 년	□□□□ 년
문1-22)	예상 교체시기 - 문1-16)에서 표기한 휴대폰의 예상 교체시기에 대해 응답해 주십시오.	① 6개월 이내 ② 6개월 ~ 1년 이내 ③ 1년 ~ 2년 이내 ④ 2년 ~ 3년 이내 ⑤ 3년 이후	① ② ③ ④ ⑤	① ② ③ ④ ⑤
문1-23)	제조사	① 삼성 ⑤ 샤오미 ② LG ⑥ 화웨이 ③ 애플 ⑦ 기타 제조사( ) ④ 착한텔레콤(구 팬택 SKY)	① ⑤ ② ⑥ ③ ⑦( ) ④	① ⑤ ② ⑥ ③ ⑦( ) ④

C. 태블릿 PC(스마트패드, 컨버터블 PC, 전자책(e-book) 리더기 포함)

문2) 귀하께서 현재 사용 중인 태블릿 PC(스마트패드, 컨버터블 PC, 전자책(e-book) 리더기, 키즈패드, 어학용패드 포함)가 있습니까? 있다면 몇 대 사용하고 있는지도 응답해 주십시오.

※ 스마트패드 : 아이패드, 갤럭시탭 등과 같이 키보드 없이 터치스크린을 이용하여 조작하는 휴대 가능한 컴퓨터
※ 컨버터블 PC : 휴대용 터치스크린 외에 키보드 등의 추가적 입력장치 연결이 가능하고, 기존 컴퓨터용 프로그램 사용 가능
※ 전자책(e-book) 리더기 : 전자책(e-book)을 볼 수 있게 해주는 기기
※ 키즈패드 : 유아 교육용으로 개발된 태블릿 PC로 기존 태블릿 PC와는 달리 다수의 유아전용 콘텐츠가 포함되어 있음
※ 어학용패드 : 어학 교육용으로 개발된 태블릿 PC로 기존 태블릿 PC와는 달리 다수의 어학전용 콘텐츠가 포함되어 있음

① 있다 (총 □□ 대) ② 없다 ➔ **문3)으로 이동**

문2-1) ~ 문2-4) 다음의 표에 현재 사용 중인 태블릿 PC 중, 많이 이용하는 태블릿 PC 순서대로 최대 2대까지 기록해 주십시오.

문 번호	항 목	보기 설명	태블릿 PC 1	태블릿 PC 2
문2-1)	태블릿 PC 사용유형 (복수응답)	① 혼자 사용 ② 가구 내 가구원과 함께 사용 ③ 회사/학교 소유 함께 사용 ④ 타 가구의 가구원(지인)과 함께 사용 ⑤ 기타( )	① ② ③ ④ ⑤( )	① ② ③ ④ ⑤( )
문2-2)	태블릿 PC 종류	① 스마트패드 ③ 전자책(e-book) 리더기 ② 컨버터블 PC ④ 키즈패드 (슬레이트PC, 탭북 등) ⑤ 어학용패드	① ③ ② ④ ⑤	① ③ ② ④ ⑤
문2-3)	최초 사용 연도 - 우측에 표기한 태블릿 PC를 사용하기 시작한 연도를 적어 주십시오.		□□□□ 년	□□□□ 년

The frequency of daily smart phone usage and SNS usage are calculated according to the KISDI personal media diary. The investigation forms of media diary are generally identical from 2015 to 2019. It varies few parts every year in order to reflect the developments of technologies and markets.

일일 매체 이용 체크리스트

문) 귀하께서는 ○월 ○일(자정 기준) 하루 동안, 아래와 같은 매체를 이용한 적이 있습니까?
각 일자별로 이용하신 매체 코드를 모두 확인해 주십시오.
만약 하나의 매체를 이용하시면서 다른 매체를 동시에 이용하셨다면, 해당 매체(두 가지)끼리 선으로 연결해 주시기 바랍니다.

※ 이용한 매체의 경우 미디어 다이어리 작성 가이드북의 해당 페이지로 이동하여 매체코드와 행위코드, 연결코드를 찾아 주십시오.
찾은 3개의 코드를 미디어 다이어리 응답지의 해당 이용 시간대에 기록해 주십시오.

매체		작성 가이드북	첫째날		둘째날		셋째날	
			매체 코드	동시 이용 연결	매체 코드	동시 이용 연결	매체 코드	동시 이용 연결
종이 매체	신문/책/잡지	➲ 3페이지	1		1		1	
	그림/사진/편지/쪽지	➲ 3페이지	2		2		2	
TV	가정용 TV	➲ 4페이지	3		3		3	
	전광판 TV(혹은 옥외 TV, G-Bus TV, 지하철 TV 등)	➲ 5페이지	4		4		4	
	휴대용 TV	➲ 5페이지	5		5		5	
	차량용 TV(승용차용 TV 수상기)	➲ 5페이지	6		6		6	
컴퓨터	데스크톱 PC	➲ 6페이지	7		7		7	
	노트북 PC	➲ 7페이지	8		8		8	
	태블릿 PC(스마트패드, 컨버터블 PC, 전자책(e-book) 리더기, 키즈패드, 어학용패드 포함)	➲ 8페이지	10		10		10	
	PDA	➲ 9페이지	11		11		11	
	내비게이션	➲ 10페이지	12		12		12	
	아웃도어 미디어키오스크(미디어폴 포함)	➲ 11페이지	13		13		13	
전화기	공중 전화기	➲ 12페이지	14		14		14	
	일반 전화기(인터넷 전화기 제외)	➲ 12페이지	15		15		15	
	인터넷 전화기	➲ 13페이지	16		16		16	
	일반 휴대폰	➲ 14페이지	17		17		17	
	PDA폰	➲ 15페이지	18		18		18	
	스마트폰(스마트폰 제어 스마트와치/밴드 등 웨어러블기기, 키즈폰 포함)	➲ 16페이지	19		19		19	
촬영 기기	디지털 카메라	➲ 17페이지	20		20		20	
	비디오 녹화기기(캠코더)	➲ 18페이지	21		21		21	
	홈CCTV/블랙박스	➲ 18페이지	22		22		22	
	드론 기기	➲ 19페이지	23		23		23	
오디오 기기	일반 라디오(라디오 전용 수신기)	➲ 20페이지	24		24		24	
	가정용 오디오 기기(포터블 오디오, 홈씨어터, 블루투스 스피커/도킹 오디오 등)	➲ 20페이지	25		25		25	
	카오디오	➲ 21페이지	26		26		26	
	오디오 레코더	➲ 21페이지	27		27		27	
	휴대용 오디오 기기(MP3 플레이어 등)	➲ 22페이지	28		28		28	
비디오 재생/녹화 기기	VCR	➲ 23페이지	29		29		29	
	DVD플레이어(블루레이, HD-DVD 플레이어 포함)	➲ 23페이지	30		30		30	
	PVR/DVR/DivX 플레이어	➲ 24페이지	31		31		31	
	휴대용 비디오 재생기기(PMP 등)	➲ 25페이지	32		32		32	
게임기	휴대용 게임기	➲ 26페이지	33		33		33	
	가정용 게임기	➲ 27페이지	34		34		34	
VR/AR 기기	VR/AR(가상/증강현실) 기기	➲ 28페이지	35		35		35	
공간 미디어	영화관	➲ 29페이지	36		36		36	
	노래방	➲ 29페이지	37		37		37	
	멀티미디어방(게임방, DVD방, 게임카페, VR/AR(가상/증강현실) 체험관 등)	➲ 29페이지	38		38		38	
	공연장	➲ 29페이지	39		39		39	
	갤러리	➲ 29페이지	40		40		40	
	박물관	➲ 29페이지	41		41		41	
	스포츠경기장	➲ 29페이지	42		42		42	

1 미디어 다이어리 작성 – 첫째날

보 기 카 드

★ 상세 설명이 필요한 경우 별도 배부한 작성 가이드북 참고 ★

매체 코드 – '어떤 매체로'

종이 매체	신문/책/잡지	1
	그림/사진/편지(쪽지)	2
TV	가정용 TV	3
	전광판 TV(혹은 옥외 TV, G-Bus TV, 지하철 TV 등)	4
	휴대용 TV	5
	차량용 TV(승용차용 TV 수상기)	6
컴퓨터	데스크톱 PC	7
	노트북 PC	8
	태블릿 PC(스마트패드, 컨버터블 PC, 전자책(e-book) 리더기, 키즈패드, 어학용패드 포함)	10
	PDA	11
	내비게이션	12
	아웃도어 미디어키오스크(미디어폴 포함)	13
전화기	공중 전화기	14
	일반 전화기(인터넷 전화기 제외)	15
	인터넷 전화기	16
	일반 휴대폰	17
	PDA폰	18
	스마트폰(스마트폰 제어 스마트와치/밴드 등 웨어러블기기, 키즈폰 포함)	19
촬영 기기	디지털 카메라	20
	비디오 녹화기기(캠코더)	21
	홈CCTV/블랙박스	22
	드론 기기	23

오디오 기기	일반 라디오(라디오 전용 수신기)	24
	가정용 오디오 기기(포터블 오디오, 홈씨어터, 블루투스 스피커/도킹 오디오 등)	25
	카오디오	26
	오디오 레코더	27
	휴대용 오디오 기기(MP3 플레이어 등)	28
비디오 재생/녹화 기기	VCR	29
	DVD플레이어(블루레이, HD-DVD 플레이어 포함)	30
	PVR/DVR/DivX 플레이어	31
	휴대용 비디오 재생기기(PMP 등)	32
게임기	휴대용 게임기	33
	가정용 게임기	34
VR/AR	VR/AR(가상/증강현실) 기기	35
공간 미디어	영화관	36
	노래방	37
	멀티미디어방(게임방, DVD방, 게임카페, VR/AR(가상/증강현실) 체험관 등)	38
	공연장	39
	갤러리	40
	박물관	41
	스포츠경기장	42

행위 코드 - '무엇을'

TV/라디오/방송 프로그램	지상파 TV방송 프로그램 시청 (MBC, KBS, SBS, EBS 및 그 계열사)	실시간 시청(재방송 포함)	1
		VOD/다시보기	2
	비지상파 TV방송 프로그램 시청 (YTN, OCN, tvN, Mnet, 홈쇼핑 등)	실시간 시청(재방송 포함)	3
		VOD/다시보기	4
	종합편성채널 TV방송 프로그램 시청 (JTBC, MBN, TV조선, 채널A 등)	실시간 시청(재방송 포함)	5
		VOD/다시보기	6
	TV 데이터방송 프로그램 시청 (날씨(날씨앤조이, 웨더채널), 교통(SBS교통정보), 증권 등)		7
	라디오 방송 혹은 음악채널 프로그램 청취	실시간 청취	8
		AOD/다시듣기	9
	옥외/지하철 방송프로그램 시청		10
영화/동영상/UCC/음악/음원/사진	영화/동영상(영화제작사, 프로덕션 등 제작) 시청 (뮤직비디오, 학습 동영상, 애니메이션 등 포함)		11
	UCC(개인창작 콘텐츠) 동영상 시청	실시간 시청(개인 생방송)	12
		다시보기/다시듣기	13
	홈CCTV/블랙박스 확인		14
	음악 · 음원 청취(MP3 등 음원 재생, 음악 듣기 서비스, 오디오북 등)		15
	사진 보기(그림, 그래픽 이미지 등 포함)		16
신문/책/잡지	신문 기사 읽기(종이 신문, 인터넷 · 전자 신문, 애플리케이션 등)		17
	책(전자책(e-book) 포함) 읽기(소설, 시, 교과서, 만화, 웹툰 등) ※ 그림 · 사진 화보 포함		18
	잡지(웹진(webzine) 포함) 읽기		19

통화/문자/이메일/채팅	통화하기(음성 통화/영상 통화)	20
	문자메시지 읽기/쓰기/보내기	21
	이메일(e-mail) 읽기/쓰기/보내기	22
	채팅/메신저 하기(쪽지 보내기, 대화하기 등, 예: 카카오톡 등)	23
온라인검색/소셜네트워크/상거래	정보콘텐츠 검색 및 이용(위치/교통/생활/상품/지식 등)	24
	소셜네트워크서비스(SNS) 이용(블로그, 미니홈피 등, 예: 트위터, 카카오스토리, 페이스북 등)	25
	온라인 상거래(온라인 쇼핑, 온라인 뱅킹, 간편결제, 자산관리, 온라인 펀딩 등의 금융서비스, 예약 서비스, 택시 호출 등)	26
게임	게임 하기(온라인 · 오프라인(자체내장) 게임)	27
문서/그래픽 작업	문서 작업(워드, 엑셀, 파워포인트, 한글 등)	28
	그래픽 작업 등(사진/동영상 편집, 그래픽 툴을 활용한 시청각 자료 작업, 프로그래밍 작업 등)	29
(종이로 된) 그림/사진/편지/쪽지	(종이) 그림 감상하기	30
	(종이) 사진(앨범) 보기	31
	(종이) 편지/쪽지 수신/발신 또는 읽기/쓰기	32
공간 미디어 활동	영화관 이용(영화 관람)	33
	노래방 이용(노래 부르기)	34
	멀티미디어방 이용(DVD 시청, 게임하기, VR/AR 체험 등)	35
	공연장 이용(연극, 뮤지컬 등의 공연 관람)	36
	갤러리 이용(전시회 관람)	37
	박물관 이용(박물관 관람)	38
	스포츠경기장 이용(스포츠경기 관람)	39
사물인터넷	가전제품, 전자기기 등 원격제어	40

연결 코드 - '어떻게 (어떤 경로를 통해)'

방송 서비스를 통해	케이블 TV 방송서비스를 통해(셋톱박스 연결 포함)		1
	IPTV 방송서비스를 통해(셋톱박스 연결 포함)		2
	위성방송 서비스를 통해(셋톱박스 연결 포함)		3
	지상파 DMB서비스를 통해		4
	지상파 방송 직접 수신을 통해(지상파 안테나/디지털컨버터)		5
전화 서비스를 통해	유선 전화 서비스를 통해		6
	이동통신 전화 서비스를 통해 (기기 고유의 통화 · 문자 서비스 이용)		7
인터넷 연결을 통해	유선인터넷(광랜, 기가인터넷, 케이블모뎀, FTTH, xDSL 등)을 통해		8
	이동통신 무선인터넷(2G, 3G, LTE 무선인터넷 서비스 이용)을 통해		9
	무선인터넷을 통해	와이파이(Wi-Fi)존, 유무선공유기 이용	10
		에그(Egg), 브릿지(Bridge), 단비 이용	11
	와이브로 서비스(휴대인터넷)를 통해		12

기타	OTT 단말기를 통해 (예: 구글 크롬캐스트, 티빙 스틱, 딜라이브 플러스 등)	13
	CD/DVD 등의 디스크를 기기(PC, DVD플레이어, 게임기 등)에서 재생하여 (비디오 녹화기기(캠코더)의 경우 테이프를 기기에서 재생하여)	14
	자체 기능 및 소프트웨어, 이미 저장된 파일을 이용하거나 외장하드, USB, NAS 스토리지 등 저장매체를 통해	15
	다른 미디어 기기(VCR, DVD플레이어, 컴퓨터, 게임기, 휴대용 오디오 등)에 직접 연결된 상태로 영상/음성 신호 등을 전달 받아	16
	정기구독하고 있거나 우편/택배/배달을 통해	17
	본인이 직접 서점/가판대 등에서 구입하거나 도서관/대여점 등에서 대여하여 (편지/쪽지의 경우 직접 받거나 건네주어)	18
	그 외 (출처 불분명, 친구/가족의 것, 회사에 있는 것 등)	19
	공간 미디어를 통해	20

첫째날 : 2019 년 월 일 요일 : □1)월 □2)화 □3)수 □4)목 □5)금 □6)토 □7)일

▶ **오늘 하루 귀하에게 특별한 일이 있었나요?** □1) 시험기간 □2) 가족 · 친척 경조사 □3) 휴가/여행 □4) 출장 □5) 특별한 일 없음

장소 응답 보기	집/직장/학교 등	교통수단	여가시설 · 서비스 · 기타	
	1) 본인 주거 공간 : 자택, 기숙사, 별장 등 2) 타인 주거 공간 : 친구 · 친척 집, 이웃 집 등 3) 직장 : 회사, 사무실 등 4) 교육시설(학생의 경우만 해당) : 학교, 학원, 도서관 등 5) 본인 주거/사업 겸용 공간 : 오피스텔 등	6) 대중교통수단 내 : 버스, 지하철, 택시, 비행기 등 7) 개인교통수단 : 자가용, 오토바이 등 8) 개인이동/대중교통 환승대기 : 인도, 지하철 플랫폼, 택시 승강장, 공항 등	9) 오락시설 : PC방, DVD방, 게임방, 노래방 등 10) 요식업시설 : 카페, 식당, 술집 등 11) 체육시설 : 헬스장, 운동장 등 12) 문화시설 : 공연장, 극장, 미술관, 박물관 등 13) 상거래 시설 : 백화점, 시장, 미용실, 병원, 은행 등	14) 종교시설 : 교회, 절, 성당 등 15) 관광휴양지 : 휴양지, 유원지, 산, 바다 등 16) 숙박시설: 호텔, 모텔, 게스트 하우스 등 17) 기타

시간대				1. 수면	2. 장소	주로 이용한 미디어			동시에 이용한 미디어		
						매체 코드	행위 코드	연결 코드	매체 코드	행위 코드	연결 코드
새벽	0시	00:00 - 00:15	1								
		00:15 - 00:30	2								
		00:30 - 00:45	3								
		00:45 - 01:00	4								
	1시	01:00 - 01:15	5								
		01:15 - 01:30	6								
		01:30 - 01:45	7								
		01:45 - 02:00	8								
	2시	02:00 - 02:15	9								
		02:15 - 02:30	10								
		02:30 - 02:45	11								
		02:45 - 03:00	12								
	3시	03:00 - 03:15	13								
		03:15 - 03:30	14								
		03:30 - 03:45	15								
		03:45 - 04:00	16								
	4시	04:00 - 04:15	17								
		04:15 - 04:30	18								
		04:30 - 04:45	19								
		04:45 - 05:00	20								
	5시	05:00 - 05:15	21								
		05:15 - 05:30	22								
		05:30 - 05:45	23								
		05:45 - 06:00	24								
	6시	06:00 - 06:15	25								
		06:15 - 06:30	26								
		06:30 - 06:45	27								
		06:45 - 07:00	28								
오전	7시	07:00 - 07:15	29								
		07:15 - 07:30	30								
		07:30 - 07:45	31								
		07:45 - 08:00	32								
	8시	08:00 - 08:15	33								
		08:15 - 08:30	34								
		08:30 - 08:45	35								
		08:45 - 09:00	36								
	9시	09:00 - 09:15	37								
		09:15 - 09:30	38								
		09:30 - 09:45	39								
		09:45 - 10:00	40								
	10시	10:00 - 10:15	41								
		10:15 - 10:30	42								
		10:30 - 10:45	43								
		10:45 - 11:00	44								
	11시	11:00 - 11:15	45								
		11:15 - 11:30	46								
		11:30 - 11:45	47								
		11:45 - 12:00	48								
오후	12시	12:00 - 12:15	49								
		12:15 - 12:30	50								
		12:30 - 12:45	51								
		12:45 - 13:00	52								
	1시	13:00 - 13:15	53								
		13:15 - 13:30	54								
		13:30 - 13:45	55								
		13:45 - 14:00	56								
	2시	14:00 - 14:15	57								
		14:15 - 14:30	58								
		14:30 - 14:45	59								
		14:45 - 15:00	60								
	3시	15:00 - 15:15	61								
		15:15 - 15:30	62								
		15:30 - 15:45	63								
		15:45 - 16:00	64								
	4시	16:00 - 16:15	65								
		16:15 - 16:30	66								
		16:30 - 16:45	67								
		16:45 - 17:00	68								
	5시	17:00 - 17:15	69								
		17:15 - 17:30	70								
		17:30 - 17:45	71								
		17:45 - 18:00	72								
	6시	18:00 - 18:15	73								
		18:15 - 18:30	74								
		18:30 - 18:45	75								
		18:45 - 19:00	76								
저녁	7시	19:00 - 19:15	77								
		19:15 - 19:30	78								
		19:30 - 19:45	79								
		19:45 - 20:00	80								
	8시	20:00 - 20:15	81								
		20:15 - 20:30	82								
		20:30 - 20:45	83								
		20:45 - 21:00	84								
	9시	21:00 - 21:15	85								
		21:15 - 21:30	86								
		21:30 - 21:45	87								
		21:45 - 22:00	88								
	10시	22:00 - 22:15	89								
		22:15 - 22:30	90								
		22:30 - 22:45	91								
		22:45 - 23:00	92								
	11시	23:00 - 23:15	93								
		23:15 - 23:30	94								
		23:30 - 23:45	95								
		23:45 - 24:00	96								

Abstract (Korean)

사회적 상호작용은 개인의 태도 및 의견, 개인의 선호, 그리고 개인의 행동에 중요한 영향을 미친다. 따라서 사회적 상호작용은 사회과학 분야에 매우 중요한 연구 대상이다. 사회적 상호작용은 크게 구전 효과와 관찰된 학습 두가지로 나눌 수 있다. 기존 경제학 분야의 선택 모형에서는 주로 관찰된 학습을 사회적 상호작용으로 간주하였다. 구전 효과는 태도나 행위의 경향성 분석에 많이 연구되어 왔지만 선택 모형에 고려하는 것은 어려움이 많았다. 그리고, 많은 선택모형에서 개인의 의사결정 과정을 더 잘 이해하고 모형의 예측력을 향상시키기 위해 개인의 잠재적 심리학 변수를 추가하고자 하였다. 사회적 상호작용이 개인의 심리학적 특성에 많은 영향을 미친다는 것은 이미 많은 연구를 통해서 입증되어 왔다. 하지만 사회적 상호작용이 개인의 잠재적 심리학 변수에 미치는 영향을 다룬 연구는 매우 드물었다. 또한, 개인의 행동 사이에 존재한 상관관계를 선택모형에 동시에 다루지 못하면 편향된 추정결과를 얻을 수 있다. 혼합형 종속변수를 동시에 추정하는 모형은 꾸준히 개발되어 왔지만 여기에 사회적 상호작용을 고려한 연구는 찾아보기 힘들다. 본 연구에서는 두가지 사회적 상호작용을 모두고려하는 다중 선택 모형을 제안하였다. 제안한 선택 모형은 사회적 상호작용을 고려할 수 있을 뿐만 아니라 내생적 다중 선택을 다룰 수 있으며 혼합적 종속변수까지 포함할 수 있다. 본 연구에서 먼저 시뮬레이션 분석을 통해서

모델의 유용성을 입증한 다음에 실증분석으로 기존 모델에서 확인할 수 없었던 사회적 상호작용이 개인의 행위에 미치는 영향을 분석하였다. 사회적 상호작용을 고려하지 않으면 추정된 결과가 편향될 수 있으며 변수의 계수를 과대 추정할 가능성이 있다는 것을 실증연구를 통해서 증명했다.

주요어 : 사회적 상호작용, 다중 선택, 혼합형 종속변수, 합성 한계 우도함수
학 번 : 2015-30854