Ph. D. Dissertation in Engineering

Using Artificial Neural Network for Consumer Choice Analysis

- Case of next-generation transportation market -

인공지능망을 통한 소비자 선택 분석
: 차세대 수송모드 서비스 시장

August 2020

Graduate School of Seoul National University

Technology Management, Economics, and Policy Program

Stephen Youngjun Park

Using Artificial Neural Network for Consumer Choice Analysis

- Case of next-generation transportation market -

지도교수 이종수

이 논문을 공학박사학위 논문으로 제출함

2020 년 8 월

서울대학교 대학원

협동과정 기술경영경제정책 전공

박영준

박영준의 공학박사학위 논문을 인준함

2020 년 7 월

위 원 장	이정동	(인)
부위원장	이종수	(인)
위 원	최재영	(인)
위 원	정기철	(인)
위 원	신정우	(인)

Abstract

Using Artificial Neural Network for Consumer Choice Analysis

(Case of next-generation transportation market)

Stephen Youngjun Park
Technology Management, Economics, and Policy Program
The Graduate School
Seoul National University

The present dissertation aims to provide insights into the application of different artificial neural network models in the analysis of consumer choice regarding next-generation transportation services (NGT). It categorizes consumers' decisions regarding the adoption of new services according to Dewey's buyer decision process and then analyzes these decisions using a variety of different methods. In particular, various artificial neural network (ANN) models are applied to predict consumers' intentions. Also, the dissertation proposes an attention-based ANN model that identifies the key features that affect consumers' choices. Consumers' preferences for different types of NGT services are analyzed using a hierarchical Bayesian model. The analyzed consumer preferences are utilized to forecast demand for NGT services, evaluate government policies within the transportation market, and provide evidence regarding the social conflicts among traditional and new transportation services. The dissertation uses the Multiple Discrete-Continuous Extreme Value (MDCEV) model to analyze consumers' decisions regarding the use of different transportation modes. It also utilizes this MDCEV model analysis to estimate the effect of NGT services on consumers' travel mode selection behavior and the environmental effects of the transportation sector. Finally, the findings of the dissertation's analyses are combined to generate marketing and policy insights that will promote NGT services in Korea.

Keywords: Artificial neural network; discrete choice model; consumer decision; demand forecasting, next-gen transportation service

Student Number: 2015-31038

Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Research Background

The improvement of computing power has expanded the application of artificial neural networks (ANN) to various fields. The driving force of artificial neural networks is that they can process almost any type of input data due to their universal function approximator, which allows the use of ANN in fields such as translation, text recognition, image recognition, and different kinds of predicting tasks. As their use has expanded, many different forms of ANN have been developed in recent years, further widening the models' application area. The main advantage of using ANN is that they are parsimonious, produce good classification results, and easily incorporate the complex relationships present in the data. However, although ANN complete the aforementioned tasks exceptionally well, they have rarely been used in the field of consumer studies. This is due to ANN's "black-box" characteristics. The learning process of ANN depends solely on the data; given input and which output to predict. The relationships between the inputs and outputs are represented by networks consisting of multiple hidden layers, nodes, and non-linear functions. Network construction is adjusted or learned through training methods such as backpropagation to make the network produce the most likely outcome based on a given input. Although researchers are able to configure the network construction, interpreting the relationship between the data by examining the network is difficult. Thus, the inside of a general ANN is considered a "black-box."

Since the field of consumer studies mainly focuses on analyzing consumer behavior and understanding the effects of certain variables on consumer preferences and behaviors, the use of ANN has been very limited. As mentioned above, the “black-box” characteristic of ANN limits the inferences researchers can make. In contrast to studies that use ANN and focus on classifying and predicting outcomes, studies regarding consumer choice have developed various econometric models to better understand the consumer behaviors and preferences that guide decision making. Such models are denoted as discrete choice models and they have been widely used in the consumer studies field to analyze different types of consumer choices. Choice models range from basic forms consisting of ordinary least squares (OLS) and logistic regression to highly advanced forms of hierarchical Bayesian models and the Multiple Discrete-Continuous Extreme Value Model. All of these models have been applied in consumer studies to develop a better understanding of consumer heuristics and consumer behavior. Put simply, the choice modeling technique’s solution methodology (closed form versus heuristic) and interpretability are superior to those of ANN. (Kumar, Rao, and Soni, 1995)

The present study attempts to incorporate ANN into the field of consumer studies by applying machine learning techniques to improve the analysis procedures of current choice analysis. Some researchers have attempted to directly incorporate ANN into consumer studies, but many of these efforts have been limited to comparing the results from the ANN to those from traditional choice models. Thus, the present study focuses on using the ANN where they are most suitable—in predicting and classifying—and incorporates the ANN results into the advanced choice models. Rather than combining the two methods into one model, the study uses both ANN and choice models to analyze different types of choice to improve the understanding of consumer decision making. Intuitively speaking, the current study uses ANN for predicting and classifying tasks, and uses choice models to make interpretations and inferences. In particular, the study uses ANN to predict consumers' intentions to adopt new products, and to identify the key variables that affect consumer choices. The proposed research framework is applied to analyze next-generation transportation modes in Korea.

Ever since the introduction of the next generation transportation services, their influence on the consumer's travel behavior, ownership of vehicle, and the usage of traditional transportation service has been growing exponentially. The next generation of mobility services can be largely divided into ride-sharing services and car sharing services which are reshaping the landscape of the mobility market. Ride sharing refers to P2P services similar to taxis but differ in that the driver is not required to have the operation license, and the vehicles are privately owned.

There are two main rider-sharing services in Korea as of January 2020. The first service is carpool service where as defined, individuals with private vehicles registered as a driver in a service can provide carpooling service. The second service is a special case of carpool preferably defined as centralized taxi service (CTS) in this study. CTS consists of commercial vehicles that are driven by non-professional drivers. The CTS designate drivers and vehicle according to the consumers' locations and their destinations. Due to their similarities with the traditional taxi service, both ride-sharing service is heavily regulated in the Korean market. The carpool services are allowed to operate only around the commuting time, limiting the total daily operation hours to four hours. However, until recently, CTS have been abusing the definition of carpool to bypass this regulation. As carpool service is defined as service provided by privately owned vehicles, CTS, which operates with commercially registered vehicles could operate 24 hours a day. Although successful in bypassing the law, CTS has been the subject of continued conflict and dispute with the taxi industry. As a result, CTS has been banned from the market as of April 2020, but it’s inevitable that similar services will enter the market soon.

Car sharing services refer to B2P services that allows customers to share the vehicles owned by the service provider. It is distinguishable from the traditional rental services in that customers can use the vehicles on minute-by-minute basis for short periods of times, only pay for the mileage covered, and are free to return the vehicle to any service designated areas in the city, providing convenient advantages over the former service. Car sharing services are largely preferred among the younger generation, who mostly use public transportation for everyday travel but require passenger cars for special purposes, such as out of city travels. This trend was quickly picked up by startups, rental companies and later by car manufacturers who have all entered the market. The following table provides comparison of mobility services currently in service in Korea.

Table 1. Type of mobility services available

Category	Car sharing	Rent car	Ride sharing	Carpool
Customer	Membership	Any customer	Membership	Membership
Available hours	24 hours	Appointed hours	24 hours	24 hours
Pickup location	Multiple locations	Branch offices	User’s current location	Agreed location with the driver
Usage cost standard	Charged every 10 minutes (includes insurance and fuel)	Half-day or 1 day (excludes insurance and fuel)	Charged by distance (includes insurance and fuel)	Charged by distance (includes insurance and fuel)
Payment method	Post-payment	Prepayment	Prepayment	Prepayment
Contract method	Contract when joining the service	Every time	Contract when joining the service	Contract when joining the service

Source: Choi and Lee, 2018

There are multiple types of car sharing services, which first started with stationary car sharing, followed by new services such as free-floating car sharing and peer-to-peer car sharing (Deloitte, 2017). Stationary car sharing is characterized by fixed return stations, where the customer is required to return the vehicle to a designated service area. Although most cities provide large number of stations, the customers are still limited in their travel behavior. In order to free customers from this restriction, free-floating car sharing services are on the rise. Although the car sharing service has just emerged around than five years ago, and the market has been booming ever since. Free-floating services allows customers to pick up the car and return the vehicle to any parking locations, as long as it is not in an illegal parking area. This allows greater level of flexibility, allowing the customers to plan for one-way trips as an alternative to a taxi. Although this type of service has higher prices, the global market share is larger than stationary services.

However, free-floating service often times faces difficulties when cooperating with local authorities to avoid parking limitations. In 2014, car2go stopped operations in London after only 18 months, as they were not able to secure parking permits in all areas of the city. The Korean government has also announced that free-floating service is under consideration, and will start trial services in the near future starting with Busan and Sejong cities (Ministry of Economy and Finance, 2018). The most recent type of service is peer-to-peer service, where the vehicle to be serviced is not owned by a certain business but owned by individuals, who offer the vehicle to be driven by others while it is not in use. However, this is limited to round trips as the user has to return the vehicle to the owner.

Recently, car sharing market has seen rapid growth globally, with large number of service providers changing the fuel type of their vehicle fleet to electric vehicles. As the European Commission have set GHG reduction goals, the entire industry is moving forward to meet the requirements (Kim, 2015). As of 2019, 66% of car sharing vehicles in Europe are either electric or hybrid vehicles. 25% of car sharing available countries have cities that only provide electric car sharing vehicles. The effort is expected to move forward even more, with the legislation requiring mobility service providers are obligated to provide environment friendly services, and be penalized otherwise.

As mentioned in some cases, the introduction of the NGT services now allows consumers to choose different types of transportation modes other than the traditional transportation services, which never occurred in the past decades. Such change will affect consumers traveling behaviors, usages of traditional transportation services, and possibly their intention to own a vehicle. Thus, careful analyses on the effect of the aforementioned services in consumer behavior and preferences are necessary to implement suitable policies and marketing strategies in the future transportation services.

1.2 Research Objective

As mentioned, this dissertation focuses primarily on applying both ANN and different types of choice models to estimate the market potential of NGT services and to analyze their effects on the current transportation service industry based on consumer preferences. Unlike previous studies that only used ANN or a choice model, this study uses both types of model to analyze different types of consumer decisions and thereby derive valuable insights regarding the market potential and effect of NGT services on the transportation market. Based on the consumer choice type, suitable models are proposed, including different artificial neural networks (ANN). The study investigates the possibility of using ANN in consumer decision studies in two specific areas: predicting consumer decisions and enhancing variable selection processes. The results derived from the ANN are compared to the results generated using traditional methods to validate the application of the ANN. This study also utilizes choice models such as hierarchical Bayesian and MDCEV to analyze consumer preferences in detail.

Consumer decision types are defined using John Dewey's buyer decision process theory. Understanding the consumer decision process has been a pivotal goal of marketing research.

One of the pioneering theories that defined the consumer decision-making process is the buyer decision process theory developed by John Dewey (Dewey, 1910). According to John Dewey's theory, there are five different stages related to consumers' decision making. The first stage is defined by problem and need recognition, where a consumer recognizes a certain need or problem that needs to be solved. This stage is important because without awareness of a problem or need, a consumer will not be motivated to choose or buy a product (Bruner and Pomazal, 1988). According to Maslow, consumer needs exist in a hierarchy whereby satisfaction levels increase as levels of need increase. Consumers can only move up the hierarchy if their needs at a certain stage are fulfilled, and individuals' tendencies to move up the hierarchy is factor that drives consumers' recognition of need (Maslow, 1943).

The second stage is defined by information search, where a consumer searches for potential solutions or products that may fulfill the need or solve the problem identified in the first stage. The development of technology and the exponential increase in the availability of information have made consumers' information searching processes easier and faster (Bunn, 1993). Once the consumer collects information regarding potential alternatives, he or she moves to the third stage of the decision process. The third stage is defined as the evaluation stage, where the consumer evaluates different alternatives using the information collected in the second stage. At this stage, the consumer does not make a final decision but tries to identify the best choice based on his or her preferences. Many different factors affect the evaluation process and many econometric models including various discrete choice models have been developed to analyze the evaluation process. Different models incorporate different factors such as consumers' demographic information, attitudes, perceptions, and involvement (McFadden, 1974; Train, 2009; Allenby and Rossi, 2005; Bhat, 2005).

After the evaluation stage, the consumer moves to the fourth stage, which is defined as the purchase decision. In the fourth stage, the consumer finally makes a choice, or purchases a product. Many regards the process from the first stage to the fourth stage as the essence of the consumer decision-making process. For example, studies like Engel et al (1973) and Block and Roering (1976) extended the decision process by emphasizing the importance of environmental influences. These studies proposed that factors such as consumer income, attitude, culture, and other social factors work as constraints that strongly affect the decision process between the first and fourth stages.

The last stage of the decision process is post-purchase behavior, which is important in retaining consumers in the future. Until the fourth stage, the consumer builds up expectations regarding the return for or utility of making a choice. However, once the decision is made and the actual utility of the alternative is realized, the consumer may experience a higher-than-expected level of satisfaction, the expected level of satisfaction, or dissatisfaction. Satisfaction level is an important factor in studies of brand loyalty and future decision making. (Blythe, 2009; Foxall, 2005). Such feedback affects the consumer's future choices and even the choices of other consumers who seek information from the consumer who made the decision. Although Dewey's consumer decision process mainly focuses on the decision to purchase a product, it is also applicable to consumer decisions to use certain services (Zeithaml and Bitner, 2003).

Thus, this study applies John Dewey's choice stage process to analyze the decision processes of consumers. The first empirical study analyzes the intentions of consumers to adopt new products/services. The second empirical study analyzes consumers' preferences regarding a specific attribute of the new products and services. The third empirical study analyzes consumers' usage distribution of products based on their preferences. The results of these three analyses are combined to provide valuable insights regarding the effects of NGT services on the transportation service market. The key findings of this dissertation have both marketing- and policy-related implications for the promotion of a healthy transportation service market that takes consumers' intentions and preferences into account. The figure below depicts this study's research framework.

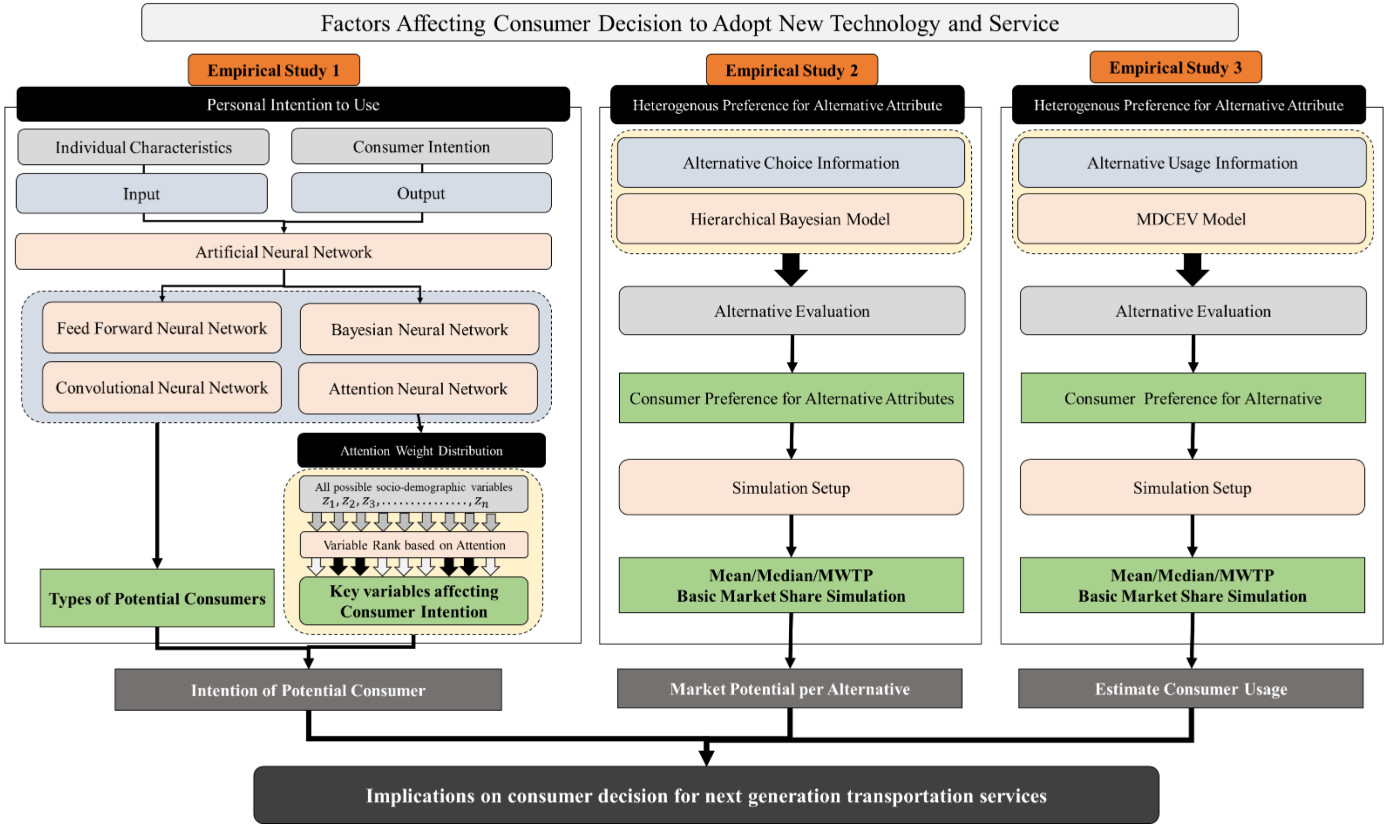


Figure 1. Research Framework

1.3 Research Outline

The dissertation is organized into five chapters. In Chapter 2, previous literatures regarding technology adoption theories are reviewed to provide theoretical background for the current research framework. Literatures regarding the component of the artificial neural network and the possibility of using it to analyze consumer decision making is reviewed. Also, various choice models that analyze consumer preferences are reviewed. Finally, studies comparing ANN and choice models are reviewed to provide the limitations of the previous studies. Based on the review, research motivation and the possible contributions of this dissertation is described.

In Chapter 3, models used to analyze different types of consumer's decisions are proposed. Models used in the first empirical study consist of the various artificial neural network used to predict the consumer's intention and identify key factors affecting them. Models used in the second empirical study is the hierarchical Bayesian model, which analyzes consumer preferences of NGT alternatives. Models used in the third empirical study consists of MDCEV model, which analyze the usage of different transportation services after the NGTs are introduced to the market. In Chapter 4, the results of the using different models are described. The first study focuses on predicting the consumer's intention of using the NGT and test the potential of using attention based ANN in identifying key features. Both traditional variable selection and ANN based methods are used to identify key features to provide the best model fit. The second study focuses on analyzing consumer preferences for attributes related to different types of NGT by using the hierarchical Bayesian model. The third study focuses on analyzing the change in consumers' usage of transportation modes when NGT are introduced to the market. In addition, the environmental effects due to the change in consumer transportation usage is estimated. The last chapter, Chapter 5, summarized the results of the empirical studies and present key findings of this dissertation. Moreover, limitations of the current study are discussed along with directions for potential research topics in the future.

Chapter 2. Literature Review

2.1 Product and Technology Diffusion Theory

Study regarding the consumer's adoption of new technology and products have been the epitome of many marketing studies. In particular, with the development of technology in recent decade, the importance of technology acceptance models has increased significantly. Many of the theories attempt to define consumers' motivation and intention to adopt behind their choice behaviors. Similarly, the present paper defined consumer decision making behavior based on different stages of choice.

In essence, the present paper takes the framework from buyer's decision process developed by John Dewey in 1910, later improved by Engel et al. in 1968 to define the choice stages, and define factors affecting different choice stages using factors from TAM and TRM. In addition, the study overcome the previous criticism towards TAM and TRM by using different methodologies to analyze consumer behaviors and provide valuable insights regarding their preferences. Before explaining the current research model, the paper overviews the previous literatures regarding the development of the technology acceptance models.

Another popular theory regarding the consumer's behavior of adopting technology is the theory of Reasonable Action (TRA) proposed by Fishbein and Ajzen in 1975. TRA defines consumer attitude as his evaluation of an object, belief as a link among the object and other attributes, and behavior as an outcome of his intention. Thus, consumer’s attitudes are affected by beliefs, which later affect the behavior of the consumer towards the object. In addition, consumer’s subjective norms also affect consumer’s behavior towards an object. Subjective norms are not only self-driven, but are defined by the community’s attitude to certain behavior towards an object. (Fishbein and Ajzen, 1975) The overall structure of the TRA is described in Figure (2).

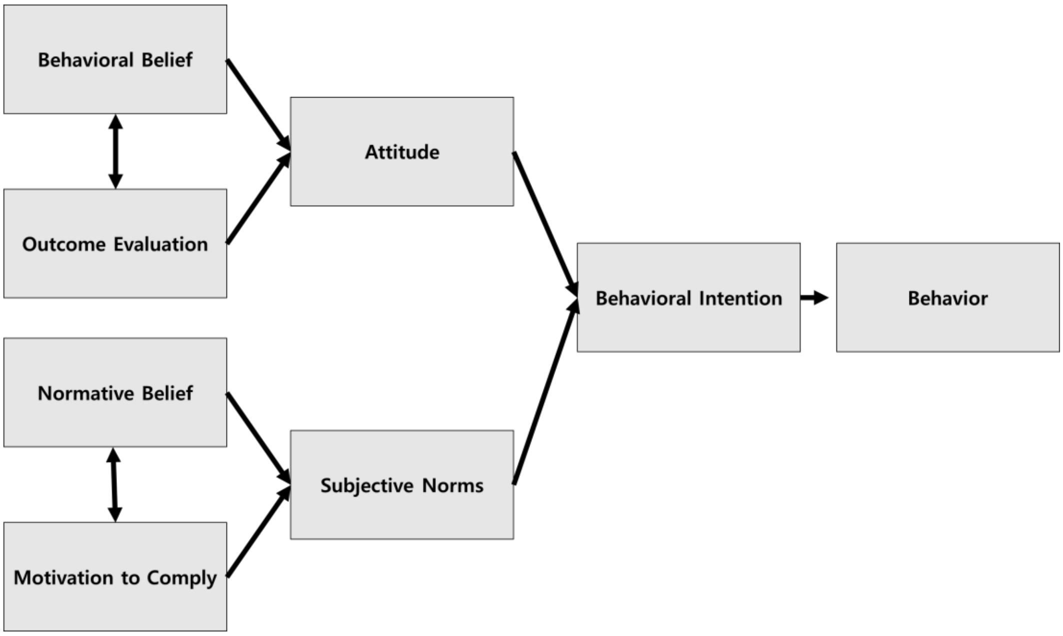


Figure 2. Theory of Reasonable Action. (Fishbein and Ajzen, 1975)

Ajzen (1991) later extended TRA to Theory of Planned Behavior, which added the perceived behavioral control in addition to attitude and subjective norms defined in TRA. This perceived behavioral control is defined as the consumer’s control that may limit their behavior towards an object or product. For example, although certain technology may seem fitting to use for some consumers, they may be limited by some factors or actions necessary to use/adopt the technology. (Ajzen, 1991) The theory of planned behavior has been greatly improved by other researchers by extending the factors that affect consumer behavior towards new product and technology. (Taylor and Todd, 1995; Shih and Fang, 2004) One of the most famous extension of TRA is the Technology Acceptance Model by Fred Davis in 1986.

The Technology Acceptance Model (TAM) developed by Fred Davis in 1986, is part of the information systems theory that model how consumers come to accept a technology and how they use it. Specifically, the first TAM was constructed to model user's acceptance of information systems and technologies. The overall procedure of the first TAM is shown in Figure (3).

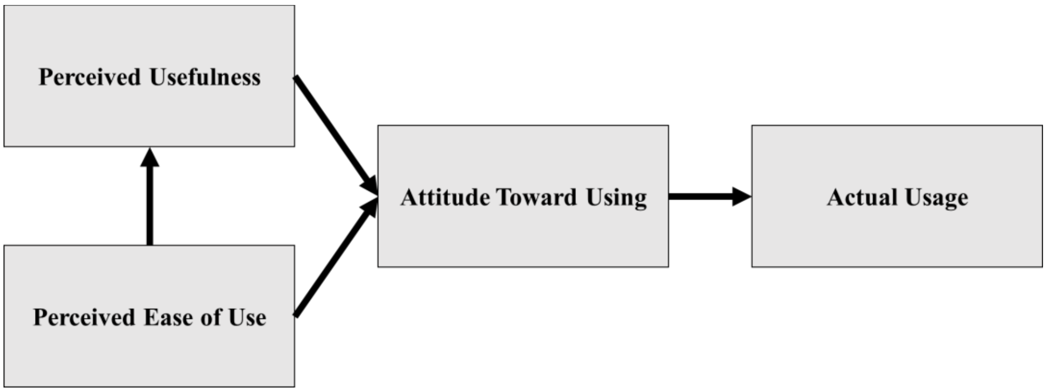


Figure 3. Technology Acceptance Model (Davis, 1986)

The general TAM included two specific beliefs: perceived usefulness and perceived ease of use. Perceived usefulness is defined by the user’s subjective belief that using such technology will improve the current task. On the other hand, perceived ease of use is defined by the expectation that such technology or innovation will be easy to use. In 1989, Davis and Warshaw improved the general TAM by defining external variables that affected two beliefs mentioned. According to the new model, any factors that affect these beliefs are defined by external variables in TAM. The modified version of TAM is shows in Figure (4). (Davis, 1989)

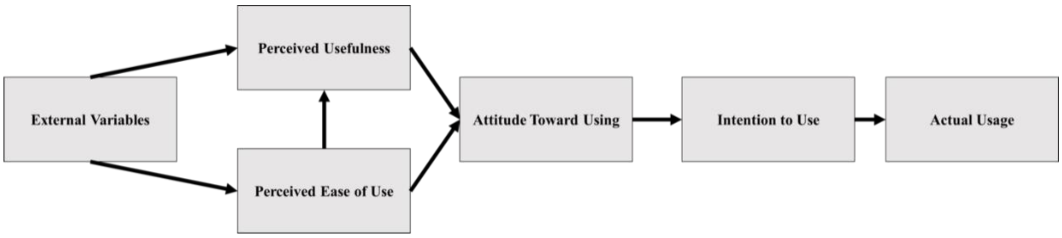


Figure 4. Modified TAM Model (Davis and Warshaw, 1989)

Another modification of TAM model was conducted by McFarland and Hamilton in 2006, where the model now included an addition of contextual variables. The authors assumed that these additional contextual variables affected the dependent variable through three mediating variables. In addition, the model also assumed a direct relation between the external variables and the system usage, not only dependent on the mediating variables such as ease of use and perceived usefulness like before. (McFarland and Hamilton, 2006)

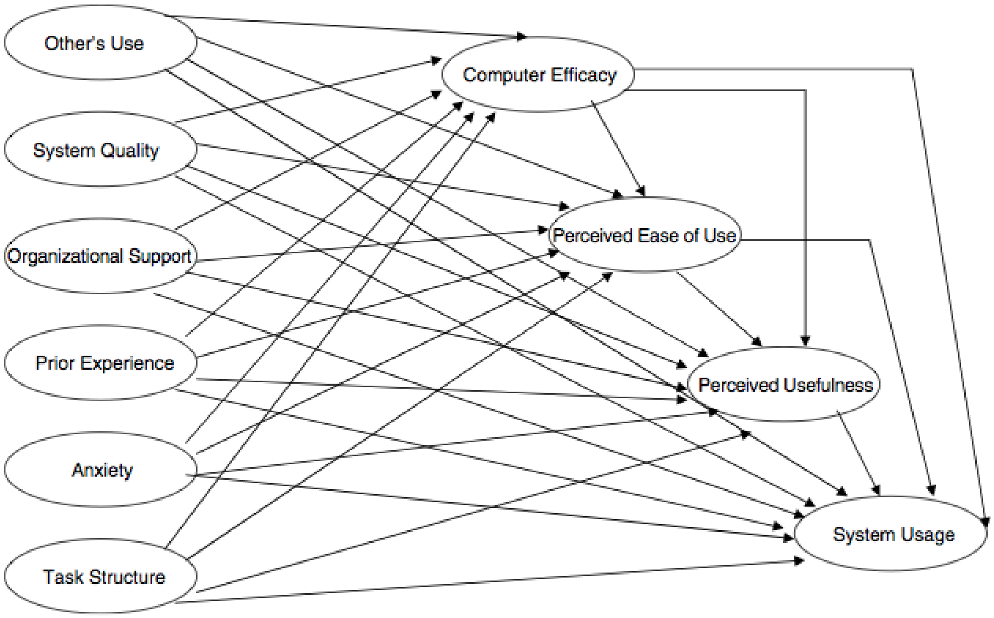


Figure 5. TAM model with contextual specificity (McFarland and Hamilton, 2006)

McFarland and Hamilton's model demonstrated that system usage was directly and significantly affected by other's use, system quality, organizational support, prior experience, anxiety, and task structure. This was also in line with the effect of the mediation terms. In some cases, the effects among contextual and the mediation factors are in the opposite direction from each other.

There were some researches that also focused on explaining the adoption behavior based on the relationship between technology and the task it performed. Goodhue et al. introduced the theory of task-technology fit (TTF) that emphasized individual's impact. The authors assumed that a good fit between task and technology is only achieved by increasing the likelihood of utilization and the performance impact of the technology to meet the task in needs. Ever since the introduction, TTF has been considered as a measurement of technology applications already released in the market. (Goodhue et al., 1995) The overall structure of the TTF is shows in Figure (6).

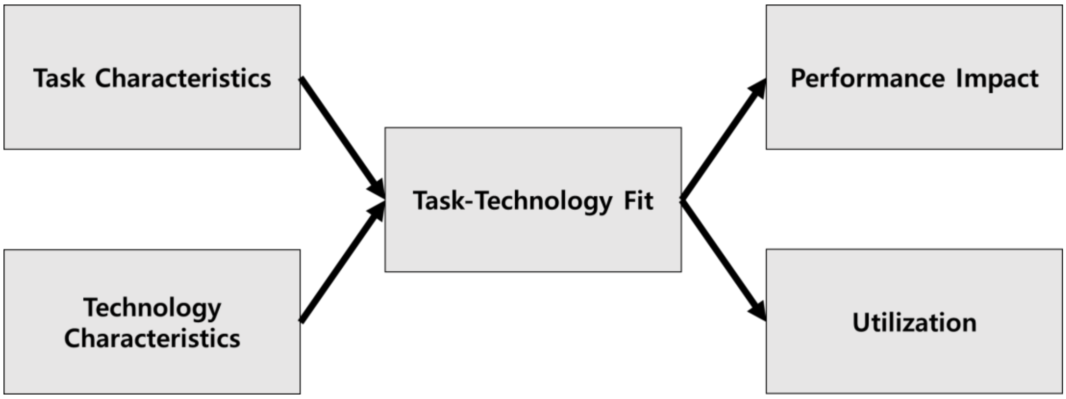


Figure 6. Task-Technology Fit (Goodhue and Thompson, 1995)

2.1.1. Extension of Adoption Models

With the development of different kinds of technology acceptance theory, there were also many studies comparing these aforementioned theories to suggest the best form of adoption model. In particular, the studies that compared the acceptance models mainly considered three types of models: TAM (Technology Acceptance Model), TRA (Theory of Reasoned Action), and TPB (Theory of Planned Behavior). Davis, Bagozzi and Warshaw's (1989) study considered TAM and TRA, comparing the advantages and disadvantages of the two model. Based on their findings, the authors proposed a unified model which was based on three determinants, perceived usefulness, perceived ease of use, and behavior intention. Their unified model suggested that social norm was not an important determinant that affected the behavior intention of the consumer as in TRA or TPB.

Study by Mathieson (1991) and Yi, Jackson, Park, and Probst (2006) proposed that social factors are also one important factor affecting consumers' adoption behavior. Thus, these studies attempted to extend TAM by adding factors from TPB to incorporate social factors in explaining the adoption behavior. In addition, study by Davis, Bagozzi and Warshaw (1989) proposed that social norms may not have any significant influence on consumers' behavior intention, especially when the technology of the focus point is specified in individual usage. On the other hand, different types of TAM were specially designed to address the factors of users' technology acceptance (Chau and Hu 2002). Thus, the studies that compared TPB with TAM confirmed that Technology Acceptance Model was easy to apply across different research settings. Other studies also noted that using TAM was more suitable than TRA or TPB. (Han, 2003; Lai and Zainal 2014; 2015) In fact, TAM model developed by Davis has been the most used framework in predicting information technology adoption (Paul, John and Pierre, 2003). However, some studies argued that TAM was too focused on only perception of technology's usefulness and convenience but also include all of the factors affecting adoption intentions beyond them. (Luarn and Lin, 2005; Lai and Zainal, 2015; Lee and Jun, 2007).

Similar to the studies shown above, there have been many types of studies that attempted to modify the adoption theories to better explain the process of consumer adoption behavior and provide meaningful interpretations through the analysis. Thus, the present research also incorporates perceived usefulness and perceived ease of use to define consumer's decision regarding the intention to use certain technology. In addition, the present study enhances the evaluation of alternatives stage and the decision of usage stage to provide insights unprovided by the previous technology acceptance models. By defining these additional choice stages, the current study is able to provide not only the consumer types that are more likely to have intention to adopt new products, but also provide key marketing and policy implementations based on consumer preferences.

All of the theories mentioned in this section attempt to explain the adoption of innovation by consumers. Compared to the present study's research framework, which is show in Figure (7), many of the previous literatures assumed that consumer behavior and intention to use certain technology or innovation was affected by their perception towards the them. Although different in form, theories such as Davis' TAM and its modifications proposed that consumer's intentions were the basis behind their choice to use certain products. This framework remains intact in the present study. Although the present research doesn't focus so much in defining the external variables affecting consumer perceptions, it focuses on explaining the relationship among different consumer characteristics and their intention to use new products or services in the future. As mentioned, by using both econometric techniques and the artificial neural network models, the present study aims to provide valuable interpretations regarding how consumer's intentions are affected by their characteristics or perceptions, and which type of consumers are more likely to adopt NGT in the near future. Different assumptions are assumed to define the focus of each choice stages.

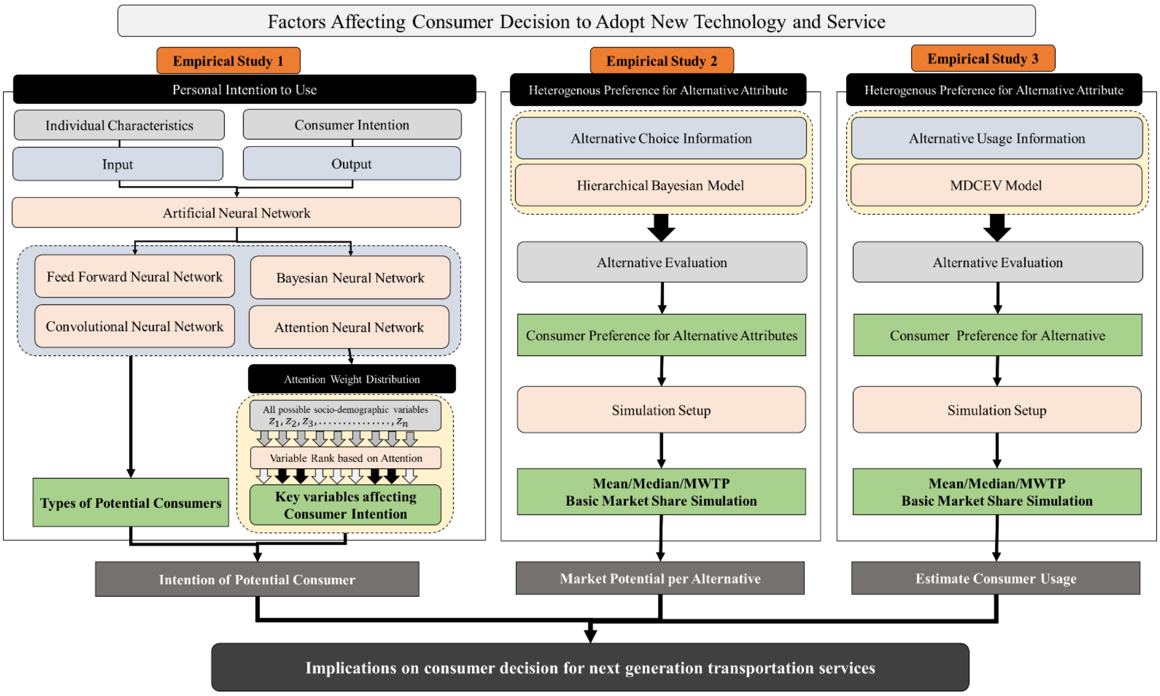


Figure 7.Research Framework

2.2 Artificial Neural Network

2.2.1 General Component of the Artificial Neural Network

Artificial neural network (ANN) are commonly used model of machine learning that inherits the biological neural network. ANN, similar to the human brain and unlike other programs, ‘learn’ to perform tasks without being programmed with solid rules. ANN was first introduced as a non-linear algorithm that mimics the process of human brain to a mathematical model. (McCulloch and Pitts, 1943; Ivakhnenko, Lapa and Valentin, 1967;

Schmidhuber, 2015) Although McCulloch's model had similar form to the current multilayer artificial neural network, consisting of interconnected neurons with synapses to compute an outcome, it's learning performance was very limited due to fixed weights of the connection. The learning performance began to improve once perceptron, which allows repeated learning and the weights to adjust was developed. (Rosenblatt, 1957) Although the key components of ANN were established for some time, real application of ANN began when the multi-layer perceptron was developed. (Minsky and Papert, 1986) A multi-layer ANN is the basic structure of the current ANN in that it allows stacking multiple layers to form a network with perceptron that learns, or adjust the weights to minimize the error of the output. With the introduction of multi-layer neural network, layers have been labeled according to their purpose. First layer where the data enter the network was defined as the input layer, the layer where the output or prediction value came out was defined as the output layer, and any layers between the has the two layers was defined as hidden layers.

Another form of ANN was introduced by Ivanknenko and Lapa in 1965, which was based on Hebbian learning. Hebbian learning created by Hebbs in 1940's was a learning hypothesis based on the mechanism of neural plasticity. (Ivakhnenko, Lapa and Valentin, 1967; Schmidhuber, 2015) The initial motivation behind ANN was to construct a system that could perform tasks in a similar way the human brain performed. However, with the development of computing power and the expansion of data availability, the application of ANN has been expanding from simple image recognition to big data analysis such as medical diagnosis, painting, modeling, and even writing novels. In case of image and text recognition, where ANN learn to identify images that contain certain objects by learning various components of data that are labeled as that image. The core framework of the ANN is the input and output data, which are labeled as such, that will train the ANN system to learn the patterns connecting them, and later produce similar output or task based on learning. (James, Witten, Hastie, and Tibshirani, 2013)

To perform a certain task, ANN consists of collection of nodes called the artificial neurons. These neurons are connected with one another, which represent a synapse in the biological brain. Each synapse gives a signal to one another, which in ANN are a real number, and the signal are computed by some non-linear function. The amount of signal certain node sends out have a weight to them, depending on the strength of the signal. The weight of the signal is adjusted during the learning process, depending on the training data, certain signals may become too weak and stop sending out signals, while some may become very strong compared to the others. Since there are numerous neurons that send out signals in ANN, they are aggregated into certain layers. The number of layers can become immensely as learning process becomes more and more complicated. Signals from the first input layer will travel throughout these layers until the signal is computed through the last output layer. Some signals may become too weak and be lost during the process.

The ANN model can be distinguished by different types of learning process. The general methodology consists of feed-forward and backpropagation method. Both learning methods can be applied to ANN with input layer, hidden layer, and the output layer. In the feed-forward model, the vector values are transferred from the input layer to the hidden layers between the input and the output layer. There is no circulating path in the feedforward model. The backpropagation model updates the weights by calculating the error between the predicted and actual output values. (Svozil, Kvanicka, Pospichal, 1997; Hecht-Nielsen, 1989)

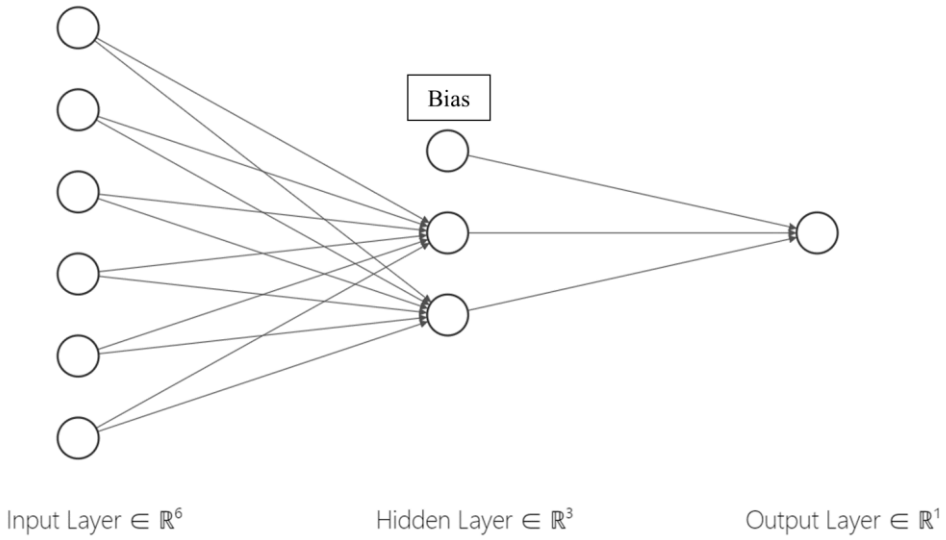


Figure 8. General form of multi-layer ANN structure with 1 hidden layer

The most popular type of ANN is the multi-layer feed-forward neural network (MLFN). (Gan et al., 2005) In the MLFN, information flows in the direction from input to the output, with three types of computational units; input layer, hidden layer, and the output layer. The input layer is where the initial input data are injected into the model and transferred into the next layer. Since the output of one layer is an input of the next layer, the output of the network can be denoted algebraically in the equation below.

$$Z = F\left(\sum_{j=1}^{J} W_j^{(l_2)} U_j\right) = F\left(\sum_{j=1}^{J} W_j^{(l_2)} \times F\left(\sum_{i=1}^{I} W_{ij}^{(l_1)} X_i\right)\right) \cdots\cdots\cdots \text{Eq. 1}$$

Where Z denote the output of the network, F denote the transfer function of the output node, $W_{ij}^{(l_1)}$ and $W_j^{(l_2)}$ denote the connection weights from input layer (i) to hidden layer (j), and to output layer in the end. The training process of the neural network is the process of calculating the weights within the neural network. Training process begins by randomly initializing connection weights when the input and the output data enter the network. Then, the network proceeds to calculate an output, which is compared to the actual output data to calculate the error. Based on the calculated error, the network adjusts the connection weights of the nodes by propagating the error backwards from the output layer. As a result, network attempts to minimize the network total mean squared error and improve the prediction accuracy by determining the best adjustments of the interconnection weights among neurons.

2.2.2 Activation Functions of Artificial Neural Network

Along with the neural nodes and layers, another important component of the ANN is the activation function. Activation functions of the node determine which output is produced in a certain node given an input within an ANN model. The main task of the activation function is to adjust the weights and bias as learning occurs, a function that progressively changes from 0 to 1. The activation function can be divided into two types: linear activation function and the non-linear activation functions. However, only non-linear activation functions are feasible in computing complex networks with small number of nodes.

The linear activation function is a simple function that assumes a linear relation between the input and the output. With its simple form, the linear activation is hardly used in the analysis, because it cannot represent the complexity of the various parameters the data contain. Non-linear activation functions are the most commonly used activation function that helps model or generalize the complexity of the data to produce an output. There are various different kinds of non-linear activation function depending on their range or curves. Most popular types of the non-linear activation functions are Sigmoid or Logistic, Tanh, ReLU, and Softplus functions.

Table 2. Types of Activation Function

Type	Equation	Derivative
Linear/Identity	$f(x) = x$	$f'(x) = 1$
Binary Step	$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$
Logistic/sigmoid	$f(x) = \sigma(x) = \dfrac{1}{1+e^{=x}}$	$f'(x) = f(x)(1-f(x))$
Tanh	$f(x) = \tanh(x) = \dfrac{e^x - e^{-x}}{e^x + e^{=x}}$	$f'(x) = 1 - f(x)^2$
Rectified Linear Unit[1]	$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ \text{x} & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ 1 & \text{for } x > 0 \end{cases}$
Softmax	$f(y_i) = \dfrac{e^{y_i}}{\sum_j e^{y_j}}$ where $j \neq i$	$f'(x) = f(x_j)(1-[f(x_i)])$
SoftPlus[2]	$f(x) = \ln(1+e^x)$	$f'(x) = \dfrac{1}{1+e^{=x}}$

Sigmoid or a logistic function is a function that are shaped like as S-shape, within the values between (0,1), and is differentiable. It is usually used in the models that predict the probability as an output. Although the S curve useful in modeling a complex learning curve that accelerates through the progress, since it is convex for values less than 0, and concave for values greater than 0, it could possess multiple optima. As a result, it may stop the neural network to stop during the learning process. Also, another downside of the sigmoid

[1] Nair & Hinton, 2010.
[2] Xavier, Bordes, Bengio, 2011.

function is that it is computational expensive and cause vanishing gradient problem. (Han and Claudio, 1995)

Tanh function is an extension of the sigmoid function, with the values between (-1,1), also in a S-shape. Similar to the sigmoid function, Tanh function is also differentiable, but also cause the vanishing gradient problem. Both sigmoid and tanh functions have potential to cause the vanishing gradient problem, which originates from squeezing too much information into a limited space. The problem occurs when the layers within a neural network become exponentially big and the gradient between the layers become too small for the training to occur. The gradient of neural networks is found using backpropagation, which are calculated by derivatives of the network while moving backward from the final layer to the initial layer. According to the chain rule, the derivatives of each layer are multiplied down the network to derive the derivatives of the initial layer. However, if the gradient become too small at certain hidden layer, the next derivative may become exponentially small, resulting in no updates to the weights and the biases within the model. Thus, since the size of the gradient are related to how the weights and the biases are updated during the training session, small gradient can halt the training of the model. (Hochreiter, 1998). As a result, different activation functions have been developed to avoid the computational difficulty and the vanishing gradient problem.

The ReLU function is a monotonic rectified linier unit activation function that is widely used in ANN modeling. Unlike the sigmoid function, ReLU has a value between

[0, ∞), meaning that it is half rectified from 0. The ReLU function is denoted by Equation (2) and the graph of ReLU activation function is shown in Figure (9)

$$R(z) = \max(0, z) \quad \text{Eq. 2}$$

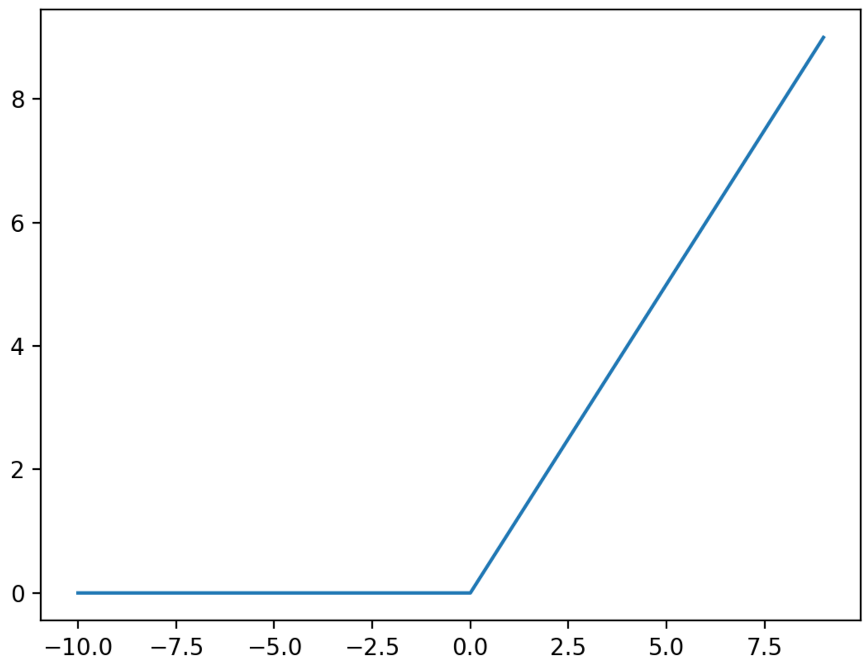


Figure 9. Shape of ReLU activation function

Using the ReLU function has many benefits over the sigmoid or tanh activation function. The benefits of the ReLU function is its computational simplicity and reduced a likelihood of vanishing gradient. Compared to computing an exponent of input for sigmoid or tanh function, the ReLU function only requires a simple calculation. Although such difference may be insignificant in ANN models with small number of layers, it significantly affects the training and performance in dense deep neural network. Also, recall that gradient of sigmoid function becomes increasingly smaller as the absolute value of input becomes larger. Such relationship can be shown in Equation (3). As a result, the sigmoid activation has chance of being saturated with relatively narrow interval of inputs where the derivative of the function is sufficiently nonzero. On the other hand, when $x > 0$, the gradient of the ReLU function has a constant value, reducing the chance of vanishing gradient and accelerating the learning process.

$$
\begin{aligned}
&Sig'(x) = Sig(x)(1 - Sig(x)) \qquad \text{as "}x\text{" grows to infinite large} \\
&Sig'(x) = Sig(x)(1 - Sig(x)) = 1 \times (1 - 1) = 0
\end{aligned}
\qquad \cdots\cdots\cdots\cdots\text{Eq. 3}
$$

However, due to assuming all of the negative values to be 0, ReLU can lead to most of the activations within the layers to become 0. Such problem is defined as a dying ReLU problem. A modification of the ReLU function in order to solve this problem is the Leaky ReLU function, which in contrast to the ReLU, assigns a nonzero slope to the negative values. The form of leaky ReLU model vary depending on how the values lower than 0 are assigned. Models such as Leaky ReLU assumes that ReLU has a very small value defined by Equation (4). The Parametric Leaky ReLU model assumes the same functional form except that it defines the value of a is learned from training in the back-propagation process. (Xu et al. 2015)

$$f(x) = \begin{cases} \dfrac{x}{a} & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \quad \cdots\cdots \text{Eq. 4}$$

2.3 Modeling Consumer Choice: Discrete Choice Model

2.3.1 Multinomial Logit Model

The first form of discrete choice was proposed by McFadden (1974), which took a form of Multinomial Logit Model. The MNL model assumes a choice situation where the consumer chooses the best alternative among different alternatives. MNL model is considered as a logit model because the probability density function of the uncertain part ε is assumed to be independent and follow an independent, identical Type 1 extreme value distribution. The difference between the random variables that follow a Type 1 extreme value distribution has a logistic distribution, making it possible to estimate the cumulative density function.

By assuming IIA (Independence from Irrelevant Alternatives) the multinomial logit model has its advantages of being easy to estimate due to it simple form. However, due to IIA assumption, MNL is unable to reflect the reality of the substitute or complimentary relationships among alternatives. The choice probability of an alternative must be changed only by adding or deducting other alternatives. Also, the MNL model assumes that every individual have the same coefficient for a particular attribute, making it impossible to reflect individual heterogeneity in preference. (Train, 2009). As a result, MNL has the advantages of being easy to use but is limited in reflecting the realistic nature of choice situation or consumer preference due to its simple and strong assumptions.

In the multinomial logit model, a consumer chooses the best alternative based on the deterministic part, which the researcher can observe, and the unobserved random part. The structure of consumer's utility is as follows

$$U_{njt} = V_{njt} + \varepsilon_{njt} \qquad \cdots\cdots \text{Eq. 5}$$

n in Equation (5) represents the individual making the choice, j represents the alternative, and t represents the choice situation. U_{njt} denote the utility of individual n choosing j at choice situation t. V_{njt} denotes the deterministic part, while ε_{njt} denotes the random part. The deterministic part of the utility can be denoted in a functional form of the alternative's attributes.

$$U_{njt} = V_{njt} + \varepsilon_{njt} = V(x_{njt}, s_{nt}) + \varepsilon_{njt} \cdots\cdots \text{Eq. 6}$$

x_{njt} denote the attributes of the alternative j, and s_{njt} denotes the attribute of the consumer n. The random utility theory simplifies the model by assuming linear form for these random part. The equation based on these assumptions are as follows

$$U_{njt} = V(x_{njt}, s_{njt}) + \varepsilon_{njt} = \beta'_{njt} + x_{njt} + \alpha'_{njt} s_{nt} + \varepsilon_{njt} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\text{Eq. 7}$$

β_{njt} represents the marginal utility of consumer n, and α_{njt} represents consumer's coefficient of the attribute. Based on this framework, the choice probability of consumer n choosing alternative i can be estimated according to Equation (8)

$$\begin{aligned} P_{\backepsilon} &= \Pr(U_{\backepsilon} > U_{nj} \quad \forall \mathrm{j} \neq i) = \Pr(V_{\backepsilon} + \varepsilon_{\backepsilon} > V_{nj} + \varepsilon_{\backepsilon} \quad \forall \mathrm{j} \neq i) \\ &= \Pr(\varepsilon_{nj} - \varepsilon_{\backepsilon} < V_{\backepsilon} - V_{\backepsilon} \mathrm{j} \neq i) \end{aligned} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\text{Eq. 8}$$

The uncertainty part $\varepsilon_{\backepsilon}$ in Equation (8) has the same structure as the cumulative density function (CDF) meaning that the probability density function of $\varepsilon_{\backepsilon}$ can be calculated by integrating $f(\varepsilon_{\backepsilon})$. The discrete choice model can take many forms according to how $f(\varepsilon_{\backepsilon})$ is defined.

2.3.2 Mixed Logit Model

Unlike the multinomial logit model introduced in the earlier section, the mixed logit model relieves some of the constraints of the model to express the heterogeneity of each consumer preferences. The MNL model assumes that every individual has the same preference for a particular attribute; however, the mixed logit model assumes a distribution on the estimated parameters for the attribute of the alternative. The researcher has the control over which distribution to assume for a particular attribute. Although the normal distribution is most commonly used, different kinds such as log-normal, truncated normal, censored normal, and many of the other distributions can be applied accordingly. For example, if the preference of a certain attribute is believed to be negative or positive, like the case of price or cost attribute of the alternative, assuming a log-normal distribution, which presumptively presume the direction of the parameter is common. (Train and Sonnier, 2005). The mixed logit models that assume a particular direction for certain parameter is called a theory-constrained mixed logit model (T-MIXL) (Keane and Wasi, 2013).

The utility function of the mixed logit model is as follows in Equation (9). The consumer utility U_{njt} consists of determinant term V_{njt} and stochastic random term ε_{njt}, which similar to the MNL, is assumed to follow the type I extreme value distribution. However, unlike the MNL, consumer n chooses the alternative j in choice situation t, where the preference parameter β_n follows the normal distribution with mean b and variance W (McFadden and Train, 2000).

$$U_{njt} = V_{njt} + \varepsilon_{njt} = \beta_n' x_{njt} + \varepsilon_{njt}$$
$$\beta_n \sim N(b, W) \qquad \cdots\cdots \text{Eq. 9}$$

The choice probability of the mixed logit model is denoted as Equation (10). The choice probability is an integral form of the multinomial logit probability $L_{nit}(\beta_n)$, where the density function of β_n is assumed to follow the $f(\beta_n \mid b, W)$ distribution.

$$P_{nit} = \int L_{nit}(\beta_n) f(\beta_n \mid b, W) d\beta_n$$
$$L_{nit}(\beta_n) = \frac{e^{V_{nit}}}{\sum_j e^{V_{njt}}} = \frac{e^{\beta_n' x_{nit}}}{\sum_j e^{\beta_n' x_{njt}}} \qquad \cdots\cdots \text{Eq. 10}$$

The likelihood function of consumer n choosing alternative i in choice situation t is denoted as $y_{nit} = 1$, and $y_{nit} = 0$ otherwise. Consumer n's likelihood function is denoted as the following Equation (11)

$$P_n = \int \prod_t \prod_i \{L_{nit}(\beta_n)\}^{y_{nit}} f(\beta_n \mid b, W) d\beta_n$$
$$Likelihood = \prod_{n=1}^{N} P_n = \int \prod_{n=1}^{N} \prod_t \prod_i \{L_{nit}(\beta_n)\}^{y_{nit}} f(\beta_n \mid b, W) d\beta_n \qquad \cdots\cdots \text{Eq. 11}$$

2.3.3 Latent Class Model

Similar to the mixed logit model, latent class model can also express consumer heterogeneity in preference. Unlike the mixed logit model, where each individual preference is estimated, the latent class model assumes that consumers' preferences are divided into several clusters or categories. Every consumer within the same cluster are assumed to have the same preference.

The utility function of the latent class model is similar to that of the multinomial logit model, consisting of deterministic $V_{njt|q}$ and random $\varepsilon_{njt|q}$ components of the alternative j. However, according to the latent class model, all of the consumers' preferences are divided into Q clusters, and the utility of $U_{njt|q}$ consumer n within category q choosing alternative j for choice task t is denoted as the following Equation (12). (Green and Hensher, 2003).

$$U_{njt|q} = V_{njt|q} + \varepsilon_{njt|q} = \beta_q' x_{njt} + \varepsilon_{njt|q} \quad \text{Eq. 12}$$

Similar to the multinomial logit model, if $\varepsilon_{njt|q}$ is assumed to follow the type I extreme value distribution, the choice probability can be expresses as follows.

$$P_{nit|q} = \frac{e^{V_{nit|q}}}{\sum_j e^{V_{njt|q}}} = \frac{e^{\beta_q' x_{nit}}}{\sum_j e^{\beta_q' x_{njt}}} \quad \text{Eq. 13}$$

Where β_q denote the coefficient parameter of consumers' preferences of cluster q. Accordingly, the likelihood function of consumer n from cluster q choosing alternative i in choice task t is denoted as $y_{nit} = 1$, and $y_{nit} = 0$ otherwise. Thus, the likelihood function of consumer n from cluster q is denoted as the following Equation (14)

$$P_{n|q} = \prod_t \prod_i \left(P_{nit|q}\right)^{y_{nit}} \quad \text{Eq. 14}$$

The likelihood is accommodated to analyze which cluster each consumer belongs to. The probability of consumer n being a member of cluster q is assumed to follow a multinomial logit form with Q th cluster as a baseline, and that individual characteristics z_n of consumer n affect consumer's cluster membership. Thus, the probability of consumer n belonging to q cluster is denoted as the following Equation (15)

$$H_{nq} = \frac{\exp(\lambda_q' z_n)}{1 + \sum_{q=1}^{Q-1} \exp(\lambda_q' z_n)} \quad \text{Eq. 15}$$

Similarly, the likelihood and sample likelihood of consumer n belonging to q cluster is denoted as the following Equation (16) and (17)

$$P_n = \sum_q H_{nq} P_{n|q} \quad \text{Eq. 16}$$

$$Likelihood = \prod_{n=1}^{N} P_n = \prod_{n=1}^{N} \sum_{q} H_{nq} P_{n|q} \quad \cdots\cdots \text{Eq. 17}$$

2.4 Modeling Consumer Heuristics in Discrete Choice Model

2.4.1 Consumer Decision Rule in Discrete Choice Model: Compensatory and Non-Compensatory Models

Consumer decision making is becoming more complicated than ever with the development of technology giving unlimited access to information, products, and services around the world. As a result, there has been a substantial amount of effort by researchers to better understand the consumer decision making behavior and incorporate consumer 's decision making heuristics into the discrete choice model. One of the well-known frameworks regarding consumer heuristics are the compensatory and non-compensatory models. Compensatory models are based on the assumption that consumer makes a choice by considering all of the attributes of an alternative. In other words, consumer chooses an alternative only after considering all of the positive and negative aspects of the attributes related to the alternative. On the other hand, the non-compensatory model assume that consumers only consider the given or only some of the attributes related to the alternatives before making a choice. There are many different forms of non-compensatory model depending on how and what the consumer considers about the alternatives. In general, there are four main models within the non-compensatory model; conjunctive, disjunctive, lexicographic, and elimination by aspect model. (Cantillo and Ortúzar, 2005).

Conjunctive models assume that consumer only view an alternative as viable option if all of its attributes satisfy consumer's minimum requirement. Disjunctive model is a relieved form of a conjunctive model in that it assumes consumer consider an alternative viable only if some of the attributes satisfy consumer's minimum requirement. In the disjunctive model, consumers consider an alternative being viable depending on whether their 'important' attributes are satisfied. The elimination-by-aspects (EBA) model was inspired by the cognitive effort minimization in the study by Amos Tversky in 1972. (Tversky, 1972) Reference 1 The EBA rule screen out the alternatives that contain attributes that has a level equal to (case of continuous), or doesn't contain (case of dummy variable) the required attribute until a single alternative exist. Models that follow elimination by aspects assume that consumer rank the attributes according to their importance and define the minimum required values to choose from. Lexicographic model assumes that consumer rank order attributes and narrow down alternatives according to the first rank. The reference-dependent model model was first proposed by Kahneman and Tversky (1979), which is based on the prospect theory that states an asymmetric reaction tendency for gain and loss domain. The model assumes that consumer have asymmetric reaction coefficients between gain and loss domain. The gain domain is domains that are preferable to the current status of the consumer (A_m) and the loss domain are not-preferable to the current status of the consumer (B_m). Reference dependent models are utilized in measuring the asymmetric reaction to positive gain and negative gain depending on the aforementioned domains, and is usually applied in analyzing consumer adoption behavior of high-tech products. (Junghun Kim, Lee, and Ahn, 2016) Random regret model is a model that analyze consumer preference based on the difference between the attribute levels of the chosen alternative and the attribute levels of the not-chosen alternatives. (Loomes and Sugden, 1982) Each of the models are defined as the following table. (Tversky, 1972; Hess et al., 2012; Chorus, 2014) Reference 2

Table 3. Decision Rules in Choice Model

Decision Rule	Mathematical Formulation of Decision Rule
Elimination-by-aspects	$y_i = 1 \leftrightarrow$ $x_{im} \geq \tilde{x}_m, \quad \forall m$ $\tilde{x}_m$: aspiration level for m-th attribute
Lexicographic	$y_i = 1 \leftrightarrow$ $x_{im} = \max_{\forall j \in C}\left[x_{jm}\right]$
Reference Dependent	$y_i = 1 \leftrightarrow V_i \geq V, \quad \forall j \in C$ $V_i = \sum_m \begin{pmatrix} -\bar{\beta}_m \max\left[0, \bar{x}_m - x_{im}\right] \\ +\vec{\beta}_m \max\left[0, x_{im} - \bar{x}_m\right] \end{pmatrix}$ $\bar{x}_m$: Reference of m-th attribute
Random Regret (I)	$y_i = 1 \leftrightarrow \mathrm{R}_i \leq R_j, \quad \forall j \in C$ $R_i = \max_{j \neq i}\left(\sum_m \max\left[0, \beta_m\left(x_{jm} - x_{im}\right)\right]\right)$
Random Regret (II)	$y_i = 1 \leftrightarrow \mathrm{R}_i \leq R_j, \quad \forall j \in C$ $R_i = \sum_m \sum_{j \neq i} \log\left[1 + \exp\left(\beta_m\left(x_{jm} - x_{im}\right)\right)\right]$

2.4.2 Choice Set Formation Behaviors: Semi-Compensatory Models

A two stage semi-compensatory model is a combination of compensatory and non-compensatory model that allows the researcher to strengthen the realistic aspects of consumer decision-making process. In a choice task with a large number of alternatives, many previous literatures have shown that consumers almost always employ a two-stage choice process (Payne, 1976; Lussier and Olshavsky, 1979). In the first stage, consumers follow the non-compensatory heuristics to select only the viable alternatives among universal choice sets. This process is consistent with the concept of choice-set formation, which reflect the fact that consumers cannot consider all of the alternatives when there are too many alternatives in the choice task. (Manski 1977; Swait and Ben-Akiva, 1986; Kaplan, Bekhor, and Shiftan, 2011) In the second stage, consumer follows the compensatory evaluation heuristic to choose the best alternative among the viable alternatives. Thus, the semi-compensatory model combines both compensatory and non-compensatory heuristics in one discrete choice model.

The general form of semi-compensatory model follows the two-stage model by Manski (1977). Manski focused on defining the structure of the random utility model which expressed the generation of choice sets. The author proposes that consumer's choice set formation behavior and choice probability of the alternative are related. Equation **(**18**)** denote Manski's model, which assumes that the probability of consumer n choosing alternative i not only depends on the choice set C_n, but the probability of the choice set formation $Q_n(C_n)$

$$P_{ni} = \sum_{C_n \in G} P(i \mid C_n) Q_n(C_n) \quad \cdots\cdots \text{Eq. 18}$$

Consumer n's choice set C_n are chosen from G, a collection of non-empty set from universal choice set M based on the probability of $Q_n(C_n)$. $P(i \mid C_n)$ represent the probability of consumer n choosing alternative i. Although this formulation is convenient in that it allows a simultaneous estimation of the choice set formation and decision behavior, the model's computational intensity increases exponentially depending on the size of M. Thus, there has been many improvements on the Manski's model.

One of the most well-known modification of the Manski model is the modification conducted by Swait and Ben-Akiva (1987). The authors modified the two stage model by representing the consumer n's choice set formation behavior as a binary decision, where $A_{ni} = 1$ if the alternative i is included in consumer n's choice set and 0 otherwise. Then, consumer n is assumed to follow the multinomial logit model for choice task. Hence, the probability of consumer n's choice set formation $Q_n(C_n)$, is modified into D_i in Equation (19).

$$Q_n(C_n) = \frac{\prod_{i \in C_n} D_i \prod_{j \in M - C_n} (1 - D_j)}{1 - \prod_{j \in M} (1 - D_j)} \quad \cdots\cdots \text{Equation 19}$$

$$D_i = \Pr(A_{ni} = 1), \ \forall i \in M \quad \cdots\cdots \text{Equation 20}$$

Along with the studies that focused on identifying the probability of the choice set formation within the choice task, studies that focused on analyzing the actual observed choice set. These studies modified the two-stage discrete model by integrating randomly distributed thresholds for attribute acceptance. In particular, Kaplan et al. (2012) proposed that decision-making process of consumer can be divided into two stages, a choice set formation stage and a choice stage as shown in figure below.

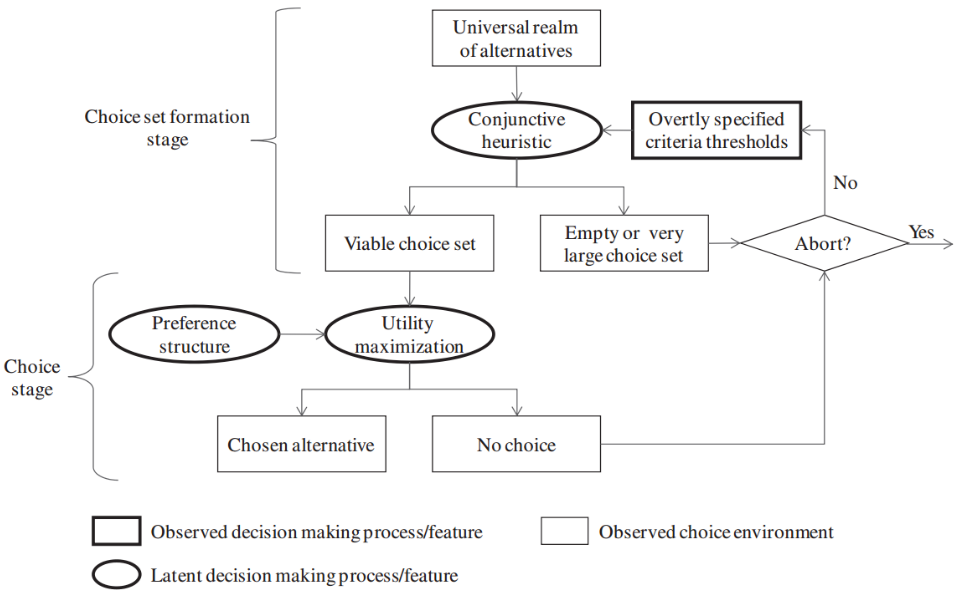


Figure 10. Two stage choice set formation and choice stage

Source: Kaplan et al. (2012)

The models assume that people behave according to non-compensatory, and in particular, elimination by aspects behavior. In other words, consumer have a threshold that are criteria that specify the set of feasible alternatives. The threshold parameters are expressed as a function of consumer's individual characteristics and the conditions of the choice task. These threshold parameters are random and different among individuals that decide whether to accept or reject certain alternatives based on their level of attributes. (Williams and Ortuzar, 1982; Cantillo and Ortuzan, 2005; Kaplan, Shiftan and Bekhor, 2012) As a result, the main difference between the traditional two-stage compensatory model and the semi-compensatory model incorporating individual threshold is that latter model the choice set reduction out of the universal choice sets rather than generating them exogenously. The frameworks of traditional compensatory and semi-compensatory models are presented in figure below.

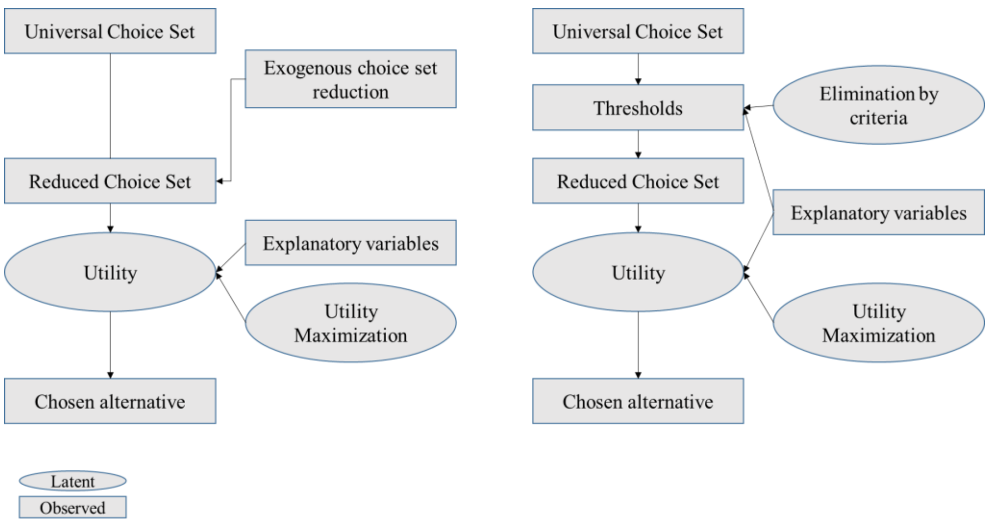


Figure 11. Traditional versus Semi-compensatory framework with threshold

Source: Kaplan, Shiftan and Bekhor, 2012

The formulation of these model is expressed as the Equation (21), where the probability of choice set formation behavior consist of different number of criterions that affect the decision of the choice set formation. (Kaplan, Bekhor, and Shiftan, 2011)

$$P_q(i \mid G) = \sum_{S \in G} P_q(i \mid S) P_q(S \mid G) \qquad \text{Eq. 21}$$

Where, $P_q(i \mid G)$ is the probability that individual q chooses alternative i and $P_q(i \mid S)$ is the probability that individual q chooses alternative i in choice set S, and $P_q(S \mid G)$ is the probability that individual q reduce the universal set of choices G to viable choice set of S during the first stage. The model assumes three types of assumptions when reducing the universal set of choices G to viable choice set of S. First, the process of elimination by aspect is traceable rather than being covert. Second, each consumer is only partially notified about the information regarding the universal set of choices G. Third, the process of reducing the choice set does not necessarily result in a single alternative. Lastly, these studies model the pattern of individual choice set formation behaviors based on the assumption that consumers' choice set formation stages is constructed by a combination of individual criteria thresholds.

The probability for each individual q to reduce the choice set to S_q is equal to the probability to select a combination of thresholds that match the formation of S_q. This relation can be denoted as the following Equation **(**22**)**.

$$P_q(S_q \mid G) = P(t^*_{1q}) \cap P(t^*_{2q}) \cap .... \cap P(t^*_{kq}) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. 22}$$

$P(t^*_{kq})$ is the probability of individual q selecting threshold t^* for the k -th criterion when $k = 1, 2, ..., K$. t^*_{kq} is a linear function of individual characteristics defined by Equation (23)

$$t^*_{kq} = \alpha_{0k} + \sum_{l=1}^{L} \alpha_l kSE_{kq} + \sum_{n=L+1}^{N} \alpha_{nk} AP_{kq} + \varepsilon_{kq} = \alpha_k{}' Z_{kq} + \varepsilon_{kq} \cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. 23}$$

Here, SE_{kq} denote individual q 's socio-demographic attributes relevant to the k -th criterion, AP_{kq} denote the perception and attitude of individual q , Z_q denote the matrix consisting of all the explanatory variables, α_k denote the vector of parameters to be estimated, and ε_{kq} denote the error term. Thus, the probability of consumer q selecting threshold t^* for the criterion k depends on the dependent variable t^*_{kq} and the distribution of ε_{kq}.

The threshold t_k^* is denoted by pre-defined or observed M_k threshold categories where $m_k = 1, 2, ..., M_k$. If the threshold parameters are assumed to follow the ordered response model, then θ_k is a set of constants that define the threshold categories and holds the relation of $\theta_1 < \theta_2 < \theta_3 < .... < \theta_{m_k} < .. < \theta_{M_k}$ for every criterion k. The unobserved dependent variable is included in the observed category m when its value is between the lower θ_{m_k} and upper bound $\theta_{(m-1)_k}$. Based on the assumption that ε_{kq} is identically and independently distributed (i.i.d) across individuals, the probability of individual q to choose threshold t^* for criterion k is denoted through a ordered-probit model form in Equation (24)

$$P(\theta_{(m-1)_k} < t_{kq}^* < \theta_{mk}) = \Phi(\theta_{(m-1)_k} - \alpha_k{}' Z_{kq}) - \Phi(\theta_{m_k} - \alpha_k{}' Z_{kq}) \quad \cdots\cdots\cdots\cdots\cdots \text{Eq. 24}$$

If the independence of the error terms across criterions hold, the unconditional likelihood of individual q to select the choice set S_q after considering k criteria for m categories can be represented as the following Equation (25). (Kaplan, Bekhor, and Shiftan, 2011) Where $d_{m_k q}$ is the indicator function equal to 1 if individual q to select the m categories for k criteria, and equal to 0 otherwise. The second stage, which consist of consumer choosing the best alternative among the chosen choice set, generally follow the utility maximization model with the multinomial logit form.

$$L_q(S_q \mid G) = \prod_{k=1}^{K} \prod_{m_k=1}^{M_k} \left[P\left(\theta_{(m-1)_k} < t_{kq}^{*} < \theta_{m_k}\right)\right]^{d_{m_k q}}$$
$$= \prod_{k=1}^{K} \prod_{m_k=1}^{M_k} \left[\Phi\left(\theta_{(m-1)_k} - \alpha_k' Z_{hn}\right) - \Phi\left(\theta_{m_k} - \alpha_k' Z_{hn}\right)\right]^{d_{m_k q}} \quad \cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. 25}$$

To put it simply the estimation of the semi compensatory model starts from the assumption that choice set formation and choice among selected alternatives are distinct mental processes. Thus, the error terms of the non-compensatory choice set formation and the compensatory choice are assumed to be uncorrelated. The combined unconditional log-likelihood for a total population of Q individuals that choose their most preferred alternative i from their selected choice set S_q can simply be denoted as Equation (26). There have been various modifications to the

$$LL = \sum_{q=1}^{Q} \ln[L_q(i \mid S_q) L_q(S_q \mid G)] \quad \cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. 26}$$

2.4.3 Modeling Consumer Usage: MDCEV Model

Multiple Discrete-Continuous Extreme Value Model is also based on the random utility maximization theory. The random utility maximization theory is a well-known microeconomic theory that states that every individual act on behalf of maximizing his or her utility. As mentioned in the earlier section, a consumer following the RUM theory chooses the best alternative based on the deterministic part and the random part of the alternatives. The model was first proposed by Bhat (2005, 2008) and is a model that examine the alternative choice usage distribution. The MDCEV model is useful in analyzing the consumer's choice and distribution of usage given multiple alternatives. The model handles the endogeneity from the previous models that analyzed multiple alternatives and simultaneously analyze the alternative choice and the usage distribution. (Spissu et al. 2009)

Many previous studies have applied the MDCEV model to analyze the consumers' preference of automobile and travel distance, proving it as a valid choice model to analyze the choice of product and usage distribution. (Bhat and Sen, 2006; Ahn, Jeong, and Kim, 2008) Some studies have applied it to analyze consumer preference of new types of vehicles such as, electric and hydrogen fuel vehicles to estimate the changes in the automobile market and vehicle usage. (Shin, Hong, Jeong, and Lee, 2012; Shin, Hwang and Choi, 2019) Other literatures have applied the MDCEV model in various other research fields such as estimating the relations between travel mode and length of stay for tourists, the relations between the type of activity and duration of activity, and the household expenditures on different goods. (Pellegrini and Scagnolari, 2019; Calastri, Hess, Daly and Carrasco, 2017)

According to the MDCEV model, the utility of consumer n when he chooses J number of alternatives among K alternatives, and chooses to distribute his usage m_j to alternative j is defined by Equation (27)

$$U_n(m_1,..,m_J,0,...,0)=\sum_{j=1}^{K}\psi(x_j)(m_j+\gamma)^{\alpha_j} \quad \cdots\cdots \text{Eq. 27}$$

Where $\psi(x_j)$ is the baseline utility from choosing j. γ is a translation parameter determining the existence of interior corner solution, where $\gamma \neq 0$ means that corner solution exists with the potential of alternative j not being used at all. On the other hand, $\gamma = 0$ means that interior solution exists, meaning that all of the alternatives are being used. (Kim, Allenby, and Rossi, 2002; Bhat, 2005) α_j is a satiation parameter between the value 0 and 1, which denote the degree of diminishing marginal utility. The satiation parameter is defined by $\alpha_j = 1/(1+\exp(-\delta_j))$. Due to the identification problem, the value of γ is generally assumed to be same of all alternatives to estimate the value of γ and α_j. (Bhat, 2008) The general form of baseline utility of MDCEV model denoted in random utility form and is assumed to be positive.

$$\psi(x_j,\varepsilon_j)=\psi(x_j)e^{\varepsilon_j}=\exp(\beta' x_j+\varepsilon_j) \quad \cdots\cdots \text{Eq. 28}$$

2.5 Difference between Artificial Neural Network and Choice Modeling

With the development of various machine learning techniques, there have been many attempts to apply the new techniques to perform similar tasks previously performed by traditional statistical tools. For example, some researches have compared the performance of ANN with the traditional methods such as SVM and Logistic regression methods in classifying task. (Moraes, Valiati, Neto, 2013; Dreiseitl, Ohno-Machado, 2002; Paliwal, Kumar, 2009) With the development in the information technology and data science, the applications of ANN have been extended to variety of business areas such as accounting and auditing, finance, marketing and production. (Vellido, Lisboa, and Vaughan, 1999) However, the technique has been rarely utilized in modeling consumer decision making. (Dasgupta, Dispensa, and Ghose, 1994) Also, many studies have compared the performance between the traditional regression models with the ANN model to predict the consumer segment based on their decisions. (Fish et al., 1995; Hu et al., 1999)

Fish et al. (1995) compared the segmentation of consumer’s purchase decision by comparing the likelihood of clustering among the discriminant analysis, logistic regression and ANN models. Hu et al. (1999) demonstrated how ANN could be applied to estimate the posterior probabilities of consumer situational choices on communication channels. Study by Paliwal and Kumar (2009) compared the performance of ANN and traditional statistical methods in different fields of applications. The study conducted a meta-analysis of studies using both ANN and traditional statistical methods such as logistic regression, multinomial logit model, and other forms of advanced regression methods. The authors overviewed the application of various ANN models in the field of accounting, health, and marketing. (Paliwal and Kumar, 2009)

In the field of accounting and finance, ANN has been applied in special cases of predicting bankruptcy, credit evaluation, insolvency pre-diction, fraud detection, property evaluation, etc. Predicting these management frauds is an important issue facing the auditing profession. Fanning, Cogger, and Srivastava (1995) developed a ANN model that could discriminate management fraud based on previous data. The model performed much better than the earlier models built using logistics analysis. Lenard, Alam, and Madey (1995) compared the performance between ANN model with gradient optimized (GRG), ANN with back-propagation, and logit model in identifying the firms with risk failure. The results indicated that neural network model learned through GRG optimizer had the highest prediction accuracy among the three models.

In the field of health and medicine, there has been increasing applications of ANN to solve classification and identification problems related to epidemiological data. Study by Gaudart, Giusiano, and Huiart (2004) utilized simulated data to compare the performance of multilayer perceptron (MLP) and linear regression in predicting patients with epidemiological data. However, at the time, the results indicated no significant difference between the prediction performances of neural networks and linear regression. Nevertheless, the studies that used ANN in the field of medicine generally concluded that neural networks were excellent tool to consider the unspecified collinearity or nonlinear functions when the real underlying regression function cannot be approached by classical methods. (Delen, Walker, and Kadam, 2005; Razi and Athappily, 2005; Song, Venkatesh, Conant, Arger, and Sehgal, 2005)

As mentioned, the application of neural networks in marketing area has gained increasing interest due to their ability to capture nonlinear relationship between the variables. Some of the marketing fields where ANN have been applied include market segmentation, market response prediction, sales forecasting, and consumer choice prediction. In the field of marketing, using ANN to estimate the price elasticities as well as the effects of changes in marketing variables have increased significantly over time. To name a few, study by Dasgupta, Dispensa, and Ghose (1994) compared the performance of logistic regression model and discriminant analysis model to a back-propagation neural network model in predicting consumer segments. More specifically, the models were compared based on their ability to identify consumer segments based upon their willingness to take financial risks and to purchase a non-traditional investment product. The results showed that ANN models were better in prediction performance, but it wasn't significant enough. Such result was opposite to the studies that compared ANN with traditional statistical models in the field of finance, where neural network models have, in general, significantly outperformed traditional statistical response models. Also, the authors acknowledged that although ANN was better in predicting the outcome, it was significantly limited in providing valuable insights required in marketing studies such as identifying the effects of variable changes on the outcome or explaining the behavioral intention of the consumer. The study by Chiang, Zhang, and Zhou (2006) also used neural networks in predicting and explaining consumer behavior in using online or offline stores. Like the other studies, the authors compared the prediction performance with the logistic regression estimations. The results indicated that neural networks significantly outperformed logistic regression models for most of the surveyed products in terms of the predicting power.

The study by Thieme, Song, and Calantone (2000) developed an artificial neural network decision support system to predict complex new product development project selection decisions. Since the concept of selecting the suitable product development decision is key to market success, the authors thoroughly examined different types of ANN and traditional statistical methods to provide firm basis of the analysis. Thus, the authors used data from over 600 projects to compare traditional methods and ANN to identify successful new products. The results showed that regardless of the type of the product being developed, the neural network models demonstrated superior predictive performance compared to the traditional statistical methods. However, the authors also acknowledged that ANN provided limited interpretations to explaining the relationship between the variables.

Similar comparison has been conducted in estimating consumer's choice for travel modes. Study by Lee, Derrible, and Pereira in 2018 compared four types of ANN with the traditional traditional multinomial logit model in predicting the travel mode choice. This study argued that the field of demand forecasting has been dominated by parametric approaches, but that non-parametric approaches (such as artificial neural networks) possessed high potential to replace them. The authors acknowledged that choice problems could be adjusted to pattern recognition problems. In this case, ANN models are easily applicable with their higher capability to identify nonlinear relationships between the input data and the outputs to predict choice behaviors. (Lee, Derrible, Pereira, 2018) The authors compared the prediction ability of multinomial logit model (MNL) with four types of ANN: MLP with backpropagation, radial basis function networks (RBFNs), probabilistic neural networks (PNNs), and clustered probabilistic neural networks (CPNNs). The results showed that ANN models outperformed MNL, with significant improvement in accuracies in predicting modes.

The aforementioned studies comparing different types of ANN and the traditional statistical methods demonstrated the ability of both models. The review clearly showed that ANN has the potential to replace traditional methods in classification and prediction tasks. As many studies pointed out, the advantage of neural networks is that it can automatically approximate any nonlinear mathematical function. Unlike the traditional statistical methods, which has difficulty in explaining complex or defining the unknown relationship among variables, ANN is capable of modeling these non-linear relationships. However, as mentioned in many previous literatures, various parameters like the number of hidden layers, number of nodes in the hidden layer within the ANN does not necessarily provide meaningful interpretations to the researchers. Also, designing the optimal configuration of neural networks can become a very time-consuming process. On the other hand, traditional statistical models with parametric assumptions allow interpretation of coefficients of the individual variables and draw inferences regarding the significance of certain variables in prediction or classification problems.

To sum up, many of the studies that compared the traditional statistical methods and artificial neural network generalized the difference between the two methods as the following statement by Kumar, Rao, and Soni (1995). “The neural network approach is parsimonious, produces better classification, handles complex underlying relationships better, and is stronger at interpolation. On the other hand, the logistic regression technique has a superior solution methodology (closed form versus heuristic) and better interpretability (p. 261-262).” As stated, ANN produced better performance results than the traditional statistical models; however, they had limited interpretability and provided little room for understanding the relationship or causality among the variables. As a result, the present paper attempt to use the best of the two models by using them in their best practice. The present model uses ANN when the task is to classify key variables in affecting consumer behavior and uses traditional choice models to make important statistical inferences of the identified variables. (Kumar, Rao, and Soni, 1995)

2.6 Limitations of Previous Studies and Research Motivation

As mentioned, the development and the increasing availability of the ANN have intrigued many researchers to incorporate the computational power of ANN into the discrete choice model. However, many studies also pointed out that there were some fundamental differences between the two methodologies, stating that the focus of the two models are inherently different. Accordingly, the focus of the discrete choice model is to provide researchers with estimation results that allow various inferences to be made. Studies using the discrete choice models are studies that aim to analyze the relation that are already partially acknowledged by the researchers, and to provide implications based on statistical evidence. This also means that discrete choice models may only work when significant relationships between variables do exist. Also, many of the discrete choice models usually require certain form of data for the analysis. Data that meets the requirements will yield important results full of rich implications, but constructing the data set is considerably a difficult task.

On the other hand, artificial neural network is more flexible in the form of data usage due to its universal function approximator. By using the artificial neural network, researchers are able to analyze many forms of data traditionally considered unfit for discrete choice models. In addition, artificial neural network, like many other machine learning models, are superior in prediction tasks compared to the discrete choice model. Although the ‘black-box’ characteristics of the model limit the amount of inference researchers can make based on the predictions of artificial neural network, it is a highly efficient tool in performing classification tasks. Wang and Zhao (2018) proposed a model that uses predictions of deep neural network to enhance the interpretability of the discrete choice model prediction. The authors argued that discrete choice models are a special case of DNN that can be treated with the DNN predictions. The authors identified various tradeoffs between the prediction ability of DNN and the interpretation ability of the discrete choice to apply to making inferences from the DNN predictions. (Wang and Zhao, 2018) Thus, as mentioned in this section, various models are used to analyze different aspects of the consumer decision. However, the aforementioned studies provide limited implications due to analyzing only certain parts of consumer's decision-making process.

Based on the previous decision process theory and adoption theory, the present study propose a hybrid approach utilizing both ANN and choice models to allow efficient analysis of consumer decision and preferences. Figure (12) describes the schematic diagram of integrated process of demand forecasting of a new product combining ANN and choice models.

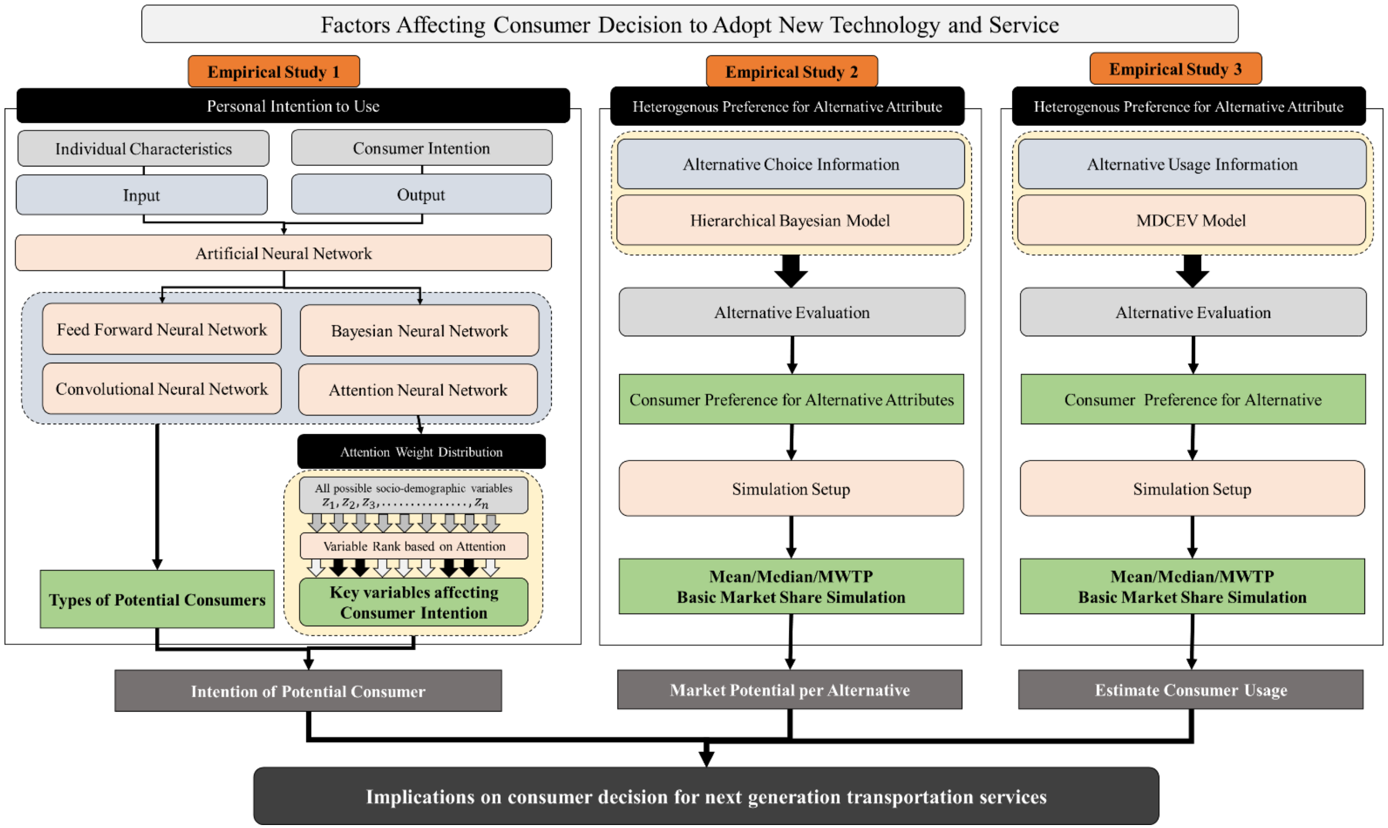


Figure 12. Schematic diagram of integrated process of demand forecasting of a new product combining ANN and choice models

The proposed research framework differentiates consumer's decision to three aspects based on the focus of the analysis, which are divided into three empirical studies. To predict the basic market size of the new services, the empirical study proposes different types of ANN that can predict the intention of the consumers by comparing the prediction performance with the traditional logistic regression model. Also, attention-based ANN model is proposed to identify key features affecting the consumer's intention to adopt new products or service. Such analysis can provide policy makers with tools that can make accurate predictions regarding the potential size of consumers for adopt new products or service and identify which factors to focus on to promote the diffusion. The second empirical study uses hierarchical Bayesian model to analyze consumer preferences towards specific attributes related to new products or service services. The analysis use key socio-demographic or consumer characteristics to identify consumer segments based on their preferences for the new products or service. The third empirical study uses the MDCEV model to provide prediction of the usage of different products after the new products enter the market. Thus, this study focuses on deriving the general implications using three different models (ANN, Hierarchical Bayesian, and MDCEV) to analyze consumer's decision-making process. In the present study, the research focus on the NGT market of Korea. The results of each empirical studies aim to propose the overall market analysis of the new transportation market in Korea by considering consumer's intention to use, preference towards different alternatives, and the distribution of usage per service when NGT are introduced to the market.

Chapter 3. Methodology

3.1 Artificial Neural Network Models for Prediction

3.1.1 Multiple Perceptron Model

The current study set up different types of ANN to find the best model that fits the data and task of the present paper, which is to estimate the consumers' intention to use NGT and identify key variables affecting that choice. The first type is the multilayer perceptron model (MLP), which has three or more layers. The MLP model consists of an input layer, one or more hidden layers, and an output layer. With an exception to the input nodes, all of the nodes within the MLP utilize the nonlinear activation function to classify the data. Most commonly used learning technique for the MLP is a supervised learning technique called backpropagation method. One of the key advantages of the MLP is that it can distinguish data that is not linearly separable. As MLP is the basic form of the ANN, the present study also used the MLP model as a base to compare the performances of different ANN models. (Rosenblatt, 1961; Rumelhart, 1986)

The first ANN proposed in the present study is a multi-layer feed-forward neural network (MLFN) model that predicts consumer's intention of using new transportation modes. The network use consumers' socio-demographic information, perception, lifestyle, attitude towards social issues and attitude towards government policies as an input to predict their intention to use new transportation mode. The model is a feed forward artificial neural network, which as mentioned earlier, an artificial neural network model consisting of interconnected neurons that receive inputs from previous neurons, process it forward if it is strong enough to produce an outcome in the last layer. One of the major benefits of the artificial neural network is its ability as a universal function approximator, or the ability to approximate any form of functions. Thus, it has been used widely in many different applications including classification, regression, clustering, and generation of images and texts. (Cybenko, 1989; Hornik et al., 1990; Kriesel, 2007; Ripley, 2007) The general design of the multi-layer neural network is depicted in the figure below.

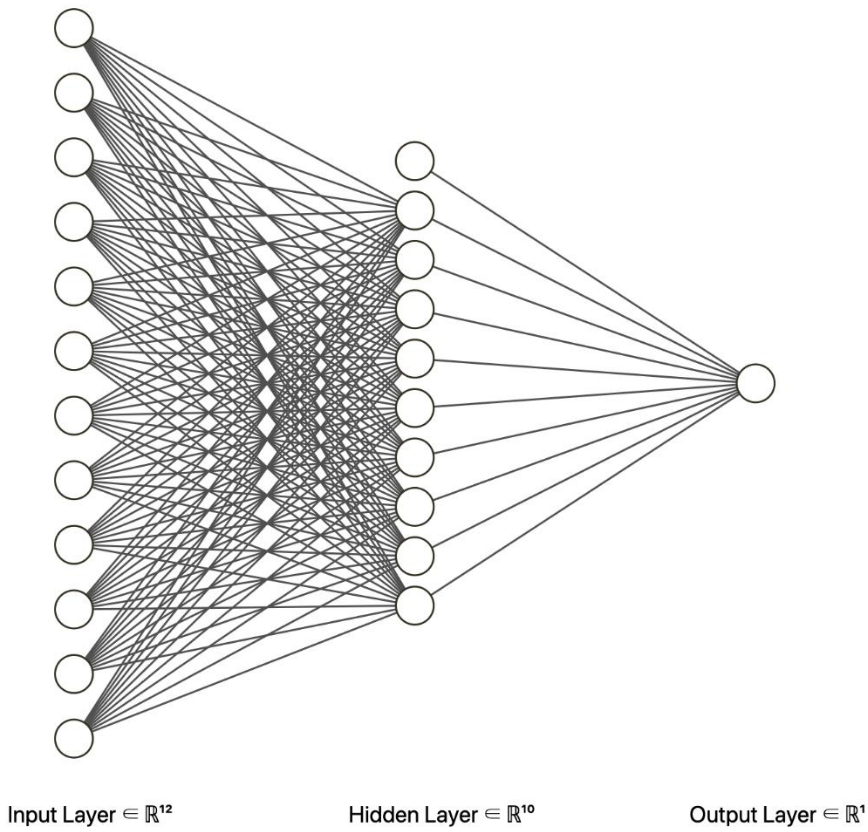


Figure 13. General form fully interconnected multi-layer neural network

The feed forward neural network in this study is constructed by one input layer, one or more hidden layers, and one output layer at the end. The structure of the model can be described as an array of $v = [v_1, ..., v_n]$ where v_i denote the size of the layer i, determined by the number of neurons in that layer. Thus, v_1 refers to the input layer, v_n refers to the output layer. The neuron o^k in layer k are connected to the neuron o^{k+1} in layer $k+1$. In the model, each neuron transforms the values sent from neurons in the previous layer with the weighted sum resulting in a temporary value t_j^{k+1} . The process can be denoted as equation below.

$$t_j^{k+1} = \sum_{i=1}^{v_k} w_{i,j}^k \times o_i^k + \theta_j^{k+1} \quad \text{Eq. 29}$$

Where $w_{i,j}^k$ is the weights of the connections nodes of the neurons i in layer k and the j neurons in layer $k+1$. θ_j^{k+1} represents the bias added to the neuron o_i^{k+1}. The weights are adjusted based on the activation functions $G(t_j)$, that which can be denoted as the equation below

$$o_j^{k+1} = G(t_j^{k+1}) \quad \text{Eq. 30}$$

In the network setup for the present study, this relation is activated by the aforementioned non-linear ReLU activation function within the hidden layers, and the last layer consist of sigmoid activation function to produce the final value at o_i^{k+1}. The ReLU activation function within the hidden layers can be described as equation below. (Hahnloser et al. 2000)

$$\mathrm{ReLU}(x) = \max(0, x) \quad \text{Eq. 31}$$

Each neuron o_i^n in the output layer denote the probability of the observation derived from the input layer being a part of the category i. Based on the sigmoid function, which (32), the probability that a certain input belongs to binary category can be denoted as the equation below.

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad \text{Eq. 32}$$

As mentioned in section 2. the values of weights and bias for each layer are derived from minimizing the loss function. The specification of the loss function completely depends on the task of the neural network. In the case of classification task of the proposed model in the present study, loss function is given by cross-entropy which is denoted by Equation (33).

$$L(\hat{p}_i, c) = -c \ln \hat{p}_i - (1-c)\ln(1-\hat{p}_i) \quad \text{Eq. 33}$$

Where c denote the indicator vector which if $c_i = 1$, then observation belongs to category 1, and if $c_i = 0$, the observation doesn't belong to category 1. $\hat{p}_i$ represents the vector of estimated probabilities. Compared to the discrete choice model, which takes into account the uncertainty and the noise within the estimation process, the ANN model is a deterministic classifier that doesn’t explicitly take the uncertainty into account. The ANN model setup and the specific data utilized in the present study will be explained in the empirical study section.

3.1.2 Convolutional Neural Network

The convolutional neural network (CNN) is another type of ANN that is commonly used form of ANN. CNN is a regularized version of MLP that considers the hierarchical pattern in data to assemble more complex patterns using smaller and simpler form. In other words, CNN simplify the complex relations within the data into a simpler form to classify the data. As a result, compared to the MLP, the scale of connectiveness and the complexity of CNN is much lower. MLP described earlier tends to overfit the data due to its full connectiveness, which is usually controlled by adding the loss function or adjusting the weights of the nodes. On the other hand, the CNN uses convolution operation by using convolutional layers and pooling layers to allow the network to be deeper with much smaller parameters. (Yann, Bottou, Bengio, Haffner, 1998)

There are some key differences between a general MLP and CNN. The Convolutional Neural Network (CNN) has the following differences compared to the existing Fully Connected Neural Network. CNN maintains the shape of input/output data of each layer. CNN can efficiently recognize characteristics of adjacent images while maintaining spatial information of the images. CNN is more efficient and accurate in Extracting and learning the features of an image/data with multiple filters. CNN contains Pooling layer to collect and enhance the characteristics of the extracted image/data. Since the filter is used as a shared parameter, there are very few parameters to train compared to a general artificial neural network. The overall process of CNN is described in the following figure.

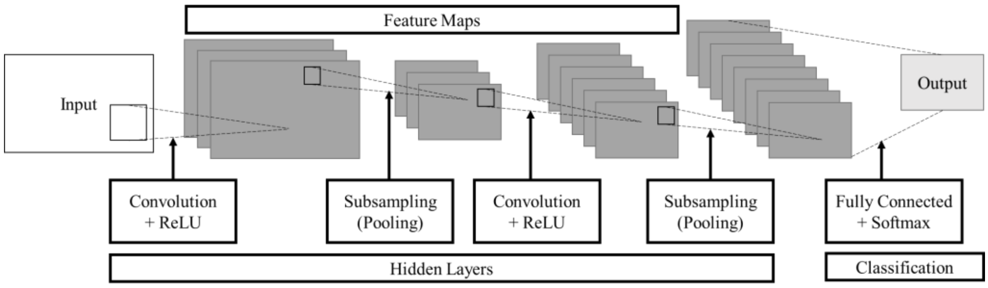


Figure 14. Overall process of CNN

A convolutional neural network contains an input layer, output layer, and multiple hidden layers. Unlike the MLP, typical hidden layer within CNN consist of a series of convolutional layers that convolve with a multiplication or other dot product. Most commonly used activation function is the ReLU function, which is usually followed by pooling layers and a fully connected layers at the end. The convolution and pooling layers are generally referred to as hidden layers because their inputs and outputs are hidden by the activation function and final convolution. Technically, the convolution layers consist of sliding dot product or cross-correlation of the inputs. This process has significance for the indices in the matrix, in that it affects how weight is determined at a specific index point.

Specifically, the convolutional layers within CNN convolve the input data and process its outputs to the next layer. Each convolutional neuron within the layer processes the data only for its receptive field. This design is more efficient than the general feed-forward network because it can deal with very large number of input data. Through this process, potential problems such as regularized weights over fewer parameters, vanishing gradient and exploding gradient shown in the backpropagation process of MLP can be avoided. (Habibi et al., 2017; Venkatesan et al., 2017; Balas et al., 2019)

Convolutional networks can include local or global pooling layers to process the underlying task. As mentioned earlier, pooling layers are able to reduce the complex dimensions of the data by merging the output of the neuron clusters at the previous layer into a single neuron. In case of global pooling, all of the neurons from the convolutional layers are processed to the pooling layer. In case of local pooling, only small clusters of the convolutional layer are processed to the pooling layer. Also, pooling can be performed based on the maximum or an average value of the outputs. Max pooling refers to pooling the maximum values of the clusters from the previous layer, while average pooling refers to pooling the average values of the clusters from the previous layer. (Dan, Meier, Masci, Gambardella, Schmidhuber, 2011; Krizhebsky, 2013; Dan, Meier, Schmidhuber, 2012)

In a fully connected neural network, each neuron receives input from every element of the previous layer. However, in CNN, neurons only receive inputs from a designated subarea of the previous layer. These subarea of input of the processed neurons is called a receptive field. Thus, receptive field is equal to the whole layer in a fully connected layer, while in CNN, the receptive field is much smaller than the entire layer. However, as the network gets deeper, meaning more layers are added to the network, the subarea of the original input in the receptive field becomes increasingly larger. Although this can make the computation of the network much higher than a general neural network, CNN simplify the task by sharing the same filter for every receptive fields. Filter in CNN are also known as kernel, which denote the vector of weights and the bias that represent particular features of the input. Sharing filter is a distinguishing feature of CNN, because it allows many neurons to use single bias and a single vector of weights across all receptive fields sharing that filter. This greatly simplify the computation compared to fully connected neural network where each receptive field have its own bias and vector weights. (Hubel, Wiesel, 1968; 2005; Fukushima, 1980).

There is a fully connected layer after the last pooling layer in CNN, which usually conduct the classification task. This layer is similar to the fully connected neural network in that it connects every neuron in the previous layer to every neuron in the next layer. In principle, it is the same as the traditional MLP network. In case of the image classification task, flattened matrix is used before the data is processed to a fully connected layer to classify the images. In present study, which deals with numbers rather than image data, such process is unnecessary.

3.1.3 Bayesian Neural Network

The Bayesian neural network (BNN) is different from the other neural networks because their weights are assigned by a probability distribution instead of a single value or point estimate. The probability distributions represent the uncertainty in the weights which can be used to express the uncertainty in the predictions. BNN is trained by variational inference to learn the parameters of the distributions instead of the weights within the network directly. (Xiang et al., 2011) This allows the network to automatically calculate the error associated with the prediction dealing with data of unknown target. Thus, BNN is useful in solving problems when there are little amount of data or there are lots of uncertainty present in the data. The main difference between the standard neural network and the Bayesian neural network is shown in Table (4), and Figure (15) denote the general concept of BNN compared to the regular MLP.

Table 4. Comparison between ANN and BNN

Standard Neural Network	Bayesian Neural Network
Parameters represented by single, fixed values (point estimates)	Parameters represented by distributions
Conventional approaches to training neural networks is proportional to approximations to the full Bayesian method	Prior distribution on the weights $P(\mathbf{w})$ and obtain the posterior $P(\mathbf{w} \mid D)$ through Bayesian learning Regularization through prior $P(\mathbf{w})$

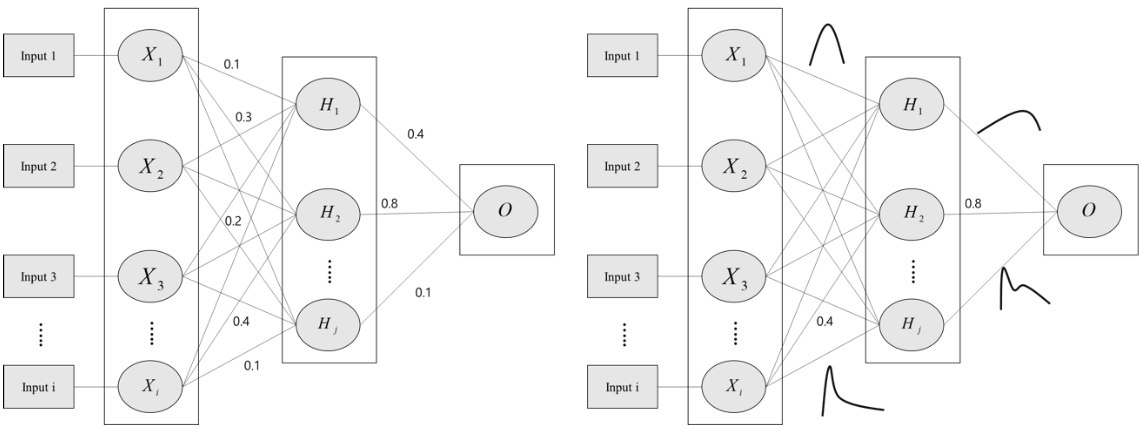


Figure 15. Deterministic FFN and distributional BNN

BNN can be viewed as a probabilistic model defined by $p(y \mid \mathbf{x}, \mathbf{w})$, where y is the set of classes and $p(y \mid \mathbf{x}, \mathbf{w})$ is the categorical distribution. The regression y is a continuous variable and $p(y \mid \mathbf{x}, \mathbf{w})$ is a Gaussian distribution. Given the training set, $D = \{\mathbf{x}^i, y^i\}$, likelihood function $p(\mathrm{D} \mid \mathbf{w}) = \prod_i p(y^i \mid \mathbf{x}^i, \mathbf{w})$ can be constructed. The likelihood function is a function of parameter $\mathbf{w}$, when maximized, gives the maximum likelihood estimate (MLE) of $\mathbf{w}$. In case of a categorical distribution, this becomes the cross-entropy error function; in case of a Gaussian distribution this is proportional to the sum of squares error function. However, simply applying MLE may lead to severe overfitting of the model. Following the Bayes theorem, the product of the likelihood defined earlier and the prior distribution $p(\mathbf{w})$ is the posterior distribution shown below.

$$p(D \mid \mathbf{w}) \propto p(D \mid \mathbf{w}) p(\mathbf{w}) \quad \text{Eq. 34}$$

Where maximizing $p(D \mid \mathbf{w}) p(\mathbf{w})$ gives the maximum posteriori estimate (MAP) of $\mathbf{w}$. Computing this posteriori regularizes the model to prevent overfitting of the model. Both MLE and MAP produce point estimate of the parameters. Using the posterior predictive distribution $p(y \mid \mathbf{x}, D) = \int p(y \mid \mathbf{x}, \mathbf{w})\, p(\mathbf{w} \mid D)\, d\mathbf{w}$ is equal to averaging the predictions from a group of neural networks weighted by the posterior probabilities of the parameters $\mathbf{w}$. The analytic solution for the posterior $p(\mathbf{w} \mid D)$ in neural network is untraceable, which makes it necessary to approximate the true posterior with a variational distribution $q(\mathbf{w} \mid \theta)$ of known function form. The parameters can be estimated by minimizing the Kullback-Leibler divergence between $q(\mathbf{w} \mid \theta)$ and the true posterior $p(\mathbf{w} \mid D)$ with respect to θ. The corresponding objective function can be expressed as the equation below.

$$F(D,\theta) = KL(q(\mathbf{w} \mid \theta) \,||\, p(\mathbf{w}) - E_{q(\mathbf{w}|\theta)} \log p(D \mid \mathbf{w}) \cdots\cdots\cdots\cdots \text{Eq. 35}$$

The second term denote the expected value of the likelihood with respect to the variational distribution and is defined as the likelihood cost. Using the KL term, the objective function can be expressed as below.

$$F(D,\theta) = \frac{1}{N}\sum_{i=1}^{N}[\log q(\mathbf{w}^{(i)} \mid \theta) - \log p(\mathbf{w}^{(i)}) - \log q(D \mid \mathbf{w}^{(\mathrm{i})}) \cdots\cdots\cdots \text{Eq. 36}$$

All of the terms in the rearranged objective function are expectation with respect to the variational distribution $q(\mathbf{w} \mid \theta)$. Thus, the objective function can be approximated by drawing samples $\mathbf{w}^{(i)}$ from $q(\mathbf{w} \mid \theta)$.

The training of BNN consist of multiple iterations of forward and backward-pass. Forward-pass denote the process of drawing a single sample from the variational posterior distribution, which is used to evaluate the objective function. The first two terms of the cost function are data-independent and can be evaluated layer-wise, the last term is data-dependent and is evaluated at the end of the forward-pass. Backward-pass denote the process of calculating gradient μ and σ through backpropagation, which is updated to optimize the model. (MacKay, 2012; Jylänki et al., 2014; Thang et al, 2015; Lobato et al., 2015)

Uncertainty related to the weights during prediction is referred to epistemic uncertainty, which can be reduced with an additional data. The epistemic uncertainty increases while training data with no or little data, and decreases while training regions with lots of data. For BNN, the epistemic uncertainty can be covered through variational posterior distribution described above. (Blundell et al., 2015; Minka, 2001)

3.2 Feature Identification Model through Attention

As mentioned in the previous studies, the present study applies artificial neural network to identify the key variables affecting the consumer choice. The present study designed a neural network that can learn the weight of each input on the produced output of the network. The model is based on the Attention mechanism within the transformer model which is one of the most recently developed form of ANN that is generally applied in dealing with natural language processing. (Polosukhin et al., 2017) The attention mechanism mimics the process of human sight mechanism into a deep learning architect. Defined by the name of the model, Attention mechanism works similar to how human identify key components when detecting and classifying different objects. When a person looks at certain objects or scene, he doesn't scan the entire scene end to end; rather he looks at a specific portion according to his needs. This can be incorporated to describe the consumer's decision-making process when making a purchase choice. As defined, human perception does not tend to process a whole information given to them, but rather focus on selective parts of the information to acquire necessary information to make a choice. Thus, the present study utilizes the attention mechanism to measure the amount of attention consumer gives to certain factors before making the decision to use NGT. Based on the attention mechanism, key variables that have the highest attention distribution are selected to check the possibility of ANN as a variable selection model in analyzing consumer choice. The attention distribution denotes the amount of weight or attention consumer give during the process of decision making according to the data.

By definition, Attention is a component within a neural network that quantify the interdependence between the input and the output, and the dependence within the input elements. There are three important components within the attention mechanism that decides the position of attention to predict certain output. These components are attention score, attention distribution, and context vector (attention output). The components are defined by the following equations.

Attention score: $e_{ij} = a(s_{i-1}, h_j)$, where e_{ij} is a scalar

Attention distribution: $a_{ij} = \dfrac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})}$, where α_{ij} is a scalar

Context: $c_i = \sum_{j=1}^{T_x} \alpha_{ij} h_j$, where c_{ij} is a scalar

The present model produces a vector for each input data, which is learned through randomly initialized parameter. Based on the form of the input, different vectors are created to produce the embedding vector. For instance, if the input consists of single response among four alternatives, the embedding vector of the chosen level of alternative used as an input vector. If the input consists of multiple responses, all of the embedded vectors are averaged to be used as an input vector. All of the embedding vectors are processed into key and value vectors.

The key, value and the query vector denote the weights of the input data, which are used to compute the context variable to produce the attention weights within the network. In short, these vectors transform the corresponding input data and the into a state vector. The query vector denotes the task at hand, which in this study is a binary classification. As the query vector is given, the network selects the key vector that match the given query vector. Based on the selected key vector, the matching value vector is selected and used in the network. The general formula for the query, key, and value vectors are defined as below.

$$Key = W_{key}e_i + b_{key}, W_{key} \in \mathbb{R}^{Dimension \times Dimension}, b_{key} \in \mathbb{R}^{Dimension} \quad \cdots\cdots \text{Eq. 37}$$

$$Value: v_i = W_{value}e_i + b_{value}, W_{value} \in \mathbb{R}^{Dimension \times Dimension}, b_{value} \in \mathbb{R}^{Dimension} \quad \cdots\cdots \text{Eq. 38}$$

Both key vectors and the given query vector are multiplied to produce the attention weight of the input data fed into the model. The produced attention weights are processed by the Softmax function and passed onto the multilayer network is denoted by the following equation.

$$\alpha_i = \mathrm{Softmax}(qk_i), \mathrm{Softmax}(x_i) = \frac{\exp(x_i)}{\sum \exp(x_j)} \quad \cdots\cdots \text{Eq. 39}$$

The derived Attention distribution α , is multiplied by the value vector of the input data through scaled-dot product to produce the weighted sum of value c. The weighted sum of value is defined as context, which can be expressed as the equation below.

$$c = \sum \alpha_i v_i \quad \text{Eq. 40}$$

By using the context as an input and the consumer's choice as an output, the multilayer network can learn the weight of the attention through backpropagation. The multilayer network is designed similar to the feedforward network described earlier. The network uses cross-entropy loss function to minimize the error in the network and apply ReLU activation function within the designated hidden layers to process the context defined by the key, value, and query vector. The cross-entropy loss function used in the current attention model is defined by the equation below.

$$\sum p_i \log \mathrm{Softmax}(o_i),\ p_i = \begin{cases} 1 & \text{if } y = i \\ 0 & \text{if } y \neq i \end{cases} \quad \text{Eq. 41}$$

The input layer of the multilayer network, where contexts are fed into the network, consit of Softmax function to distribute the weight of the context. The output layer depends on the classification task at hand. For the current paper, the output consist of binary classification. The formulation of the hidden layers and the output layer are defined by the equations below.

$$h = \mathrm{ReLU}\left(W_h v + b_h\right), W_h \in \mathbb{R}^{Dimension \times Dimension}, b_h \mathbb{R}^{Dimension} \quad \text{Eq. 42}$$

$$o = W_o h + b_o, W_h \in \mathbb{R}^{Dimension \times Dimension}, b_o \mathbb{R}^{Dimension} \quad \text{Eq. 43}$$

The overall process of the attention model incorporating the defined components are depicted in Figure (16).

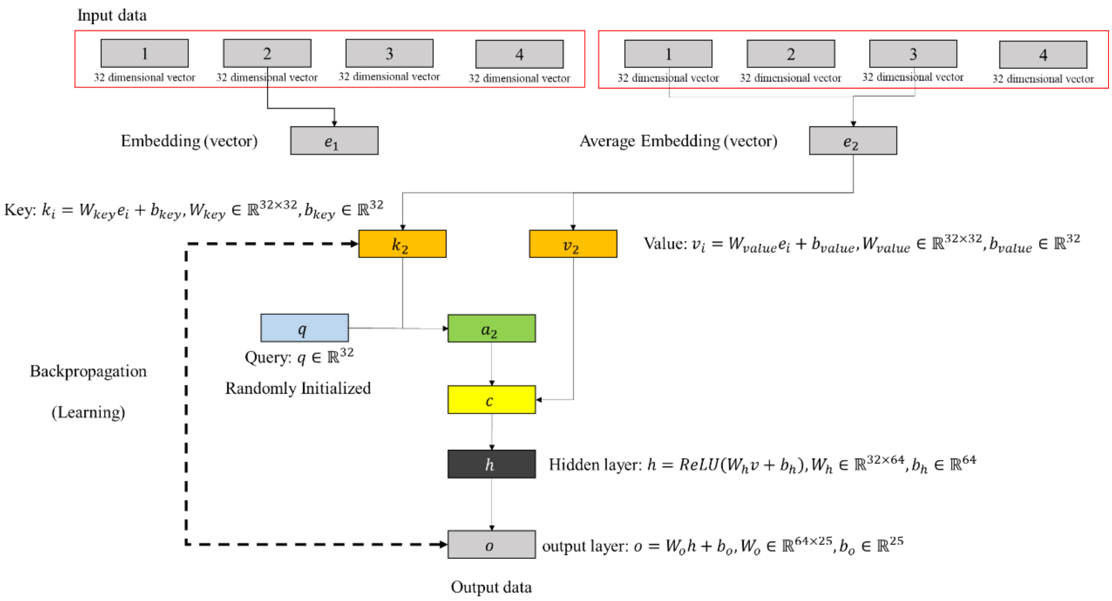


Figure 16. Procedure of attention-based ANN model

Thus, by training this attention model, the weight of the context and the input of the models can be derived. The attention value indicate how much certain information was influential in deciding the output of the model. When the input data consist of consumers characteristics and the output is their decision, the defined model can produce the amount of weight each characteristics had on the decision of the consumer.

3.3 Hierarchical Bayesian Model

The second empirical study applies the Hierarchical Bayesian model to analyze the consumer preference for different types of automobile-based transportation services, or the new types of transportation services in Korea. The hierarchical Bayesian model, proposed and developed by Allenby and Rossi in 2006, has been widely used to analyze heterogeneity of consumer preference more flexibly. The hierarchical Bayesian model is based on the multinomial logit model mentioned previously. However, unlike the multinomial logit model and similar to the mixed logit model, it incorporates heterogeneity in consumer preference. The main advantage of using the hierarchical Bayesian method is that it presents individual level parameters.

Hierarchical Bayesian model incorporates heterogeneity of consumer preference in the utility by assuming a functional form for preference parameter. According to Allenby and Rossi (2006), hierarchical Bayesian model has two levels. The higher level consists of individual's part worth, which is described by multivariate normal distribution characterized by a vector of means and a matrix of covariance. Given the individual's part worth, the lower level calculates the choice probability for certain alternative following the multinomial or other forms of logit model. Compared to the traditional multinomial logit models, hierarchical Bayesian model allows the researcher to investigate individual part worth values by incorporating heterogeneity among consumers. (Jang, 2014)

Accordingly, the probability of β_n of the multinomial logit model is depicted in equation below.

$$\Pr(j)_n = \frac{\exp(x_j\beta_n)}{\sum_k \exp(x_k\beta_n)} \cdots\cdots \text{Eq. 44}$$

In the hierarchical Bayesian model, β_n is denoted as Equation (45)

$$\beta_n = \Gamma' z_n + \zeta_n, \qquad \zeta_n \sim N(0,\Sigma) \cdots\cdots \text{Eq. 45}$$

Where z_n are covariates and Γ is the coefficient explaining the relationship between β_n and the covariates. ζ_n are random part that captures all of the unobserved heterogeneity not explained by the covariates, and is assumed to be multivariate normal distribution. (Allenby and Ginter, 1995). Covariates usually consist of consumer's socio-demographic information such as age, gender, income, perception of an alternative, attitude towards policy or products, etc.

The estimation of hierarchical Bayesian model is based on the Bayesian estimation method following the Bayes' rule. An example of a Bayes' rule is depicted in

$$P(\theta \mid Y) = \frac{P(Y) \times P(Y \mid \theta)}{P(Y)} \cdots\cdots \text{Eq. 46}$$

Here Y and θ denote the observed data and parameter to be estimated. $P(Y \mid \theta)$ represents the likelihood function of the realization probability of the data based on the prior parameter $P(\theta)$. Lastly, $P(\theta \mid Y)$ denote the posterior distribution, which is the updated prior distribution by the likelihood. If $P(Y)$ is assumed to be constant to make the posterior a probability distribution, Equation (46) can be expressed as equation below.

$$P(\theta \mid Y) \propto P(\theta) \times P(Y \mid \theta) \quad \text{Eq. 47}$$

In many cases, Markov Chain Monte Carlo (MCMC) Gibbs sampler is used in Bayesian estimation procedure by the following order. (Choi, 2009)

$$\begin{aligned} &\Gamma \mid \Sigma, \beta_n \\ &\Sigma \mid \beta_n, \Gamma \\ &\beta_n \mid \Gamma, \Sigma \end{aligned} \quad \text{Eq. 48}$$

The conditional distribution of each variable is denoted as Equation (49), where the prior distribution Γ is assumed to be normal, and the distribution of Σ is assumed to be inverse-Wishart.

$$\Gamma \mid \Sigma, \beta_n, Z \;\; \forall n = \gamma \mid \Sigma, \beta, Z \;\; \forall n \sim Normal(\gamma^*, S) \quad \text{Eq. 49}$$

where

$$\beta = \left(\beta_1', \beta_2', ..., \beta_n', ..., \beta_N'\right)'$$

$$\gamma^* = S\left(Z^{*\prime}(I \otimes \Sigma^{-1})\beta\right)$$

$$S = \left(Z^{*\prime}(I \otimes \Sigma^{-1})Z^*\right)^{-1}$$

$$Z^* = \left(Z_{l=1} \otimes I, Z_{l=2} \otimes I, ..., Z_{l=n} \otimes I\right)$$

$$\Sigma \mid \beta_n, \Gamma \ \ \forall n \sim \textit{Invert Wishart}\left(K+N, (KI+N\bar{S})/(K+N)\right)$$

Where, $\bar{S} = \left(1/N\right)\sum_n \left(\beta_n - \Gamma z_n\right)\left(\beta_n - \Gamma z_n\right)'$ and K is the number of random variables.

3.4 Multiple Discrete-Continuous Extreme Value Model

The third empirical study applies the Mixed Multiple Discrete-Continuous Extreme Value Model to analyze consumer preferences for different transportation mode and the distribution of their usage. (Bhat 2005, 2008; Ahn et al. 2008; Shin, Hong, Jeong and Lee, 2012) According to the MDCEV model, as mentioned in the previous section, the consumer chooses the best distribution of usage according to his or her preference of the alternatives in order to maximize his or her utility. The model utilized in this analysis if the model proposed by Bhat (2008). Bhat assumed that random part ε is independent and follows the type I extreme value distribution, making it possible to estimate the choice probability and usage of alternatives chosen by individuals. The utility function of the MDCEV model of the present study is as follows (40)

$$U(x) = \sum_{k=1}^{K} \frac{\gamma_k}{\alpha_k} \psi_k \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\} \cdots\cdots\cdots\cdots \text{Eq. 50}$$

$U(x)$ denote the utility, which is quasi-concave, increasing, continuously differentiable function. K denote all of the available alternatives, j denote all of the alternatives chosen by the individual, and m denote the usage of the chosen alternatives. ψ_k denotes the baseline utility representing the marginal utility when none of the alternatives are consumed. Thus, the marginal utility can be denoted as the equation below.

$$\frac{\partial U(x)}{\partial x_k} = \psi_k \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k - 1} \cdots\cdots\cdots\cdots \text{Eq. 51}$$

x_j denotes the observed characteristics of the alternative j, γ_k is the constant term that determines whether or not the interior or a corner solution is possible in the utility maximization equation and represent a satiation parameter. This parameter converts the asymptote of the indifference curve according to $(0, \ldots, 0) \rightarrow (-\gamma_1, \ldots, -\gamma_K)$. $\gamma_k > 0$ denote the case when the slope is not 0, meaning there is a chance of corner solution. In other words when $\gamma_j > 0$, the model can be set to use only j products among K goods. The parameter can also serve as a satiation parameter, where as the value of γ_k becomes larger, the absolute value of the slope of the indifference curve increases, which can be interpreted as it having a relatively strong preference (low satiation). The satiation parameter can be parameterized by the following equation.

$$\gamma_k = \exp(\mu_k) = \exp(\varphi'_k \mathrm{w}_k) \text{ .... Eq. 52}$$

Where w_k is the individual characteristic. α_k is a variable that affects the ratio of diminishing marginal utility for product k relative to the usage or investing time. As a result, large α_k denote that product k has low satiation. Where $\alpha_k = 1$ means no satiation or constant marginal utility. As $\alpha_k \to 1$, satiation becomes lower, increasing the consumption of product k, and as $\alpha_k \to 0$, satiation becomes higher, leading to linear expenditure system. When the satiation becomes too high, the utility function becomes like the equation below. When $\alpha_k \to -\infty$, it's said that there is full satiation.

$$U(x) = \sum_{k=1}^{K} \gamma_k \psi_k \ln\left(\frac{x_k}{\gamma_k} + 1\right) \text{ .... Eq. 53}$$

Since $0 \leq \alpha_k \leq 1$ by definition, it can be parameterized as the following equation.

$$\alpha_k = \frac{1}{1+\exp(-\delta_k)} = \frac{1}{1+\exp(-\theta_k y_k)} \quad \text{Eq. 54}$$

Where, y_j is the individual characteristic affecting the satiation of product j.

There are some empirical identification issues related to the function form of the utility. The exact functional mechanism of γ_k and α_k are different. While γ_k control satiation through translating, α_k denote satiation by exponentiating the consumption quantity. As a result, identifying both γ_k and α_k simultaneously is difficult, making it necessary to fix one of the parameters to estimate the other. The process of estimating begins by assuming a given value for ψ_k, and identifying the effect of γ_k and α_k through different combinations each parameters. If the estimation uses γ_k-profile, it assumes that $\alpha_k \to 0$, and estimate the best fitting value of γ_k. On the other hand, it the estimation uses α_k-profile, then it assumes $\gamma_k = 1$ to estimate the best fitting value of α_k. If the combination profile of γ_k and α_k are easily detectable, meaning one of the parameters are significantly higher or lower than the other one, using different profile estimation yields well-fitting model. However, if there are no significant difference in the combination profile of γ_k and α_k, meaning both of the parameters are large or small, there is difficulty in choosing the model that fits.

Assuming that baseline utility includes the stochastic term, the utility function can be expressed as the equation below. Where z_k denote the vector of socio-demographic variables, and ε_k denote the error term representing the effect of all unobserved characteristics of product k on the baseline utility.

$$\psi\left(z_k,\varepsilon_k\right)=\psi\left(x_k\right)\cdot e^{\varepsilon_k}=\exp(\beta' z_j+\varepsilon_j) \cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. 55}$$

Substituting the newly defined baseline utility to the utility function, Equation (53) is transformed into the equation below.

$$U\left(x\right)=\sum_k \frac{\gamma_k}{\alpha_k}\left[\exp\left(\beta' z_k+\varepsilon_k\right)\right]\cdot\left\{\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k}-1\right\} \cdots\cdots\cdots\cdots\cdots \text{Eq. 56}$$

Where the budget constraint is defined by $\sum_{k=1}^{K} e_k = E$, with E denoting the total expenditure consist of $e_k = p_k x_k$, with p_k denoting the unit price of k. Accordingly, considering the budget constraint, the utility maximization problem for the modified utility function can be estimated by applying Lagrangian method and Kuhn-Tucker condition. In addition, assuming that random part ε_j follows the Type I extreme value distribution, the probability of consumer choosing alternatives i 's from K alternative is expressed in the equation below.

$$L = \sum_k \frac{\gamma_k}{\alpha_k}\left[\exp\left(\beta' z_k + \varepsilon_k\right)\right]\left\{\left(\frac{e_k}{\gamma_k p_k}+1\right)^{\alpha_k} - 1\right\} - \lambda\left[\sum_{k=1}^{K} e_k - E\right] \cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. 57}$$

$$\begin{aligned}&\left[\frac{\exp\left(\beta' z_k + \varepsilon_k\right)}{p_k}\right]\left(\frac{e_k^*}{\gamma_k p_k}+1\right)^{\alpha_k - 1} - \lambda = 0,\ if\ e_k^* > 0,\ k = 1,2,\ldots,K,\\ &\left[\frac{\exp\left(\beta' z_k + \varepsilon_k\right)}{p_k}\right]\left(\frac{e_k^*}{\gamma_k p_k}+1\right)^{\alpha_k - 1} - \lambda < 0,\ if\ e_k^* = 0,\ k = 1,2,\ldots,K,\end{aligned} \cdots\cdots\cdots\cdots\cdots \text{Eq. 58}$$

When the value of e_k * for $K-1$ alternative is determined, expenditure on the other product is automatically calculated. Therefore, assuming product 1 is nonzero amount consumption, the utility function can be estimated as the equation below.

$$\lambda = \frac{\exp\left(\beta' z_1 + \varepsilon_1\right)}{p_1}\left(\frac{e_1^*}{\gamma_1 p_1}+1\right)^{\alpha_1 - 1} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. 59}$$

The logarithm of substituting λ from the Lagrangian to Kuhn-Tucker condition yields the following derivation.

$$\begin{aligned}
&V_k + \varepsilon_k = V_1 + \varepsilon_1 \quad if \;\; e_k^* > 0 \;\; (k = 2, 3, ..., K), \\
&V_k + \varepsilon_k < V_1 + \varepsilon_1 \quad if \;\; e_k^* = 0 \;\; (k = 2, 3, ..., K), \;\; where \\
&V_k = \beta' z_k + (\alpha_k - 1)\ln\left(\frac{e_k{}^*}{\gamma_k p_k} + 1\right) - \ln p_k \;\; (k = 1, 2, 3, ..., K)
\end{aligned} \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. 60}$$

Despite the difference in defining the utility function, the overall structure of the MDCEV model is similar to the previous model. (Bhat, 2005) The model assumes that ε_k follow a i.i.d. extreme value distribution and that ε_k and z_k are completely independent. Also, ε_k is assumed to follow a scale parameter σ where there is no price difference, $\sigma = 1$. Two different model setups are used to define the selection of alternatives according to γ_k-profile and α_k-profile. Both model setups are defined in the equation below.

$$\begin{aligned}
&V_k = \beta' z_k + (\alpha_k - 1)\ln\left(\frac{e_k{}^*}{p_k} + 1\right) - \ln p_k \quad (k = 1, 2, 3, ..., K), \; when\ the\ \alpha - profile\ is\ used,\ and \\
&V_k = \beta' z_k + \ln\left(\frac{e_k{}^*}{\gamma_k p_k} + 1\right) - \ln p_k \quad (k = 1, 2, 3, ..., K), \; when\ the\ \gamma - profile\ is\ used
\end{aligned}$$

Eq. 61

As mentioned, in case when there is no significant difference between the two-satiation parameter, they can be estimated by fixing the value of one parameter. The utility function incorporating the scaler factor is represented below.

$$\tilde{U}(x)=\sum_k \frac{\gamma_k}{\alpha_k^*}\left[\exp\left(\sigma\times\left(\beta' z_k+\varepsilon_k\right)\right)\right]\cdot\left\{\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k^*}-1\right\} \quad \text{Eq. 62}$$

Where $\alpha_k^*=\sigma\left(\alpha_k-1\right)+1$.

The KKT condition of the utility function with a scaler can be derived as the following equations.

$$\begin{aligned} V_k^*+\sigma\varepsilon_k &=V_1^*+\sigma\varepsilon_1 \quad if \;\; e_k^*>0 \;\; \left(k=2,3,...,K\right), \\ V_k^*+\sigma\varepsilon_k &<V_1^*+\sigma\varepsilon_1 \quad if \;\; e_k^*=0 \;\; \left(k=2,3,...,K\right), \end{aligned} \quad \text{Eq. 63}$$

Which can be transformed to

$$\begin{aligned} V_k^* &=\sigma\beta' z_k+\left(\alpha_k^*-1\right)\ln\left(\frac{e_k^{\ *}}{\gamma_k p_k}+1\right)-\ln p_k \;\; \left(k=1,2,3,...,K\right) \\ &=\sigma\beta' z_k+\sigma\left(\alpha_k-1\right)\ln\left(\frac{e_k^{\ *}}{\gamma_k p_k}+1\right)-\ln p_k \;\; \left(k=1,2,3,...,K\right) \end{aligned} \quad \text{Eq. 64}$$

The probability of consumption for M (>=1) products among K products in the model with fixed satiation parameter is as follows.

$$P\left(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0\right) = |J| \int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty} \left\{ \prod_{i=2}^{M} \frac{1}{\sigma} \lambda \left[\frac{V_1 - V_i + \varepsilon_1}{\sigma} \right] \right\} \times \left\{ \prod_{s=M+1}^{K} \Lambda \left[\frac{V_1 - V_s + \varepsilon_1}{\sigma} \right] \right\} \frac{1}{\sigma} \lambda \left(\frac{\varepsilon_1}{\sigma} \right) d\varepsilon_1 \quad \text{............ Eq. 65}$$

Where λ is a standard extreme value density function, and Λ is a standard extreme value cumulative distribution function. Thus, the Jacobian takes the following form

$$|J| = \left(\prod_{i=1}^{M} c_i \right) \left(\sum_{i=1}^{M} \frac{1}{c_i} \right), \ where \ c_i = \left(\frac{1 - \alpha_i}{e_i^* + \gamma_i p_i} \right) \quad \text{...................................... Eq. 66}$$

Equation (65) can be expressed as the following equation, which collapse to a multinomial logit form when $M = 1$ and there are no satiation effect.

$$P\left(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0\right) = \frac{1}{\sigma^{M-1}} \left[\prod_{i=1}^{M} c_i \right] \left[\sum_{i=1}^{M} \frac{1}{c_i} \right] \left[\frac{\prod_{i=1}^{M} e^{V_i/\sigma}}{\left(\sum_{k=1}^{K} e^{V_k/\sigma} \right)^M} \right] (M-1)! \quad \text{Eq. 67}$$

The mixed MDCEV model relieve the constraint assuming a same form of error terms. As a result, error term ε_k is divided into ς_k and η_k. ς_k follow independently, identically Gumbel distribution with scale parameter of σ. η_k is correlated between alternatives, and is assumed to be heteroscedastic scale. Thus, η_k can be defined as

$\eta = (\eta_1, \eta_2, ..., \eta_k)'$, $\eta \sim N(0, \Omega)$. Based on these assumptions, the probability of purchasing the first M products out of K products are estimated according to the following equation. Where $F(\cdot)$ is a multivariate cumulative normal distribution.

$$P\left(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0\right) = \int_\eta \frac{1}{\sigma^{M-1}} \left[\prod_{i=1}^{M} c_i\right] \left[\sum_{i=1}^{M} \frac{1}{c_i}\right] \left[\frac{\prod_{i=1}^{M} e^{(V_i+\eta_i)/\sigma}}{\left(\sum_{k=1}^{K} e^{(V_k+\eta_k)/\sigma}\right)^M}\right] (M-1)!\, dF(\eta)$$

Eq. 68

Where c_j is defined by the following Equation (69), and the deterministic part V_j are denoted with the satiation parameter as Equation (70)

$$c_j = \left(\frac{1-\alpha_i}{m_i^* + \gamma}\right) \quad \text{Eq. 69}$$

$$V_j = \beta' x_j + \ln \alpha_j + (\alpha_j - 1) \ln(m_i^* + \gamma) \quad \text{Eq. 70}$$

According to Equation (70), if the consumer is assumed to only choose one alternative, that is $M = 1,\ \alpha_j = 1,\ m_i = 0$ for $i \neq j$, the probability becomes a traditional multinomial logit model. Thus, we can confirm that MDCEV model is an extended model of the logit model that reflects the continuous variable.

The mixed MDCEV model used in the present study is an expansion of the general MDCEV model it that it reflects the heterogeneity of consumer preference in the model. Similar to the mixed logit model mentioned in section 2.1, the mixed MDCEV model assumes distribution for marginal utility parameter (β). By assuming a distribution of the parameter, the model can estimate the variance of the marginal utility, which considering a statistically significant parameter, can reflect the heterogeneity of the consumer preference. According to Equation (70) satiation parameter α and constant term γ is bounded by exponential relation, which makes it necessary to fix one parameter to estimate them. Since the present study aims to analyze consumer preference for different transportation mode, it is acceptable to assume that no mode is selected, meaning that corner solution may exist. Thus, to derive more implications, the present model assumed a fixed value of 1 for constant term α to estimate the satiation parameter γ. (Cho et al. 2016)

Although the mixed MDCEV model allow the researcher to analyze consumer preference in more detail, its probability is difficult to estimate using the classical maximum likelihood method. As a result, many previous studies have used the Bayesian estimation method to estimate the probability. There are multiple advantages of using the Bayesian estimation method compared to the maximum likelihood method. (Train and Sonnier, 2005; Train, 2009) First, Bayesian estimation method can avoid one of the weakness of the maximum likelihood method, estimating a biased maximum likelihood from setting an incorrect initial value. Also, the results from Bayesian estimation can be converted into not only classic estimation results, but their meaning can be easily compared. In addition, when complex computational difficulty exist in the estimation Bayesian estimation can achieve consistent results efficiently while using fewer constraints than the classical estimation method.[3] (Koop, 2003; Allenby and Rossi, 1998; Hurber and Train, 2001; Train, 2009) Rather than using a general random extraction, Bhat (2005) recommend using a pseudo extraction technique such as the Halton sequence. The present study follows this method to derive the posterior distribution of coefficients that maximize the likelihood, and verified the statistical significance level by using the mean and variance of the estimated posterior distribution.

[3] Details regarding the Bayesian estimation method is provided in section 2.1.4

Chapter 4. Empirical Analysis: Consumer Preference and Selection of Transportation Mode

4.1 Empirical Analysis Framework

The next series of empirical study apply the methodologies described in Chapter 3, including different types of artificial neural network and choice models to analyze the consumer decision process to adopt new products. The focus of the present study is analyzing consumer choice regarding their adoption, alternative decision, and usage of the NGT service. The empirical study is conducted according to the framework shown in Figure (17).

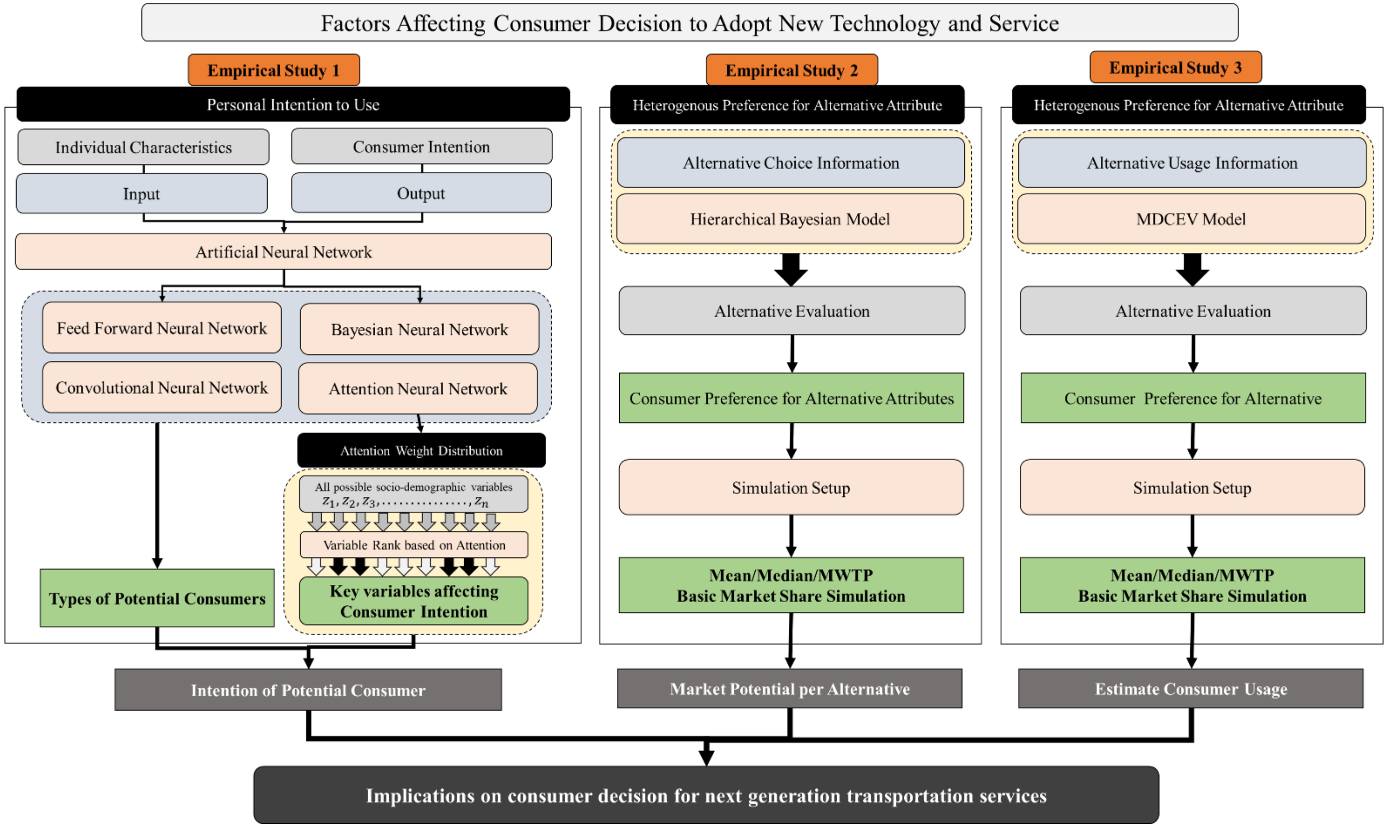


Figure 17. Empirical Framework

Accordingly, the first study aims to predict the consumer's intention of using the NGT services by using artificial neural network models proposed in chapter 3. The prediction performance of the proposed ANN model is compared to the traditional logistic regression model to validate the possibility of using ANN in consumer decision studies. In the process, the attention model is used to rank influents of all of the variable related to consumer characteristics. The ranked variables are compared to the variables selected by the traditional step-wise variable selection methods. The variables that show highest weights can be assumed to be significant factors affecting the consumer's intention to adopt NGT.

The second study aims to analyze consumer's preference towards different types of automobile-based transportation services. The analysis focuses on identifying consumers' preferred form of NGT services. Simulation analysis is conducted to analyze the effect of NGT services entering the market based on different market conditions. Implementations regarding conflicts involving the traditional transportation services and the NGT services are inferred from the analysis.

The third study aims to analyze the changes in consumer's usage of different transportation modes when the NGT services enter the market. Different schemes of NGT services entering the market are analyzed based on consumer preferences to predict the future of transportation service market.

4.2 Data

4.2.1 Overview of the Survey

Data used in the empirical studies are collected through an online survey. The survey was conducted by a professional survey company (Macromil Embrain) in Korea from December 2019 to January, 2020. The respondents consisted of 1,000 adults between the age of 20 and 59 years living in Korea. The survey respondents were chosen according to the purposive-quota sampling method based on the respondents' socio-demographic factors such as age, gender, income, and their geographical location to represent the actual population of Korea. (Teddlie and Yu, 2007; Kim et al., 2016; Moon et al., 2018) We employed Korean demographic statistics in 2019 from the Korean Statistical Information Service (http://kosis.kr), which confirmed that a component ratio of survey respondents follows a component ratio of the Korean population aged 20–59 years. In other words, the survey covers a representative sample of the Korean population. The small difference between component ratios of the sample and the population was attributable to the limitations of the online survey. If we had surveyed 1,000 people using a random sampling method, the resulting sample would not have been guaranteed to represent the Korean population. The general statistical description of the survey respondents is represented in tables.

Table 5. Socio-Demographic of Survey Respondents

Category	Description	Survey Respondent (Share)	National population (Share)
Gender	Male	512 (51.2%)	16,161,076 (51.5%)
	Female	488 (48.8%)	15,215,175 (48.5%
Age	20~29 years	220 (22.0%)	7,004,966 (22.3%)
	30~39 years	227 (22.7%)	7,446,677 (23.7%)
	40~49 years	272 (27.2%)	8,408,883 (26.8%)
	50~59 years	281 (28.1%)	8,515,725 (27.1%)
Region	Seoul	199 (19.9%)	6,134,124 (19.6%)
	Busan	63 (6.3%)	1,999,969 (6.4%)
	Daegu	47 (4.7%)	1,479,391 (4.7%)
	Incheon	59 (5.9%)	1,848,927 (5.9%)
	Gwangju	29 (2.9%)	913,158 (2.9%)
	Daejeon	29 (2.9%)	933,785 (3.0%)
	Ulsan	23 (2.3%)	726,805 (2.3%)
	Gyeongi	266 (26.6%)	8,239,788 (26.3%)
	Gangwon	29 (2.9%)	856,994 (2.7%)
	Chungbuk	30 (3.0%)	953,883 (3.0%)
	Chungnam(Sejong)	39 (3.9%)	1,456,358 (4.6%)
	Jeonbuk	32 (3.2%)	1,009,475 (3.2%)
	Jeonnam	32 (3.2%)	954,197 (3.0%)
	Gyeongbuk	47 (4.7%)	1,502,162 (4.8%)
	Gyeongnam	64 (6.4%)	1,977,700 (6.3%)
	Jeju	12 (1.2%)	389,535 (1.2%)

Other characteristics regarding the profession or occupation of the consumers and their education level were also included in survey. Table (6) denote the distribution of occupations and education level of the survey respondents.

Table 6. Occupation and Education level

Category	Description	Number	Ratio
Occupation	Self-employed	80	8%
	Sales/ service	55	5.5%
	Skilled labor	23	2.3%
	General work	19	1.9%
	Office job	510	51.0%
	Management	40	4%
	Freelancer	71	7.1%
	Housewife	91	9.1%
	Student	54	5.4%
	Unemployed	30	3.0%
	Other	27	2.7%
Education Level	Middle/high school	162	16.2%
	Junior college	140	14.0%
	College	606	60.6%
	Graduate school	92	9.2%

The first part of the survey contained questionnaires regarding consumer's own automobile, including the second vehicle if they had any. Each consumer was asked to provide specific information regarding the vehicle type, purchase type, whether it was new or used, time of purchase, the model of the vehicle, body type, fuel type, and the annual usage of the vehicle. The ratio of different types of purchase and the fuel types of vehicle are depicted in Figure (18) and Figure (19).

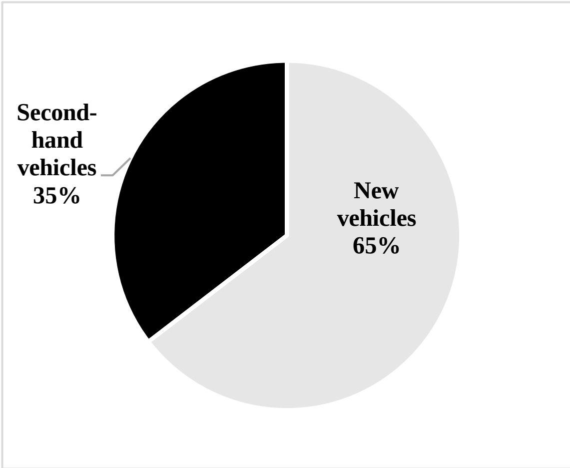


Figure 18. Type of Vehicle Purchase

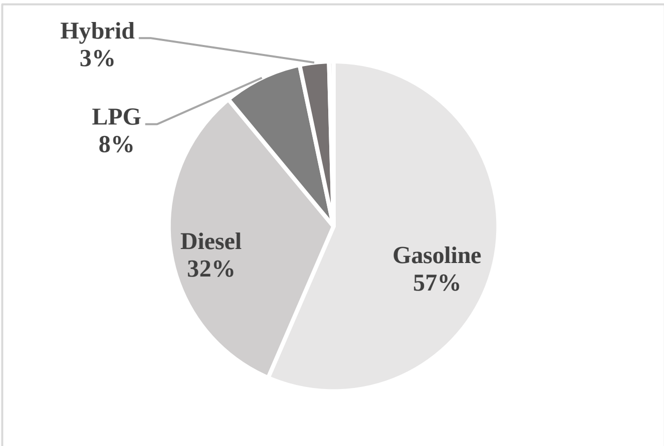


Figure 19. Fuel Type of Vehicles

All respondents owned at least one vehicle in their households. 64.6% of respondents purchased new vehicles, while 35.4% purchased second-hand vehicles. Of these vehicles, 51.3% were gasoline, 29.5% diesel, 7% LPG, 2.6% hybrid, and 0.4% electric vehicles.

In addition, consumers were required to answer multiple questionnaires regarding their recharging pattern of the vehicle and smartphone to collect information regarding the charging patterns of the consumers. Since the charging pattern of the smartphones were more likely to replicate the charging pattern of an electric vehicle it was included in the survey. (Moon et al., 2018) Next, consumer's intention of buying an environmentally friendly vehicle, such as hybrid, electric vehicle, and hydrogen fuel vehicles, as well as attributes that were important for making the decision was collected. Consumers' were given various factors such as high purchasing price, lack of charging infrastructure, short travel distance, lack of large body type, technical uncertainty, different driving experience, etc. to choose from. This information was collected to identify key factors that affected consumer's intention and their actual choice to not purchase the environmentally friendly vehicles, such may be important to analyzing the expansion of the new types of vehicles in the consumers' perspective in the future.

Of the respondents, approximately 3% owned alternative fuel vehicles, different questionnaires were given to identify the main reason to purchasing the alternative fuel vehicles. Figure (20) represent the reasons for purchasing an alternative fuel vehicle. According to the response, 28% of the respondents owning alternative fuel vehicles made a purchase to protect the environment, 24% to save fuel cost, 44% because of purchase subsidy, and 4% due to convenient charging infrastructure.

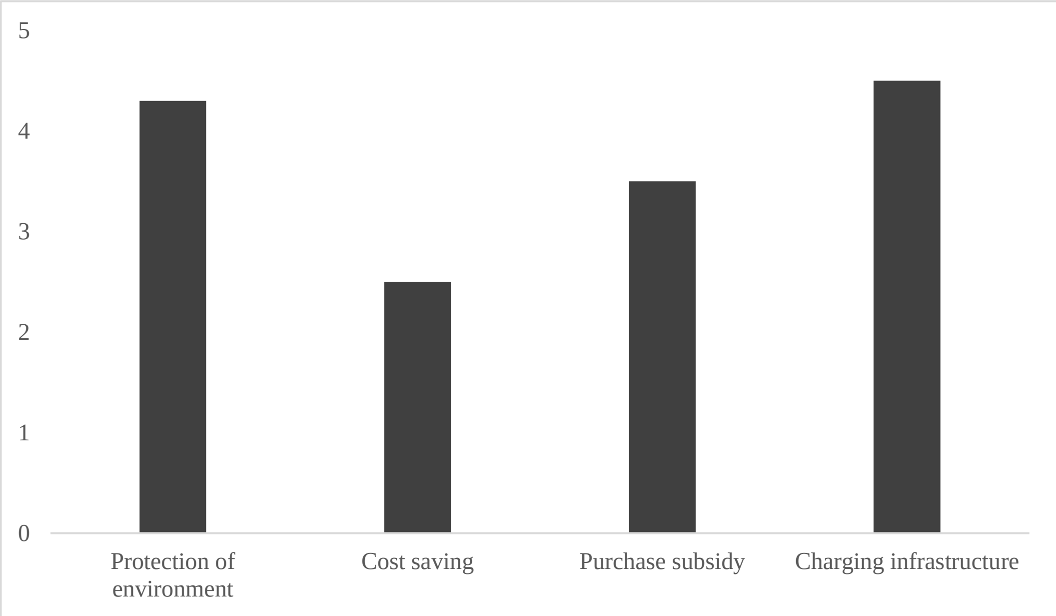


Figure 20. Reason for purchasing alternative fuel vehicle

On the other hand, all of the consumers not owning an alternative fuel were also asked to identify the reasons for not purchasing the alternative fuel vehicle. Figure (21) shows the reasons for not for not purchasing the alternative fuel vehicle. 80.2% of the respondents who did not own an alternative fuel vehicle answered that they have intention to purchase one in the future. Of the respondents who revealed no intention to purchase an alternative fuel vehicle, 74.4% responded that they will not make a purchase due to high purchase price of the vehicles, 12.2% due to lack of infrastructure, 7.8% due to short driving distance, 4.4% due to lack of vehicle classes, and 1.1% due to complicated government support process.

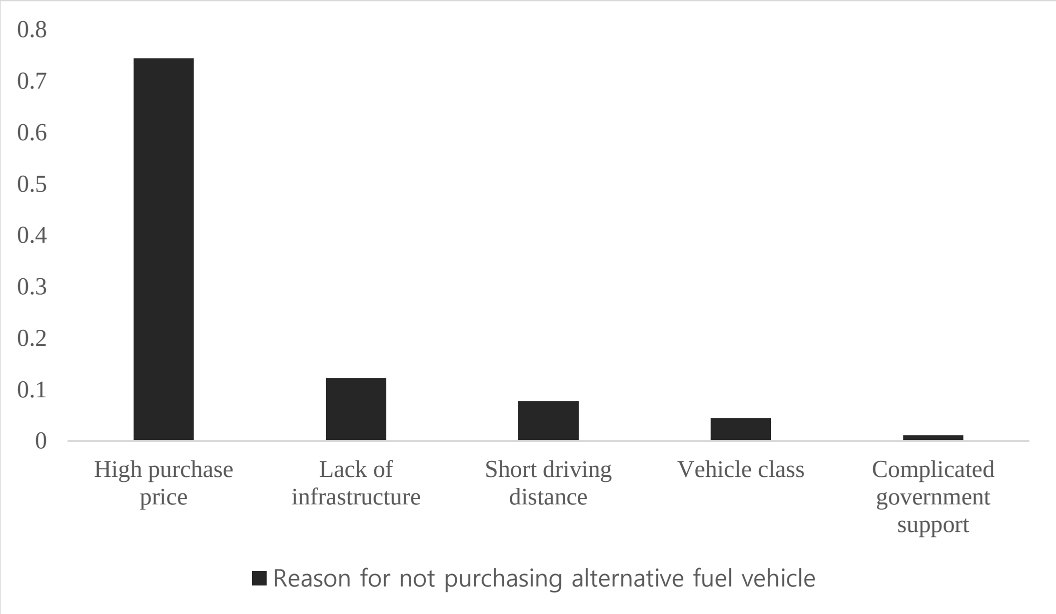


Figure 21. Reason for not purchasing alternative fuel vehicle

In addition, consumers were questioned on their general selection of transportation modes for different travel purposes. The survey questionnaires in this section focused on collecting information about the travel patterns and usage of different transportation modes. Transportation modes consist of private automobile, carpool service, car sharing service, rider sharing service, bus, subway, taxi, train, any other possible alternative. As new transportation modes are relatively unknown to the public compared to the traditional modes, additional information was provided. For more detail graphical representation and online links containing information about the actual services were provided to the respondents. As for the travel pattern, respondents were required to provide information regarding their average travel time during the day, the frequency of travel for different purposes, and how much distance they traveled in one trip. Figure (22) show the ratio of the purpose of travel for the respondents, and Figure (23) show the type of transportation modes mostly preferred by the respondents for all types of travel purpose.

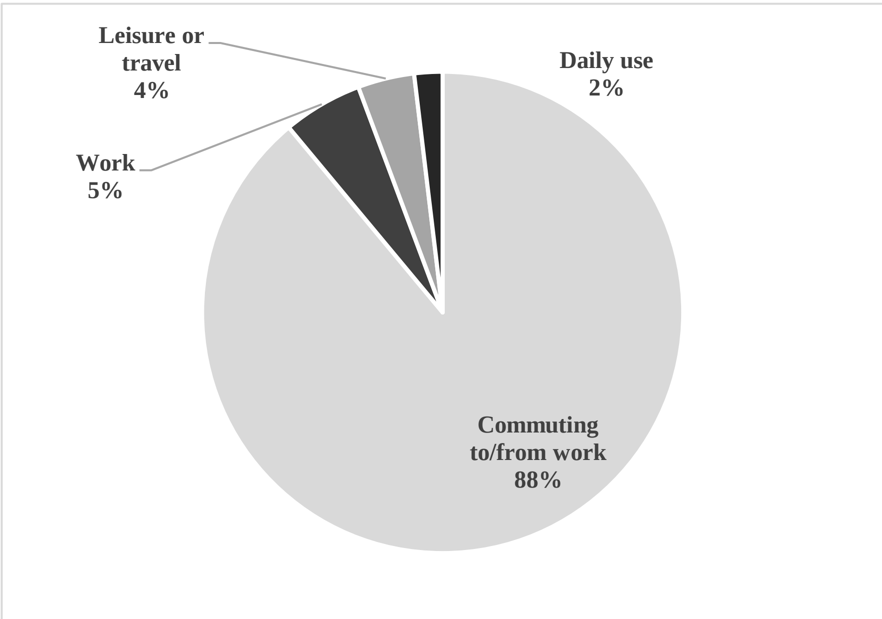


Figure 22. Purpose of Travel

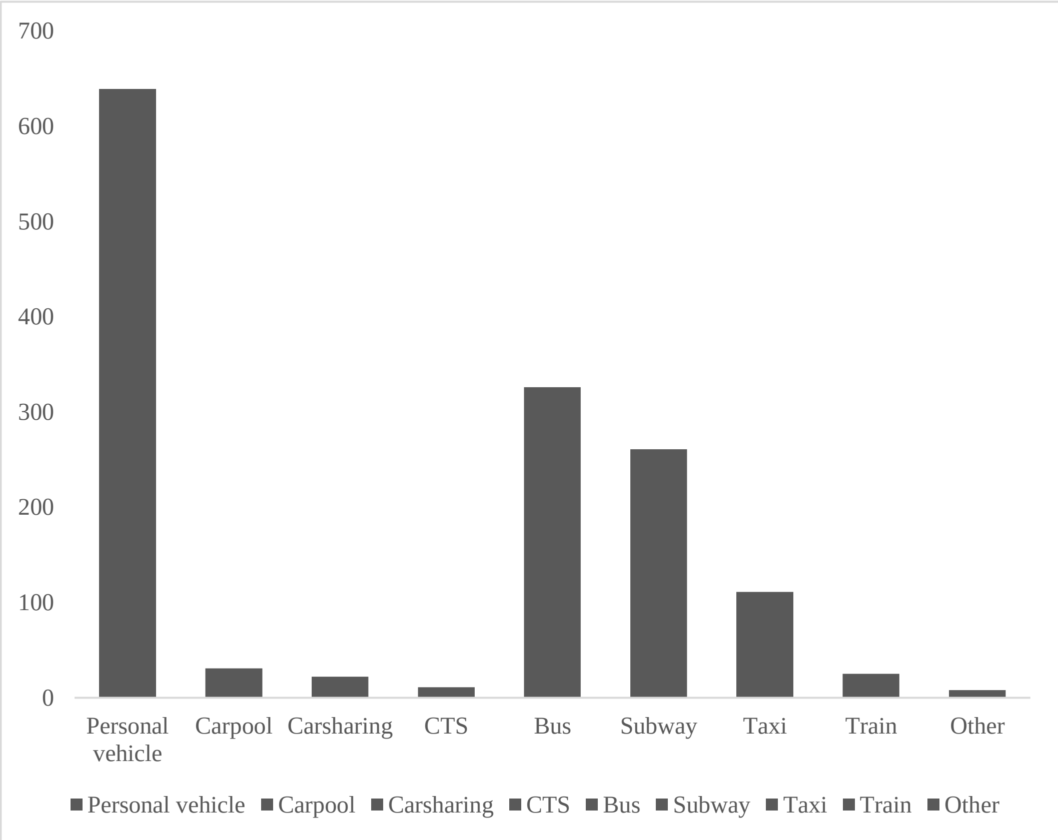


Figure 23. Favorite mode of travel

When the respondents were asked the purpose of traveling, approximately 88% of the respondents traveled to commute to and from work. 5% for work purpose, 4% for either leisure or travel, and 2% for daily purposes such as grocery shopping. When asked which type of mobility service they use, most of the respondents used personal vehicle for travel, followed by bus, subway and taxi. The results showed that not many respondents used mobility services such as carpool and car-sharing on daily basis.

The survey collected data regarding consumers' perception towards social issues related to the new transportation modes based on 5-point Likert scale, one being very pessimistic and five being very optimistic. Some of the social issues related to the new transportation modes are insurance policy for carpool service, qualification of the carpool driver, conflict between taxi industry and rider sharing service, etc. Social issues related to the new transportation modes are discussed in detail in the second empirical study. The survey also conducted a conjoint experiment regarding the types of automobile-based transportation services. Data from the conjoint experiment are used to analyze the consumer preference and choice probability among taxi, carpool, car sharing, and rider sharing services. Detailed information regarding the design of the conjoint experiment and the attributes included in the survey are discussed in the second empirical study. Lastly, the respondents were questioned on their demographic (gender, age, education level, number of family members, household income etc.), knowledge of government policies, perception of environmentally friendly vehicles, and lifestyle.

4.3 Empirical Study I: Consumer Intention Use

4.3.1 Research Motivation and Goal

The research focus of the first empirical analysis is to construct an ANN model that can accurately predict consumer's intention to use new types of transportation based on their socio-demographic data. Research field regarding consumer choice has been developed for decades and there were huge improvements in reflecting actual choice situations to estimate their preferences. However, many of the choice models require specific form of data that are hard to design. In general choice analyses, estimating consumer preference for certain product or service often requires stated preference data collected by conducting conjoint analysis designed to provide the respondents with hypothetical alternatives representing the product or service. For example, the second empirical study uses the data from the discrete choice experiment specifically designed to analyze consumer preferences on different types of automobile-based transportation service. Although the result of such analysis provides meaningful interpretations regarding consumer preferences for specific attributes of the new services, it is difficult and expensive to make such data for every analysis. On the other hand, the ANN model is excellent in making predictions with any form of data due to its universal function approximator characteristic. Thus, if the available data is limited and the prediction task is simple, it may be a more suitable to use ANN to make direct inferences. (Dreiseitl and Ohno-Machado, 2002; Abdolmaleki, Yarmohammadi, and Gity, 2004; Paliwal and Kumar, 2009)

In particular, this study tries to investigate the possibility of using ANN to predict consumer's intention to adopt new transportation services. Although many previous studies have analyzed consumers' preferences related to different aspects of transportation services, analyzing consumers' intentions to adopt particular service has been considered as more complicated. (Rezvan, Jansson, and Bodin, 2015; Wang et al., 2016; Wang and Zhao, 2017) The intention of consumer may be affected by multiple non-linear causalities such as the public's atmosphere towards the new transportation services, the level of consumer's need, past experience, and so forth. These causalities and the intention may have a linear relationship, but its highly likely that its composed of non-linear relationships, making it a complex decision to analyze using only traditional regression methods. As mentioned above, artificial neural network is a nonlinear model that can be trained to optimize almost various forms of data. (Wang, Raj, Xing, 2017) Thus, this study uses the ANN model to predict the consumer's intention, which may involve multiple causal factors.

To simplify the analysis, this study assumes that consumers who have intentions who already recognized and searched the presence of new types of transportation modes based on their intention. As described in the overview of the past technology adoption theories, consumer's decision to choose certain product and technology is affected by their intention. In addition, it can be assumed that a consumer decides to buy or use certain product and services only if he is aware of both needs and the alternatives that satisfy his particular need. The present study assumes that consumers with an intention to use new types of transportations are consumers who needs transportation services and have already searched information regarding the new services. In other words, consumer will have intention to use the new services only if he needs to travel, and if he is already aware of the new services. If a consumer doesn't need to travel, or if he is unaware of the new services due to low searching, he will have no intention to use the new services. Thus, this study tries to implement an optimized algorithm using multi-layer ANN to predict the consumer's intention. In addition, the prediction performance is compared to the traditional regression model to investigate the possibility of using ANN to predict consumer's intention based on simple information regarding consumer's socio-demographic variables and perceptions. (Lee, Kang, and Shin, 2017)

Among the traditional regression analysis models, logistic model is selected based on the used data and the form of output this analysis aims to produce. The prediction performance of the logistic regression analysis and the ANN models are analyzed by comparing the area under a ROC curve. The receiver operating characteristic (ROC) curve plots the true positive rate, defined as sensitivity, to the false positive rate, defined as specificity, for all possible cutoff values. (Altman, 1991; Swets, 2014; Krzanowski and Hand, 2009) The area under a ROC curve (AUC) is widely used to measure the accuracy of a classification test. The AUC can have a value between 0.5 (where the prediction performance is 0) and 1 (where the prediction performance is perfect) Thus, the general rule of thumb is that higher AUC values imply better prediction performance. AUC value can be interpreted as the average true positive rate, or the average sensitivity, across all possible false positive rate. Thus, this study attempts to provide quantitative evidence that ANN are suitable in analyzing social or consumer related phenomena that are non-linear in nature. In addition, by constructing ANN that accurately predicts the consumer intention, this study estimate the potential consumer type of the new services.

4.3.2 Data and Model Setup

4.3.2.1 Data Set to Predict Intention

Four different models are used to predict the adoption intention of the consumers. The models are traditional logistic, feed-forward multilayer network, convolutional neural network and Bayesian neural network. All of the models use the same input variables to compare the model performance in predicting the intention of the consumers. Table (7) denote all of the variables related to consumer characteristics included in the survey.

Table 7. Potential consumer characteristics included in analysis

Type		Variable
Input	Demographic	Age
		Gender
		Ownership of vehicle
		Household income
		Distance travelled (Monthly)
		Education
	Travel/Driving Habit	Prefer driving by self
		Travel during commute time
	Policy for NGT	Regulation towards Carpool Operation hours
		Sharing Transportation Driver regulation
		Conflicts between NGT
		Need to define service boundaries within transportation sector
		Need of government intervention for NGT
	General Policy	Belief in government policies
		Involved in government policy making
	Environment	Consider environmental effect of transport

		Prefer eco-friendly vehicle
Output	Intention	Intention to adopt NGT in the future

Four different models are used to predict the adoption intention of the consumers. The models are traditional logistic regression, feed-forward multilayer network, convolutional neural network and Bayesian neural network. As mentioned in Section 3.1. all of the later 2 ANN models are different versions of the feedforward neural network that extends the use of the multi-perceptron components. To validify the comparison, all of the models use the same input and output data to train, validate, and test the prediction analysis. When using artificial neural network, it is essential to split the data set into train, validation, and test set. Train data set is the sample of the data used to fit the model. This is data set the model sees and learns from, which denotes the process of adjusting the weights and biases within the neural network. The validation set is the sample of the data used to provide an unbiased evaluation of the model fit by the training dataset. In essence, the model uses the validation data to provide insights on how to set up the hyperparameters, such as the number of nodes, hidden layers and so forth, but the model doesn't learn or adjust the bias and weights from the validation dataset. The test dataset is used to provide an unbiased evaluation of the final model fit done by the training dataset. When the model is fully fit, the prediction or outcome based on the test set can provide significant implications regarding the task at hand. (Brownlee, 2017; James et al., 2013; Ripley, 1996)

The models use the input variables, which include basic socio-demographic variables to responses such as their awareness of government policies related to transportation market, social conflict related to the new types of transportation, perception towards driving, and involvement in government policies. The models used 80% of the data to be trained, which were modified according to the validation set, which consisted of 10% of the data. Different number of hidden layers and nodes were tested to find the best fitting hyperparameters to maximize the model fitness.

The hidden layers consisted mostly of ReLU activation function and the sigmoid (logistic) function was utilized as the output activation function. The equations of the two activation functions are based on the equations described in section 3, which are denoted below.

ReLU function: $$\mathrm{ReLU}(x) = \max\left(0, x\right) \qquad \text{Eq. 71}$$

Sigmoid function $$\sigma(x) = \frac{1}{1+e^{-x}} \qquad \text{Eq. 72}$$

Regardless of the number of layers and the amount of input data used, backpropagation algorithm is utilized to adjust the weights of the nodes to minimize the error of the model. With this model setup, consumer's intention to use the new types of transportation modes are predicted. Figure (24) depicts the general structure of the ANN models designed in this study.

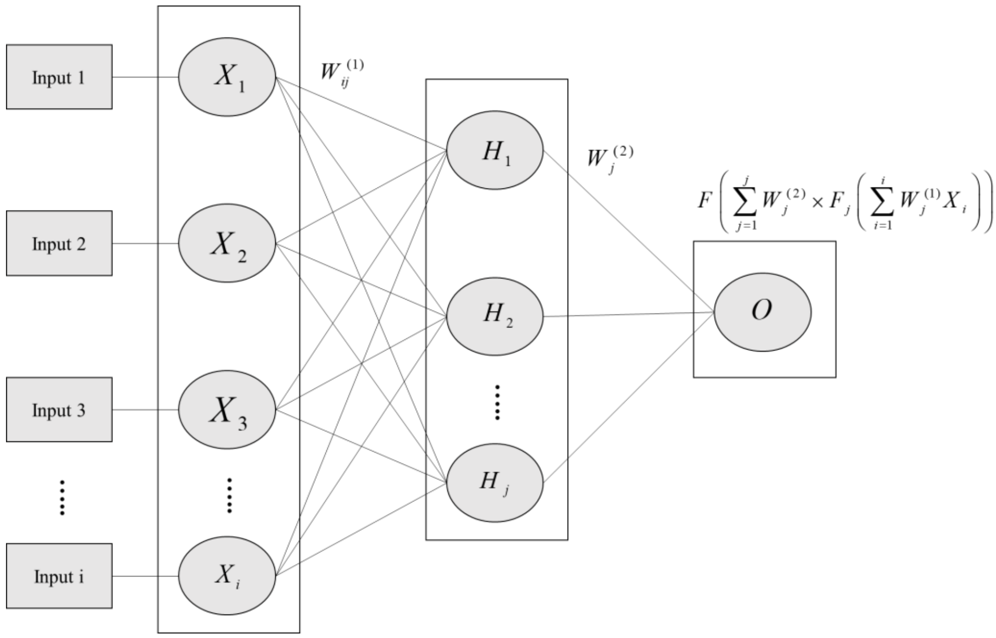


Figure 24. ANN Model Setup

The present study used the traditional logistic regression to compare the prediction performances and evaluate the possibility of using ANN to predict consumer's intention to use new types of transportation services. Logistic regression model is selected based on the utilized data and the binary classification of the dependent variable. (Dreiseitl and Ohno-Machado, 2002; Barthélemy, Dumont, and Carletti, 2018) The logistic regression measures the relationship between a categorical dependent variable and one or more independent variables by estimating the probabilities using a logistic function. The logistic function is the cumulative distribution function of the logistic distribution. The logistic regression model is defined by the equation below.

$$\log\left(\frac{p}{1-p}\right)=\beta_0+\beta_1 X_1+....+\beta_n X_n \quad \text{Eq. 73}$$

Where X_n denote the independent or predictor variables, p denotes the probability that dependent variable Y is equal to 1, and β are regression coefficients of the independent variables. The regression coefficients describe the amount of the effect of the corresponding independent variable on the outcome. The effect of the independent variables on the outcome is measured by using the odds ratio of the independent variable, which represents the factor by which the odds of an outcome change for a unit change in the independent variable. Table (8) denote all of the models used in this section.

Table 8. Models used to predict consumer intention

Model	Input	Output
Logistic Feed-Forward Network Convolutional Neural Network Bayesian Neural Network	Consumer Data defined by Table 7	Consumer intention to adopt NGT

4.3.2.1 Variable Selection Criterions

In this section, the traditional variable selection methods and developed attention model described in Chapter 3 are compared to identify key variables affecting the consumer choice. The present study conducts variable selection using the ANN-Attention model and the traditional methods of comparing the odds-ration and step-wise logistic regression to derive the best subsets of the variables that affect the intention choice of the consumers. Variable selection model is conducted to identify key variables that affect the dependent variable, which is the intention to adopt NGT in the future.

For the traditional method, comparing the odds ration based on the logistic regression and the step-wise method using different information criterions were used. While the odds ratio only compared the significance of the variables and their effect on the dependent variable, the step-wise logistic regression used different set of information criterions to find the best subset of variables. The information criterion is a function of a logistic regression model's explanatory power and complexity. There are many information criteria to evaluate the logistic regression models: Mallow's C_p, R^2_{ADJ} (adjusted), Akaike's information criterion (AIC), Akaike's corrected information criterion (AICc), and Bayesian information criterion (BIC or Schwarz criterion SBC). (Sheather, 2009; Izenman, 2008; Akaike, 1974; Hurvich and Tsai, 1989; Mallows, 1973; Schwarz 1978)

The R^2_{ADJ} is an extension of the R^2 measure of the model's explanatory power. The formulations of the R^2 and R^2_{ADJ} are shown below.

$$R^2 = 1 - \frac{RSS}{SST} \quad \text{Eq. 74}$$

$$R^2_{ADJ} = 1 - \frac{n-1}{n-k-1}\frac{RSS}{SST} \quad \text{Eq. 75}$$

Where RSS is the residual sum of squares and SST denote the total sum if squares. Note that there is an additional penalty factor for the unnecessary predictor for the R^2_{ADJ} , where n denote the sample size and k denote the number of predictors in the model. Accordingly, as R^2_{ADJ} increases, the model fit increases.

Compared to R^2_{ADJ} , the AIC is measured in the opposite direction, whereas the value of AIC decreases, the model fit increases. AIC measures the explanatory power of the model based on the maximized log likelihood of the predictor coefficients and error variance. The penalized factor is denoted by the addition of the number of predictors shown below.

$$AIC = 2\left\{-\log L\left(\hat{\beta}_0, \hat{\beta}_1, ...\hat{\beta}_p, \hat{\sigma}^2 \mid Y\right) + k + 2\right\} \cdots\cdots \text{Eq. 76}$$

Which can be expressed in terms of normal distribution likelihood as below.

$$AIC = n\log\frac{RSS}{n} + 2k + n + n\log\left(2\pi\right) \cdots\cdots \text{Eq. 77}$$

The AICC is an improved version of AIC, which corrects the bias within the AIC. AICC is a simplified version of AIC that is applied when evaluating small sample or there are large number of predictors in the analysis. The formulation of the AICC is shown below.

$$AIC_C = AIC + \frac{2(k+2)(k+3)}{n-(k+2)-1} \quad \text{Eq. 78}$$

Mallow's C_p criterion is similar to the AIC in that smaller value of the criterion is preferred to the larger one. The formulation of Mallow's C_p is shown below.

$$C_p = (n-m-1)\frac{RSS}{RSS_{FULL}} - (n-2p) \quad \text{Eq. 79}$$

Where m is the number of possible predictors without the intercepts, RSS_{FULL} is the RSS under the model containing all of the predictors, and p is defined by $p = k+1$. According to Mallow's C_p, good model would have $C_p \approx p$ and a full model will always satisfy the criterion. (Hocking, 1976; Mallows, 1973)

The BIC criterion was introduced by Schwarz and Raftery. BIC is a similar criterion to the AIC but it adjusts the penalty term to incorporate the complexity of the sample size. The definition and the simplified formulation of the BIC is shown below.

$$BIC = -2\log L\left(\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_p, \hat{\sigma}^2 \mid Y\right) + \left(k+2\right)\log n \quad \text{Eq. 80}$$

Which can be reduced into the following equation.

$$BIC = n\log\frac{RSS}{n} + k\log n + n + n\log\left(2\pi\right) \quad \text{Eq. 81}$$

There are many arguments regarding which of the information criterion to use. However, many of the studies that attempts to find the most suitable information criterion suggest that choosing a model merely on $R^2{}_{ADJ}$ may lead to overfitting. As a result, these studies propose that criterion should be chosen by minimizing the RSS for the model's predictor size k. This is the most general way for the selection process. Since all of the other terms except for predictor k are constant for the same predictor size, the best model that minimizes the RSS can be obtained.

There are many algorithms to select variables based on the information criterions described before. The present study applies two types of algorithm: forward selection and backward elimination. Although these algorithms work well with most of the criterions, Mallow's C_p criterion is an exception. All of the information criterion mentioned have an intrinsic ordering among their values: AIC is best with smallest value, $R^2{}_{ADJ}$ is best with the largest, Mallow's C_p is best when it is close to the number or predictors, and BIC is best with small value.

Forward selection is an iterative procedure, where each iteration adds a predictor to the model with only intercept terms. During the iteration process, model with the optimal predictors that derive the best information criterion value is chosen. Each criterion values are recorded when a predictor is added to the model, and if the value improves after the iteration, the process continues until there are no improvement. If there is no predictor that improves the information criterion value than the previous iteration, the algorithm stops and converge to the model if the previous iteration.

Backward elimination is also an iterative procedure similar to the forward selection. However, unlike the forward selection which starts from only intercept term, backward elimination starts with the model including all of the predictors. For each iterative process, certain predictors are excluded from the model and the information criterion values are recorded accordingly. Iteration process of excluding predictor continue until there are no improvements in the criterion value. Based on the values of all of the possible combinations, the set of predictors that yield the best criterion value is selected. (Lindsey and Sheather, 2010) For the present paper, three information criterions, AIC, AICC, and BIC will be used to test the model fit for the logistic regression.

4.3.3 Result and Discussion

4.3.3.1 Intention Prediction Result

This section predicts the intention of consumers to use NGT services through different types of ANN models and the logistic regression model. The comparison results identify the advantages and disadvantages of the ANN incorporating non-linear form to the traditional logistic regression model which assumes linear relationship. Table (9) denote all of the variables related to consumer characteristics included in the survey.

Table 9. Consumer characteristics included in analysis

Type		Variable
Input	Demographic	Age
		Gender
		Ownership of vehicle
		Household income
		Distance travelled (Monthly)
		Education
	Travel/Driving Habit	Prefer driving by self
		Travel during commute time
	Policy for NGT	Regulation towards Carpool Operation hours
		Sharing Transportation Driver regulation
		Conflicts between NGT
		Need to define service boundaries within transportation sector
		Need of government intervention for NGT
	General Policy	Belief in government policies
		Involved in government policy making
	Environment	Consider environmental effect of transport
		Prefer eco-friendly vehicle
Output	Intention	Intention to adopt NGT in the future

All of the variables presented in Table (9) were used as an input and output for the four models. These data is used to train and test the different types of ANN models. Among the survey data, 80% are randomly selected as training data for the model estimation and learning. Remaining 10% of the (80 respondents) data are used as the validation set and the last 10% of the data are used as a test data to validate the prediction capability of the model. The data are all normalized by embedding it to 64 vectors. The testing process was conducted 50 times and the average result is reported. For the ANN model, the current analysis designed the general feed-forward multilayer perceptron model (FFN), convolutional neural network (CNN), and the Bayesian neural network model to compare the prediction ability among the models.

The train loss represents the error in the training set of data, which is calculated by the cross-entropy loss function in the models. The loss is used the calculate the gradients, which adjusts the weights of the neural networks to train the model. As training proceeds, the loss is minimized, and the accuracy of the train set increases. However, increasing the training process may result in overfitting of the model. Overfitting occurs when the training process of neural network is stretched too long, resulting in low loss and high training accuracy, but huge error in predicting the testing set. Compared to the discrete choice model, which takes into account the uncertainty and the noise within the estimation process, the ANN model is a deterministic classifier that doesn't consider the uncertainty into account. The settings of the models were determined to avoid the overfitting of the models. With this in account, Table (10) denote the results of the ANN models of the study. Figure (25) ~ Figure (27) denote the ROC curve of the models and Figure (28) ~ Figure (30) denote the loss of models.

Table 10. Result of ANN models

Model	Train Loss	Train Accuracy	Test Accuracy	AUC
Logistic	0.0076	86.62%	81.00%	0.6631
FFN	0.0003	99.88%	93.00%	0.7855
CNN	0.0019	96.50%	89.00%	0.7996
BNN	0.0008	98.50%	96.00%	0.8475

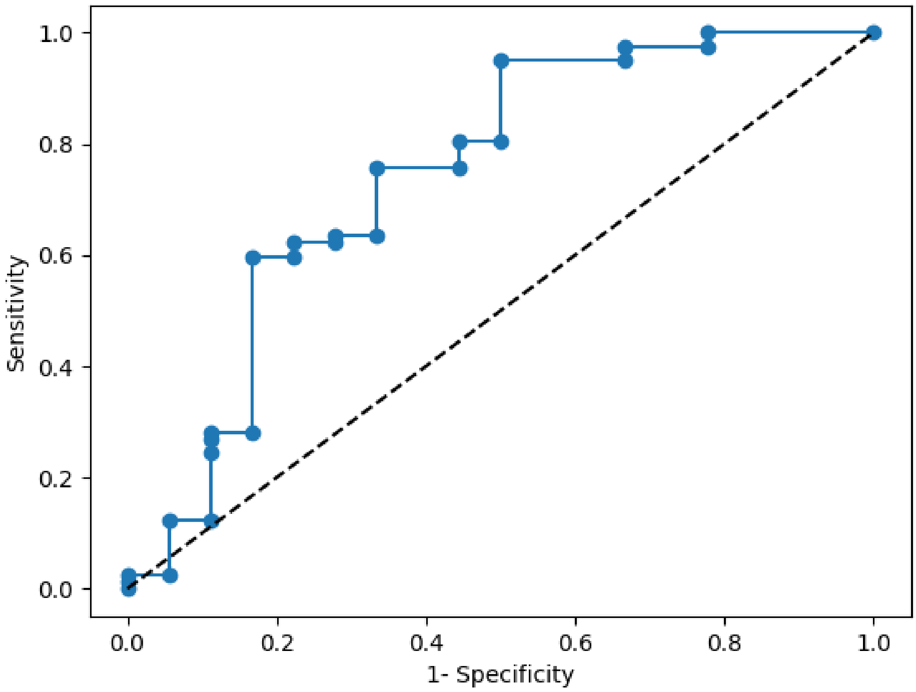


Figure 25. ROC curve of FFN

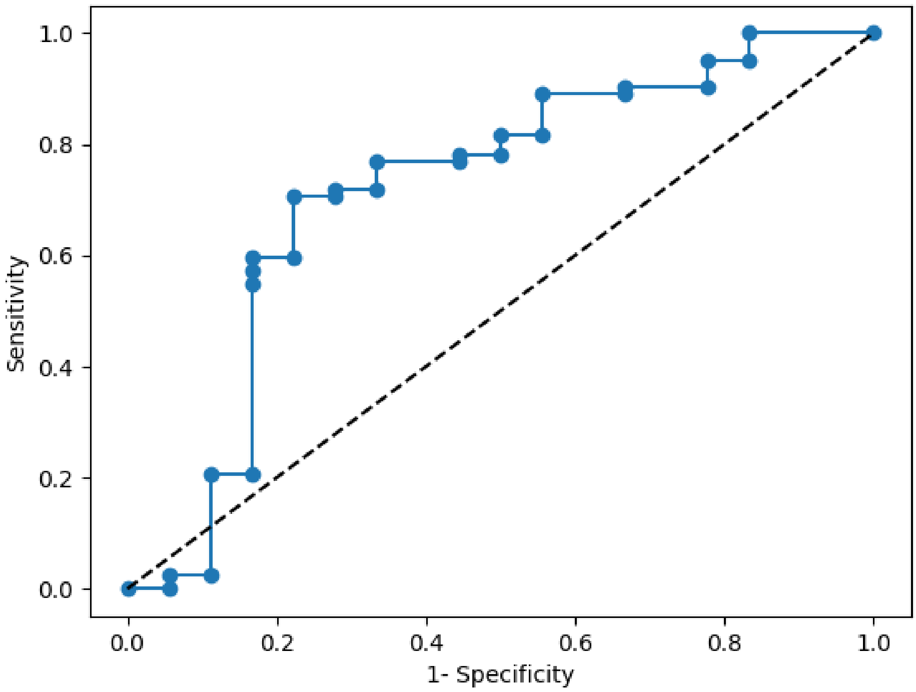


Figure 26. ROC curve of CNN

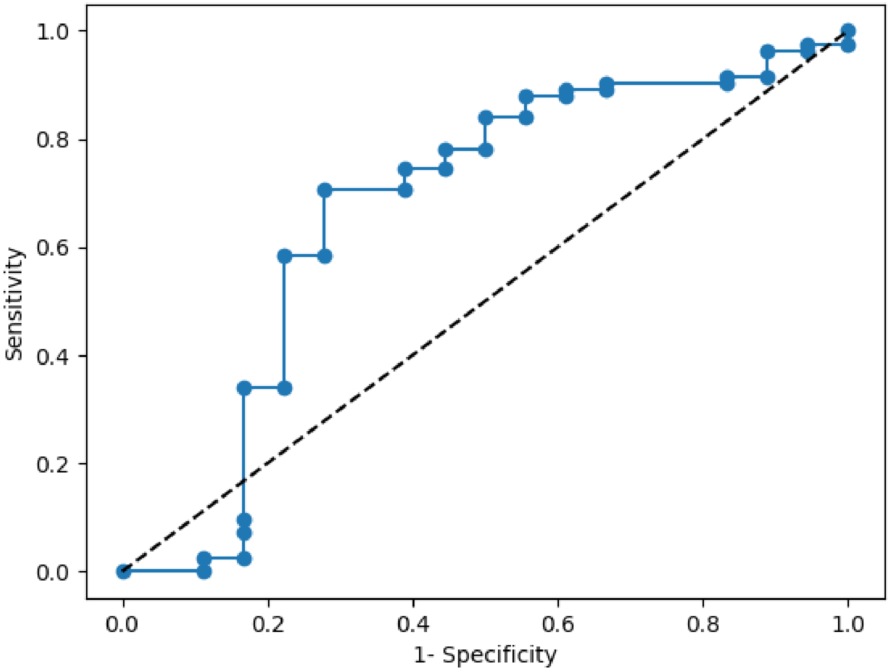


Figure 27. ROC curve of BNN

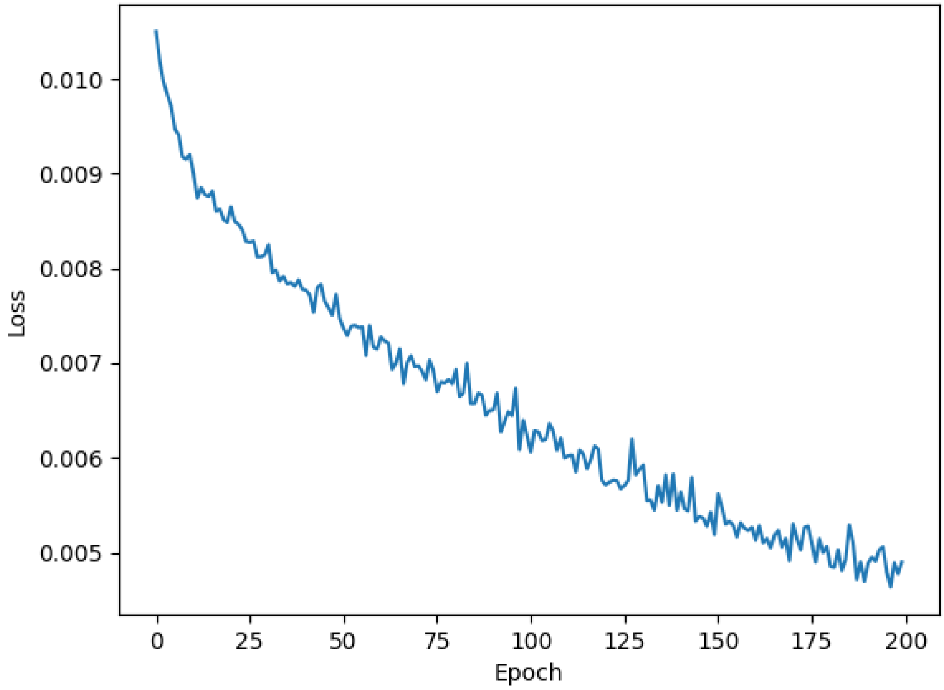


Figure 28. Loss function of FFN

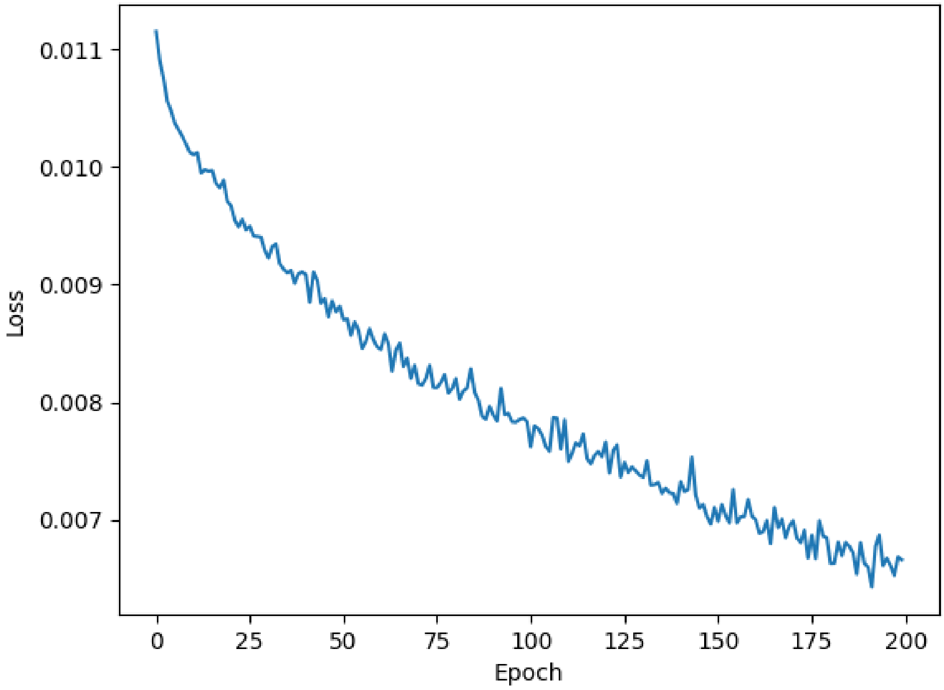


Figure 29. Loss function of CNN

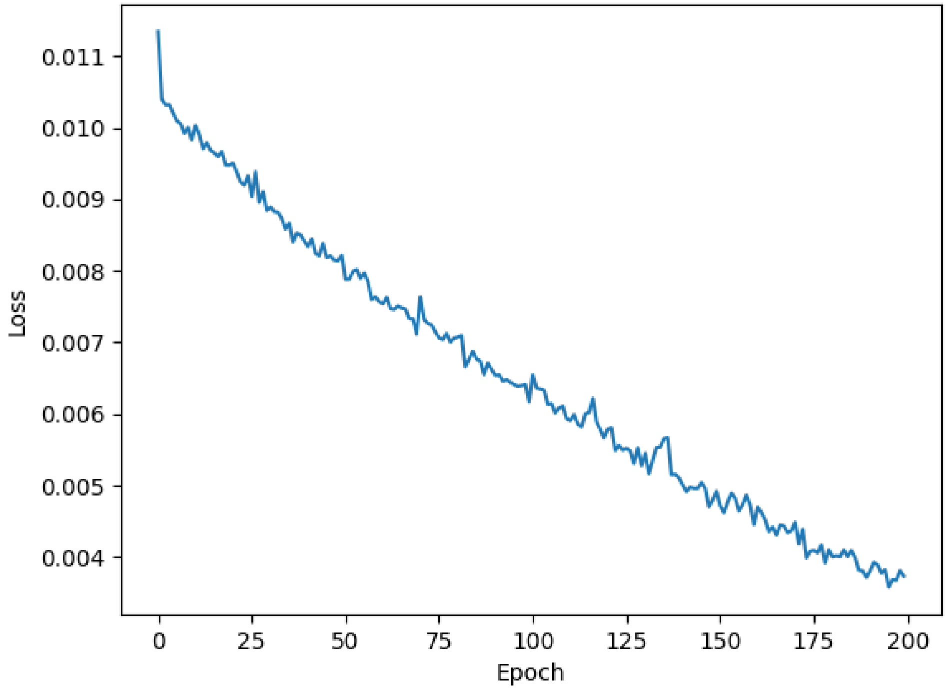


Figure 30. Loss function of BNN

All of the models yielded better AUC value to the AUC value of the logistic regression, showing that they had similar or better prediction performance with the traditional regression models. In the training process of all of the ANN models, increasing epoch and neurons of the layers resulted in higher training predictions but at one point the AUC and the test accuracy of the model decreased. This highlights the importance of avoiding the overfitting problem. If the size of the training data is increased and the number of learning epochs is increased, the prediction rate may increase, but it is still up to the researchers to identify the right settings of the model for the task at hand. (Lee, Kang, and Shin, 2017)

For this study, AUC values of all of the ANN model were superior to the logistic regression models. This indirectly indicates that ANN and other forms of deep learning may also be effective in predicting consumer intentions of using the new transportation types based on simple socio-demographic data. The BNN model showed the highest AUC value and the prediction performance. To summarize, the comparison of ANN and logistic regression show that all types of ANN models can produce similar or better AUC values to the logistic model. Comparison between the logistic regression with more independent variables may show that ANN is more than efficient and capable of modeling different types of data to produce better prediction performance.

4.3.3.1 Feature Identification Result

This section identifies the advantages and disadvantages of the ANN model to select the best fitting variables for the analysis. To compare the model fit of variables selected by the attention-based ANN model and the variables selected based on previous literatures and researcher intuitions, step-wise selection method is conducted. The chosen variables though the ANN model and traditional methods are both locked in during the step-wise method to provide insights on which method is more applicable in variable selection. The advantages of applying the attention-based ANN for pre-screening all of the potential variables are provided by comparing the model fit of the logistic regression of variables selected by step-wise method through two models.

First, the following variables are selected through reviewing numerous previous literatures regarding consumer preference for transportation modes and new technology. (Moon et al., 2017; Lavieri, et al. 2017; Abraham et al., 2016; Egbue and Long, 2012; Ewing and Sarigöllü, 2000; Hackbarth and Madlener, 2013; Train and Winston, 2007)

Table 11. Consumer characteristics included in analysis

Type		Variable
Input	Demographic	Age
		Gender
		Vehicle Ownership
		Household income
		Distance travelled (Monthly)
	Travel/Driving Habit	Travel during commute time
Output	Intention	Intention to adopt NGT in the future

Since NGT consist of all different transportation modes, it is adequate to choose the variables selected by the previous literatures. Next, the input variables are processed through attention-based ANN model to identify the rankings of the variables according to their weights. Table (12) denote the result of weights for every variable according to the attention-based ANN model.

Table 12. Attention weights of variables

Rank	Variables	Attention Weight
1	**Vehicle Ownership**	0.0701

2	**Distance Traveled**	0.0665
3	**Participate in Policy Making**	0.0641
4	**NGT Policy Necessary**	0.0633
5	**Household Income**	0.0631
6	**Carpool Regulation**	0.0625
7	NGT Issue	0.0599
8	Trust in Government	0.0588
9	Personal Driving	0.0582
10	Pro-Ecovehicle	0.0578
11	Education	0.0577
12	Pro-environment	0.0567
13	Age	0.0558
14	Driver Regulation	0.0549
15	Transportation Category Necessary	0.0508
16	Commute Time	0.0506
17	Gender	0.0504

Based on the attention weights, the top six variables that had the most attention for the consumer's intention to adopt NGT were Vehicle Ownership, Distance Traveled, Participate in Policy Making, NGT Policy Necessary, Household Income, and Carpool Regulation. Two different logistic regressions based on the variables chosen by different methods were conducted. According to the step-wise regression of intuitive variables, the principle socio-demographic variables shown in many of the previous literatures were chosen to be included in order. The first inclusion variable specified is the gender variable, followed by age, and vehicle ownership. (Moon et al., 2017; Lavieri, et al. 2017; Abraham et al., 2016; Egbue and Long, 2012; Ewing and Sarigöllü, 2000; Hackbarth and Madlener, 2013; Train and Winston, 2007) The following table denote the step-wise procedure based on the inclusion variables. The stepwise logistic regression is conducted by StepReG package in R. (Li, 2020)

Table 13. Step-Wise of Intuitive derived features (AIC Criterion)

Model # of variables	Variables Included	Variables added[4]	Information Criterion AIC
1 (6)	Gender	Participate in Policy Making Distance Traveled Driver Regulation NGT Issue Household Income	**598.59**
2 (7)	Gender Age	Participate in Policy Making Distance Traveled Driver Regulation NGT Issue Household Income	599.90
3 (8)	Gender Age Vehicle Ownership	Participate in Policy Making Distance Traveled Driver Regulation NGT Issue Household Income	600.58
4 (6)	Gender Age Vehicle Ownership Household Income Distance Travelled Commute Time	Participate in Policy Making Driver Regulation NGT Issue	602.58

[4] The order of the variables added

Table 14. Step-Wise of Intuitive derived features (BIC Criterion)

Model # of variables	Variables Included	Variables added	Information Criterion BIC
1 **(3)**	Gender	Participate in Policy Making Distance Traveled	**620.77**
2 **(4)**	Gender Age	Participate in Policy Making Distance Traveled	633.54
3 **(5)**	Gender Age Vehicle Ownership	Participate in Policy Making Distance Traveled	637.40
4 **(6)**	Gender Age Vehicle Ownership Household Income Distance Travelled Commute Time	Participate in Policy Making	644.30

The result showed that the stepwise procedure following both AIC and BIC criterion added Participate in Policy Making and Distance Traveled in all cases. No matter what variable was set as the initial inclusion, variable, the best fitting models in every case included Participate in Policy Making and Distance Traveled variables.

According to the step-wise regression of ANN-derived variables, including the variables with the highest weights yielded the following variables and their information criterion values.

Table 15. Step-Wise of ANN derived features (AIC criterion)

Model # of variables	Variables Included	Variables added	Information Criterion AIC
1 (6)	Vehicle Ownership	Participate in Policy Making Distance Traveled NGT Issue Household Income Driver Regulation	**598.15**
2 (6)	Vehicle Ownership Distance Traveled	Participate in Policy Making NGT Issue Household Income Driver Regulation	598.15
3 (6)	Vehicle Ownership Distance Traveled Participate in Policy Making	NGT Issue Household Income Driver Regulation	598.15
4 (6)	Vehicle Ownership Distance Traveled Participate in Policy Making Household Income Driver Regulation	NGT Issue	601.22

Table 16. Step-Wise of ANN derived features (BIC criterion)

Model # of variables	Variables Included	Variables added	Information Criterion BIC
1 (3)	Vehicle Ownership	Participate in Policy Making Distance Traveled	**620.56**
2 (3)	Vehicle Ownership Distance Traveled	Participate in Policy Making	620.56
3 (3)	Vehicle Ownership Distance Traveled Participate in Policy Making	None	620.56
4 **(6)**	Vehicle Ownership Distance Traveled Participate in Policy Making Household Income Driver Regulation NGT Policy	None	638.72

The stepwise regression added variables with the top attention weights to the model excluding the NGT issue variable. The process was recorded until top ten variables were included in the model. The step-wise procedure converged after top six variables were added to the model, indicating no further improvement in model fit occurred after total of 6 variables were included in the model. The top six variables are Vehicle Ownership, Participate in Policy Making, Distance Traveled, NGT Issue, Household Income, and Driver Regulation. Among the added variables, 4 of the variables excluding the first included variables were amongst the top ten variables with the most attention weights. Table (17) and Table (18) denote the logistic regression and their model fitness. The logistic regression is conducted by STATA/SE 13.1 program using Logit command, which is based on maximum likelihood estimator method.

Table 17. Logistic regression analysis result (Based on previous studies and intuition)

Independent Var.[5]	Coef.	Std. Err.	z	P>z	95% Confidence Interval	
Male	-0.316	0.229	-1.380	0.168	-0.765	0.133
Participate in Policy Making	-0.501***	0.128	-3.910	0.000	-0.752	-0.250
Distance Traveled	-0.204***	0.054	-3.800	0.000	-0.309	-0.099
Driver Regulation	-0.209*	0.113	-1.850	0.064	-0.429	0.012
Household Income	-0.224*	0.121	-1.850	0.064	-0.462	0.013
NGT Issue	-0.135	0.092	-1.480	0.139	-0.315	0.044
Constant	2.296***	0.745	3.080	0.002	0.836	3.757
Log-Likelihood = -291.9502						
AIC = 598.599				BIC = 632.953		

*** 99%; ** 95%; * 90% significance level

[5] Dependent variable is the consumer intention to use new types of transportation service

Table 18. Logistic regression analysis result (Based on Attention-based ANN)

Independent Var.[6]	Coef.	Std. Err.	z	P>z	95% Confidence Interval	
Vehicle Ownership	0.477	0.401	1.190	0.234	-0.309	1.263
Participate in Policy Making	-0.506***	0.128	-3.970	0.000	-0.756	-0.256
Distance Traveled	-0.254***	0.062	-4.120	0.000	-0.375	-0.133
NGT Issue	-0.213*	0.120	-1.780	0.075	-0.447	0.021
Household Income	-0.155	0.095	-1.630	0.102	-0.341	0.031
Driver Regulation	-0.185*	0.111	-1.660	0.096	-0.403	0.033
Constant	1.946***	0.744	2.620	0.009	0.489	3.404
Log-Likelihood = -294.072						
AIC = 598.156				BIC = 632.51		

*** 99%; ** 95%; * 90% significance level

As shown in the comparison, the model fitness of two models are almost identical. However, the step-wise model including the variable derived from the attention-based ANN is marginally better. Also, the variables included in both of the models consist mainly of variables that had high attention weights. In the model with intuitive variable as the initial inclusion variable, all of the models added the Participate in Policy Making and Distance Traveled variable, which were among top 3 variables with the highest attention. Also, in the model with ANN-derived variables, step-wise model derived the best model fit with six variables, where five variables consisted of variables with high attention weights. Thus, at least in this case, this shows that attention-based ANN model can be used

[6] Dependent variable is the consumer intention to use new types of transportation service

as a prescreening tool when there are large number of variables or features to choose from. Although using intuitive variables yield similar model and fitness, choosing such variables may be difficult when the subject of the analyzing case carry lots of uncertainty. For example, when there are no previous literatures of reference to select the intuitive variables as shown in the study of transportation modes, the weights of the variables derived by the proposed model can serve as a guide in selecting significant features.

In addition, the proposed model can be used as a screening method to make the process of designing survey more efficient. More specifically, when designing a discrete choice experiment, it is essential to include the key features that effectively represent the subject of the study to the respondents. For example, using insignificant or inappropriate variables in the conjoint analysis can yield results with greatly misguiding inferences. As mentioned earlier, the main purpose of consumer studies using choice models are to make accurate inferences adequate enough to derive marketing and policy implications. Thus, including the right features within the survey data or selecting the adequate data is essential to conducting a consumer study. Studies similar to the current study, where the subject of the focus have been studied for in the past and a firm theoretical background regarding the necessary features may not need such screening method. However, when the subject of the focus as relatively new or contain hundreds of potential variables to choose from, the attention-based ANN model can be very useful. Since the current study only provides one case, which doesn’t necessarily fit the described cases, further studies should be conducted to test the true ability of the proposed model.

All in all, the comparison shows that step-wise logit regression using the variables selected by attention-based ANN model show better model fitness according to the AIC and BIC criterion. Although limited to this case, the result shows that attention model is capable of selecting variables that are more fitting with higher prediction power than the traditional variable selection model. However, as mentioned, further research must be conducted to validate the use of attention model in variable selection as a whole. As the current analysis only conducts step-wise selection method among various variable selection methods, it cannot be said that these variables derive the best outputs.

4.4 Empirical Study II: Heterogenous Preference for Alternative Attribute

4.4.1 Research Motivation and Goal

The purpose of this empirical study is to estimate consumer preference for different types of new transportation modes. Thus, this study represents the evaluation stage of consumers' decision-making process with an intention to identify the most preferred form of new transportation mode based on consumer preference. The study uses discrete choice experiment and hierarchical Bayesian model to analyze heterogeneous consumer preferences for different attributes related to the new transportation modes. Also, along with identifying preferences for key attributes of the new services and their potential market size, this study attempts to provide quantitative evidence of the sources to some of the social conflicts in implementing the new transportation services.

As mentioned in the earlier section, introduction of the new transportation service may cause a shift in the paradigm of how consumers travel on daily basis, and affect their intention to own a vehicle. As a result, many social conflicts, especially conflict among the traditional and the new transportation services has become one of the key issues for the government promoting the sharing economy. Although there have been many previous studies that analyzed the market potential of the new transportation modes, few have provided inferences regarding the source or solutions to social issues related to implementing them. Thus, the present case's goal is to not only analyze consumer preferences of the new transportation services, but derive important inferences related to the social conflicts in promoting them.

Some previous studies have analyzed consumer preference for different transportation mode, but have been limited in analyzing the substitution patterns among the traditional transportation mode and the new sharing types of transportations. (Bhat and Sen, 2006; Jian, Rashidi, and Dixit, 2017; Kim et al., 2015; Park and Park, 2015; Jang, 2015; Seo et al., 2018 Reference 3)

The consumer preferences are used in the scenario analysis to estimate the choice probability of the new transportation modes and determine the shape of the potential automobile transportation service market in the future. Also, different scenario analysis is conducted to explain the reasons behind the social conflict related to the implementation of the new transportation modes. The results of this empirical study can be a guideline for policy makers and service providers to promote stable yet effective diffusion strategy of new transportation modes.

4.4.2 Data and Model Setup

The present study uses discrete choice experiment data and the hierarchical Bayesian model explained in section 3 to analyze consumer preferences of new automobile-based transportation services defined by taxi, carpool, car sharing, and driver sharing. For the analysis, stated preference data was collected from conjoint sets consisting of four alternatives.7 Four distinct type of automobile-based services were included in the conjoint card with six key attributes and levels were identified to distinguish the types of the automobile-based transportation services. Key attributes were identified to distinguish the types of services and each respondent was given graphical descriptions about the attributes and their level in the conjoint questionnaire before making a choice. Table (19) describe the key attributes and their levels in the conjoint survey.

The discrete choice experiment can avoid the multicollinearity problem in common regression analysis of observational studies. (Thyne et al., 2006; Haider and Ewing, 1990) This is because an discrete choice experiments are constructed based on orthogonal fractional factorial design, which ensure that there are no correlation among the attributes. (Danaher, 1997) In addition, discrete choice experiments allow researchers to

[7] Discrete choice experiment on the new transportation services was included in the survey described in the Data section

combine a number of attributes related to service and to control the alternative sets (Thyne et al., 2006; Haider and Ewing, 1990).

Table 19. Attributes in Automobile Based Transportation Service Conjoint Card

Attribute	Description	Level
Cost of Travel	The cost of travel per different service Calculated assuming a 10km one-way trip	200KRW/km 500KRW/km 800KRW/km
Fuel Type	Fuel type of the service vehicle	Gasoline Diesel LPG Hybrid EV
Body Type	Body type of the service vehicle	Small/Compact/Subcompact Mid/Full SUV/RV
Wait Time	Average time necessary to use the service vehicle	1 minutes 10 minutes 15 minutes 20 minutes
Driver Type	Driver type of the service vehicle Self Non-professional driver: Carpool and CTS Professional Driver: Taxi	Self Non-professional driver Professional driver
Commercial Vehicle	Whether the service vehicle is commercial or privately owned	Commercial Non-commercial

The attributes shown in Table (19) are based on the automobile services provided in Korea as of January 2020. The cost of travel attribute denotes the average cost of using different automobile-based transportation services. The attribute level of the cost attribute is assumed to be between 200KRW/km to 800KRW/km, which is based on the service cost of representative companies in Korea. (GreenCar, SoCar, Taxi, etc.) The attribute levels of fuel type attribute are based on vehicles provided by the automobile transportation services. Currently there are discussion within the policy makers to restrict the fuel type of all of the NGT vehicles to eco-friendly vehicles. However, since the present conjoint analysis focused on representing the current market, which consists of all fuel types included in Table (19). Total of five types of fuel were assumed in this discrete choice experiment: gasoline, diesel, LPG, hybrid, and electric. Gasoline, diesel, and LPG vehicles are operated by internal combustion engines and use a petroleum-based fuel; gasoline, diesel, or LPG fuel, respectively. Hybrid vehicles are operated by both an internal combustion engine and electric motor, but use only a petroleum-based fuel (usually gasoline). Hybrid vehicles generate electric energy using their own braking system instead of connecting to a power grid to recharge the battery. EVs are mainly powered by electricity, which is supplied from an external power source. In general, EVs plug into a power grid for electric energy to recharge the battery.

The attribute levels of the body types of vehicles provided by the automobile transportation services. The experiment assumed that vehicle types could be categorized into three types: small/compact/subcompact vehicles, mid/ full-size vehicles, and sport utility vehicles (SUVs)/recreational vehicles (RVs). The wait time attribute represents the average time required to use different types of automobile services. In case of car sharing service, we assumed the wait time included the time of the trip for the consumer to get to the service station, while the other services only denoted the actual wait time. Both driver type and commercial vehicle attributes are used to differentiate the types of services. For instance, alternative representing the carpool service will have non-professional for the attribute level of driver type, and non-commercial for the attribute level of commercial vehicle. As for CTS, the levels will be non-professional driver and commercial, while for CTS, the levels will be commercial, respectively. Thus, various combinations of these two attributes can derive many different forms of automobile-based transportation services.

Based on the attributes shown in Table (19), the number of possible combinations of alternatives with six NGT related attributes and the levels of each attributes is 1,080 ($3 \times 5 \times 3 \times 4 \times 3 \times 2$). However, it is practically challenging to use all the 1,080 alternatives to analyze the public preferences. Thus, this experiment established 32 alternative cards using the orthogonal fractional factorial design mentioned above.[8] Such procedure ensured the orthogonality of each attribute within and between alternatives, avoiding any multicollinearity problem. With these alternatives, eight choice sets consisting of four alternatives representing different types of NGT services are constructed. Each respondent was then asked to choose their most preferred alternative in each choice set shown in Figure (31). In addition to reducing the choice alternatives, employing a fractional factorial design also permitted us to avoid the multicollinearity problem. In other words, since the discrete choice experiment with a fractional factorial design has zero correlation among attributes (Danaher, 1997, Thyne et al., 2006, Verlegh et al., 2002), it is possible to estimate precise coefficients without the multicollinearity problem. (Huber and Zwerina, 1996)

[8] Factorial design was conducted by using IBM SPSS program's orthogonal design package. The package ensures orthogonality among the produced sets of alternatives

▣ Questionnaire 1

Attribute		Type A	Type B	Type C	Type D
1. Cost of Travel		800KRW/km	200KRW/km	200KRW/km	200KRW/km
2. Fuel Type		LPG	Gasoline	Diesel	EV
3. Body Type		Small/Subcompact /Compact	Mid/Large	SUV·RV	SUV·RV
4. Wait Time		5 Minutes	5 Minutes	5 Minutes	10 Minutes
5. Driver Type	5-1. Driver	Professional	Self	Non-professional	Self
	5-2. Vehicle	Commercial	Non-commercial	Commercial	Non-commercial
Choose the most preferred one					

Figure 31. Example if choice experiment for NGT services (translated from Korean)

4.4.3 Result and Discussion

4.4.3.1 Consumer preference analysis: Automobile based transportation service

This section reports consumer preferences for key attributes of the automobile based transportation service in Korea. Table (20) denote the consumer preferences for key attributes of automobile based transportation service. The estimation of hierarchical Bayesian model was computed through CBC Hierarchical Bayes Module Sawtooth Software. Version 5.2.8.

Table 20. Preferences for attributes of automobile-based transportation service

		β					
		α	Γ_{male}	Γ_{age}	Γ_{own_car}	Γ_{hh_income}	$\Gamma_{\ln(mileage)}$
Cost (100KRW/km)		-0.2323***					
		0.6159***	-0.1546**	-0.0057	0.0871	-0.0107	-0.0299
Fuel Type (base = hybrid)	Gasoline	-0.3406***					
		-0.7346**	-0.0131	0.0046	0.6211***	-0.0228	-0.0422
	Diesel	-0.9163***					
		0.1183	-0.3489***	-0.0171**	0.2240	-0.0419*	-0.0278
	LPG	-0.5404***					
		0.0551	-0.0535	-0.0077	0.2092	-0.0225	-0.0587*
	Electric	0.2882***					
		0.8812***	-0.0382	-0.0053	0.6548**	-0.0306	-0.1415***

Vehicle class (base = SUV)	~ Compact	-0.4196*	0.0704	0.0011	-0.2362***	0.0684	-0.0152	0.0201
	Mid-size ~	0.1206	0.2239**	-0.0096**	-0.1693***	-0.1930	0.0170	0.0139
Waiting time (min)		6.1584***	0.7587	0.0181	-0.0046***	1.0249	-0.1942***	-0.4926***
Driver (base = taxi driver)	Self	-0.4028	-0.2028*	0.0140***	0.1173***	-0.0061	0.0095	0.0026
	Not taxi driver	0.9946***	-0.2038**	-0.0044	0.3263***	-0.4839**	0.0204	-0.0079
Commerial Vehicle (base = Non-commercial)	Commercial	0.7325*	-0.1169	-0.0041	0.5448***	0.1732	-0.0092	-0.0127

*** 99%; ** 95%; * 90% significance level

Table (20) shows the estimated mean of βs. All of the posterior distributions of βs do not include zero at the 99% level, which indicates that all the estimated means are statistically significant. In the case of fuel type attribute, electric vehicles are most preferred, followed by hybrid, gasoline, LPG and diesel fuel in the respective order. The results are in line with the general norm regarding environmental awareness, as electric vehicles have the largest magnitude of preference, which is generally considered as the most environment-friendly fuel type. For vehicle type attribute, SUVs are most preferred, followed by mid-size or above vehicles and compact or smaller vehicles. For driver type, non-professional drivers are most preferred, followed by personal driving and professional driver. Here, non-professional drivers refer to drivers representing individuals who may be operating carpool or CTS. Also, low preference for professional driver may reflect public's negative perception towards taxi drivers for being aggressive on the road or being impolite to customers. Lastly, the results show that the consumers prefer commercial vehicles over non-commercial vehicles.

The estimation results of the parameter Γ are also shown in Table (20). If the signs of β_k and Γ_k are the same, preference for attribute covariates represented by k will be more sensitive when z increases and less sensitive when z decreases (Park and Koo, 2016). The results indicate that compared to females, males are more sensitive to price, prefer diesel vehicles compared to hybrid vehicles, and prefer larger/mid-sized vehicles compared to SUV. In case of driver type, female showed stronger preference for non-professional driver and driving personally compared to male. In other words, although male consumers also preferred to be driven by non-professional driver or driving personally, they were less sensitive to using taxi compared to female consumers. In case of age, as consumers' age gets older, they prefer diesel vehicles compared to hybrid vehicles, prefer SUV than mid/large size vehicle, prefer driving personally than being driven by professional drivers. Consumers with own private vehicles prefer gasoline and electric vehicles compared to hybrid vehicles and prefer non-professional drivers rather than professional drivers. This indicate that consumers with own vehicle has a tendency to drive personally or use services with professional driver rather than using services with non-professional drivers. Consumers with higher household income are more sensitive to waiting time than those with lower household income. In case of monthly mileage, consumers with higher monthly mileage have lower magnitude of preference toward LPG and electric vehicles compared to hybrid vehicles and are very sensitive waiting time.

Having discussed the parameter Γ, the above results can be interpreted in a different way in terms of the attributes. For example, consumers who are more sensitive to cost are males. Such interpretation can reveal which consumer groups have the highest preference for the key attributes. Compared to hybrid vehicles, consumers with private vehicles prefer gasoline vehicles the most, young female consumers most prefer diesel vehicles, consumers with lower monthly mileage prefer LPG vehicles the most and those with private vehicles and lower mileage rate show highest preference for electric vehicles. Consumers with higher household income and higher monthly mileage are more sensitive to waiting time than consumer with lower household income. Female consumers with no private cars prefer non-professional drivers compared to professional drivers. No distinctive consumer group showed significant preference for commercial vehicles.

4.4.3.2 Consumer preference analysis: Willingness to Pay and Relative Importance

This section reports the marginal willingness to pay for each attribute and the relative consumer preferences for key attributes of the automobile-based transportation services. Table (21) denote consumer's willingness to pay (WTP) and relative importance (RI) of the attributes of automobile-based transportation service.

Table 21. WTP and RI of attributes of automobile-based transportation service

		MWTP using Mean beta (KRW/km)	Median MWTP using indiv. Beta (KRW/km)	RI (%)
Cost		-	-	36.9
Fuel type (base = hybrid)	Gasoline	-146.6	-151.4	31.9
	Diesel	-394.4	-520.0	
	LPG	-232.6	-255.2	
	Electric	124.0	101.5	
Vehi. Class (base = SUV)	~ Compact	-101.7	-111.3	6.3
	Mid-size ~	-72.9	-103.4	
Wat time		-2.0	-0.0	1.8
Driver (base = taxi driver)	Self	50.5	57.6	8.6
	Not taxi driver	140.5	156.2	
Car	business	234.5	246.7	14.4

WTPs and RI can be calculated for each respondent based on the estimated β_k and Γ_k. The calculated consumer RIs revealed that cost (36.9%) is the most important deciding attribute in choosing the automobile-based service, followed by fuel type pf the vehicle (31.9%), whether the vehicle is commercial or not (14.4%), type of driver (8.6%), vehicle body type (6.3%) and the amount of wait time (1.8%) in the respective order. This indicates that changes in cost and operation method leads to the greatest variability.

The MWTP for fuel type is the highest with significant amount of difference. The negative WTP of diesel vehicles (-394.4 KRW/km) or the positive WTP of electric (124.0

KRW/km) vehicles is especially large in amount. In the case of vehicle class, the MWTP is consistently low, and the MWTP for waiting time is too miniscule to be considered significant.

4.4.3.3 Scenario Analysis Setting

Scenario analysis based on consumer preferences and the actual market conditions in Korea is conducted to analyze the choice probabilities for different types of automobile-based transportation services. The analysis begins by examining the attribute levels provided by the current automobile-based transportation services in Korea. Two assumptions are made for the analysis; distance of the trip for the consumer choosing the transportation service, regulation level of limiting the operation hours of the carpool service. Unlike the taxi or CTS, which is operational and available at all time, the carpool service's operation time has been limited to 'commute time' by government regulation.[9] The commute time are between 7~9 a.m. and 6~8 p.m. which only allows four operation hours for any form of carpool service. Since choice probability for the scenario is heavily dependent on the scenario setting, it is important for researchers to set the scenario as close as possible to the real market condition. (Axsen, 2017; Hardman 2018; Al-Alawi and Bradley, 2013) As a result, the scenarios in this study incorporate the regulation effect by assuming it has effect on the availability ratio of the carpool service to the taxi and CTS in

[9] The regulation has been in practice since July 2018, and although many carpool services are attempting to bypass this regulation, it is illegal to operate outside the commute time.

the scenario. In addition, to demonstrate the effect of this regulation, different scenarios with various level of regulations are analyzed and compared to the market state with no regulations. By showing the trends of changes in choice probability for each alternative under different market conditions, this analysis identify substitution patterns among different services and the effect of government regulations in the new automobile-based transportation service market.

The trip used as a base in the scenario analysis is about 10 km trip in Seoul from Gangnam to Geumcheon regional office respectively. Table (22) describes the value of each attributes of the baseline scenario. The duration of wait time is assumed to be same, since variation of time vary significantly for consumers. Table (23) denote the condition of the scenarios analyzed in this empirical test. After the first scenario, car sharing service is excluded from the analysis because it is a different type of transportation. T, S, P, DS denote the automobile-based transportation services; T being taxi, S being car sharing, P being carpool, and CTS being CTS. The first three scenarios only reflect the choice probability of consumers with the actual attribute levels of the services in Korea. The last six scenarios reflect the regulation effect by adjusting the level of availability of carpool service. The effect of regulation, or the level of availability for carpool service compared to taxi is divided into three levels to demonstrate the trend of changes in the choice probability due to the regulation measure. Mid-level of regulation is defined by 50% availability of carpool service compared to taxi, and low-level of regulation is defined by 90% availability of carpool service compared to taxi.

Table 22. Attribute levels of baseline scenario

		Taxi	Car Sharing[10]	Carpool[11]	CTS
Cost (1,000KRW)12		10	6.5	8	14
	Gasoline	0	1	1	0
	Diesel	0	0	0	1
Fuel type	LPG	1	0	0	0
	Electric	0	0	0	0
	Hybrid	0	0	0	0
Vehicle Type	Small	0	1	0	0
	Big	1	0	1	0
	SUV	0	0	0	1
Wait-time (minutes)		10	10	10	10
Self		0	1	0	0
Non-professional driver		0	0	1	1
Professional Driver		1	0	0	0
Commercial		1	1	0	1
Non-commercial		0	0	1	0

[10] SOCAR is one of the biggest car sharing service providers in Korea
[11] POOLUS is major carpooling platform in Korea
[12] Average of the actual cost of 10km trip

Table 23. Conditions of different scenarios

Scenario	Taxi	Carpool	CTS	Car Sharing
Baseline	✓	✓	✓	✓
1	✓	✓	-	-
2	✓	✓	✓	-
		+ Regulation effect[13]		
1+(Low)	✓	✓	-	-
1+(Mid)	✓	✓	-	-
2+(Low)	✓	✓	✓	-
2+(Mid)	✓	✓	✓	-

✓: Indicate the inclusion of service in the analysis

4.4.3.4 Scenario Analysis Result

This section presents the results of the scenario analysis described in the previous section. Table (24) show the choice probabilities according to different scenarios.

[13] (Low = 10%, Mid= 50%, High=90%) indicate the level of regulation, high regulation means the availability of carpool is very limited and vice versa

Table 24. Choice probability of different alternatives

Scenario	Taxi	Carpool	CTS	Car Sharing
Baseline	16.9%	26.5%	7.54%	49.1%
1	39.0%	61.0%	-	-
2	33.3%	51.9%	14.8%	-
		+ Regulation effect		
1+(Low)	39.0%	61.0%	-	-
1+(Mid)	69.5%	30.5%	-	-
2+(Low)	33.3%	51.9%	14.8%	-
2+(Mid)	51.2%	26.0%	22.8%	-

The baseline scenario shows the choice probability of four types of automobile-based transportation service with no regulation in practice. In other words, the baseline probability represents the choice probability with the actual attribute levels in Korean market, with all services being 100% available. In such situation, car sharing service has the highest probability (49.1%), with carpool (26.5%), taxi (16.9%), and CTS (7.54%). However, as mentioned in the previous section, car sharing is excluded from the other scenarios because it is similar to rental services rather than the other transportation services in the analysis. The rest of the scenario analyze the choice probabilities based only on taxi, carpool, and CTS alternatives. Figure (32) and Figure (34) show the changing trends of each alternative based on different market and regulation level.

Figure (32) show the changes in the probability of taxi, carpool, and CTS alternatives from scenario 1 to 2, where CTS enters the market.

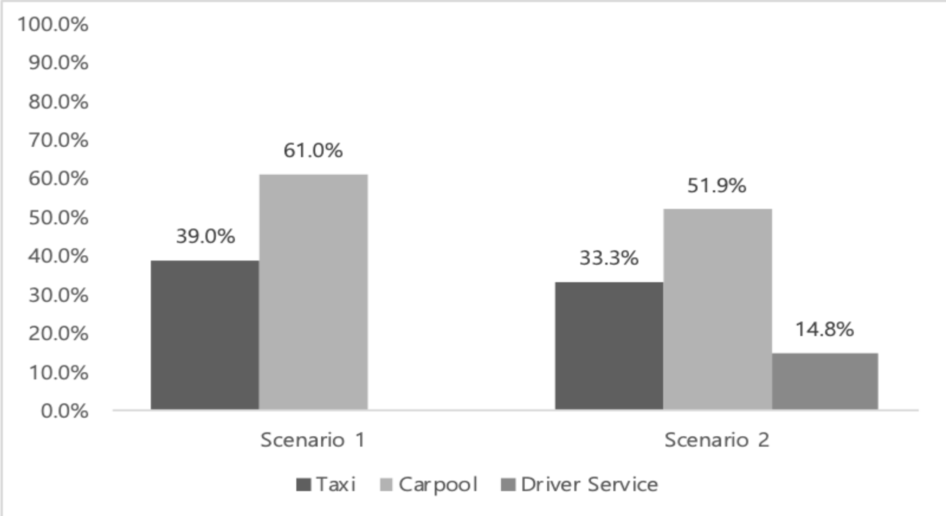


Figure 32. Effect of CTS Entering the Market

Figure (33) show the changes in the probability of taxi and carpool service (Scenario 1), with different level of the regulation on the carpool service.

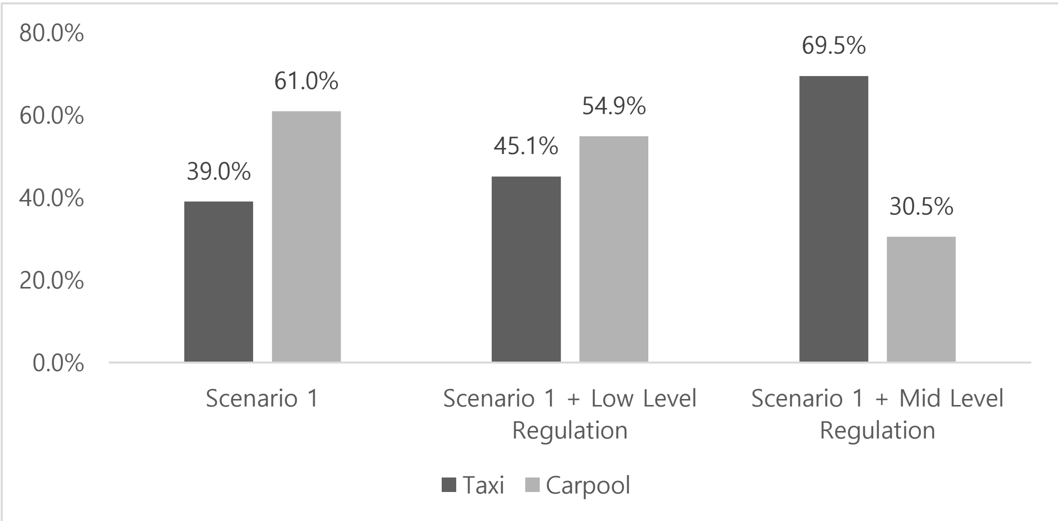


Figure 33. Regulation Effect on Market with Taxi and Carpool

Figure (34) show the changes in the probability of taxi, carpool service, and CTS (Scenario 2), with different level of the regulation on the carpool service.

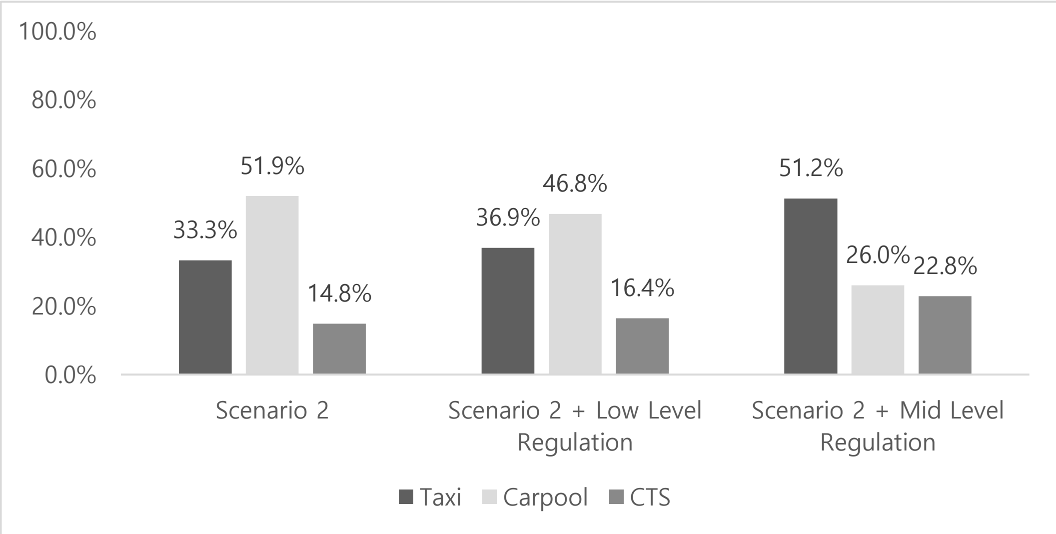


Figure 34. Regulation Effect on Market with Taxi, Carpool, and CTS

The results of the second empirical study analyze the consumer preference towards the different forms of NGT services, providing valuable insights to the marketers and service providers of the NGT service. Also, the scenario analysis result based on the consumer preferences demonstrate that consumers have a high tendency to choose carsharing, carpool and CTS service over taxi when all of the NGT services are available, and even when they are under mild regulations. The substitution patterns recognized in the scenario analysis indicate that consumers will likely adopt the NGT services, and even replace their current use of taxi when they are diffused.

4.5 Empirical Study III: Heterogenous Preferences for Alternative

4.5.1 Research Motivation and Goal

The purpose of this empirical study is to estimate the impact of new types of transportation on consumer travel mode selection behavior and its environmental impact. The study elaborately analyzes the impact of new transportation modes represented by car-sharing, carpool, and ride-sharing services on the transportation sector. This study inspects various aspects of the new transportation methods that replace existing transportation methods such as buses, subways, taxis, and private vehicles. Some previous studies have analyzed consumer preference for different transportation mode, but have been limited in analyzing the substitution patterns among the traditional transportation mode and the new sharing types of transportations. (Bhat and Sen, 2006; Jian, Rashidi, and Dixit, 2017; Kim et al., 2015; Park and Park, 2015; Jang, 2015; Seo et al., 2018 Reference 4) The result of this analysis will help understand the consumer preference towards different transportation modes and explain the complicated relationship among them. Also, estimating the environmental effects of usage based on consumer preferences and transportation usage may serve as a guidance to implement policies that increase environmental welfare during the diffusion of new transportation modes.

4.5.2 Data and Model Setup

Empirical analysis of consumer preference towards different transportation mode and their usage are conducted. To successfully reflect consumers' behavior of using transportation, which consists of distributing the total distance travelled according to their preference, it is necessary to use a model that can estimate consumer behavior and usage distribution simultaneously. As a result, the present study uses the Mixed MDCEV (multiple discrete-continuous extreme value) model described in section 3.4 to analyze consumer preferences and usage of transportation modes. After analyzing the consumer preference for transportation modes, scenario analysis assuming the different introduction time of new transportation modes are conducted. By executing the scenario analysis, various inferences regarding the effect of new transportation modes in the usage of existing transportation modes are analyzed. Then, based on the estimated usage for different transportation modes, environmental effect is calculated to analyze the environmental benefits of new transportation modes entering the market.

Survey data regarding monthly distance travelled for each transportation modes and various socio-demographic data are used in the analysis. Six transportation modes are defined for the analysis; these consist of traditional and new transportation modes such as car-sharing, carpool, and ride-sharing services. Socio-demographic information such as gender, age, monthly household income, and status of travel during commute time is used to differentiate consumer preferences of transportation modes. Descriptions of the data used in the analysis is represented in Table (25)

Table 25. Consumer data used in the Mixed MDCEV analysis

Survey	Description
Distance travel by transportation modes	
Private automobile	Monthly distance travelled by consumer according to the service
Taxi	
Car sharing service	
Car pool service	
Rider sharing service	
Public (Subway/Bus)	
Demographic Variables	
Gender	0 if male, and 1 if female
Age	Age of the respondent
Income	Household income of the respondent
Commute time	1 if respondent travel during commute time 0 if respondent does not travel during commute time (commute time is defined as 6 a.m~9 a.m and 6 p.m~9 p.m)

The estimation is conducted by the Mixed MDCEV written in GAUSS, the code is based on the GAUSS code provided by Bhat. The code follows the Bayesian estimation. (Bhat, 2008).

4.5.3 Result and Discussion

4.5.3.1 Consumer Preference Analysis: Transportation Modes

Table (26) represent heterogeneous consumer preferences for transportation modes using the mixed MDCEV model. Table (27) represent the satiation parameter of six transportation modes in the analysis.

Table 26. Baseline utility of consumer preference for transportation modes

Type	Variables	Baseline utility	
Personal Car (baseline)		Mean	Variance
Public (Subway + Bus)	ASC	0.3535**	0.0319***
	Female	0.7918***	0.0255***
	Age	-0.6273***	0.2455***
	Income	-0.2095**	0.2734***
	Commute Time	0.3574***	0.028***
Taxi	ASC	0.0506	0.0295***
	Female	0.8632***	0.0228***
	Age	-1.4483***	0.1559***
	Income	-0.5738***	0.1878***
	Commute Time	0.2595	0.0228***
Sharing	ASC	0.452**	0.0239***
	Female	-0.3327**	0.0192***
	Age	-1.8644***	0.098***
	Income	-1.0243***	0.1296***
	Commute Time	-0.6778***	0.0206***
Driver	ASC	0.0747	0.0188***
	Female	-0.1054	0.0193***
	Age	-2.4228***	0.0618***
	Income	-1.7411***	0.0823***
	Commute Time	-0.0564	0.0172***
Carpool	ASC	-0.2729**	0.023***
	Female	-0.683**	0.0197***
	Age	-1.9795***	0.0914***
	Income	-1.2227***	0.1111***
	Commute Time	-0.4664***	0.0204***

Note 1: ASC is the Alternative Specific constant

Note 2: Female variable is coded in a way that male is 0 and female is 1, (+) coefficient may indicate that females preferred the transportation more than male
Note 3: ***,**,* refer to significance level at 99%, 95%, 90%

Table 27. Satiation parameter for transportation modes

Transport	Satiation	
	Mean	Variance
Personal Car	0.1371***	0.0019***
Public Transit (Bus/Subway)	0.0587***	0.0005***
Taxi	0.0374***	0.0005***
Car-sharing Service	0.0728***	0.0005***
CTS	0.2712***	0.0009***
Carpool	0.1202***	0.0006***
Gamma (γ)	17.7492***	52.0471***

Note 3: ***, **, * refer to significance level at 99%, 95%, 90%

Various inferences can be made from the estimated results of the baseline utility. The coefficients of the 'Female' variable is positive for public and taxi, compared to negative coefficient for car-sharing service, CTS, and carpool service. Such result indicate that male may consist the majority of the consumers for new transportation modes. Female consumers preferred traditional transportation modes compared to male, while statistically, did not prefer any of the new transportation modes compared to male. In particular, female consumers showed strong distaste for carpool service than male consumers. Such difference may be due to the fact that females are more reluctant to ride a vehicle operated by non-screened individual. As described in section 2, the carpool service represents a form of transportation service operated by any individual with their own car. Unlike taxis, which are only operated by certified drivers officially registered to a taxi company or government, carpool drivers consist of any individual drivers with private cars and registered to the service center. Thus, the negative coefficient of Female variable for carpool service may represent female consumers' fear of meeting harsh or even dangerous driver.

The coefficient of the 'Age' variable is consistently negative for all of the transportation mode meaning that as age of the consumers became younger, they preferred different transportation modes to private vehicle. Consumers preference based on age was CTS, carpool, car sharing, taxi, public, and private vehicle. Such result indicates that younger consumers prefer new transportation services to the traditional services. In particular, younger consumers' preference the CTS the most, which provides relatively young professional drivers with pick-up service compared to taxi. Although both doesn't require the consumer to drive, consumers who are young is more likely to prefer new CTS compared to taxi.

The coefficient of the 'Income' variable is statistically significant and negative for all of the transportation modes. These results indicate that consumers with lower income preferred car-sharing and carpool service compared to other transportation modes. According to the estimation, as consumer's income became lower, the preferred order of transportation modes was CTS, carpool, car sharing, taxi, public transportation, and private vehicle. Such result may be due to the tendency of consumers with lower income to prefer transportation modes that doesn't require them to drive. Both CTS and carpool services are automobile transportation modes that provide separate drivers, compared to private vehicles and car sharing services that require the consumer to drive. In other words, consumers with higher income preferred own vehicle the most compared to any other transportation modes.

The 'Commute Time' variable indicated whether the consumer frequently travelled during commute time, which was defined in the survey as 6 a.m. to 9 a.m., and 6 p.m. to 9 p.m. during the day. The coefficients of commute time are statistically significant for public, car sharing, and carpool service. Contrary to the positive coefficient for the public transportations, consumers had negative coefficients for both of the new transportation modes. In particular, consumer who usually travelled during commute time tended to prefer the traditional transportation modes such as bus and subway. Such result may be due to the uncertainty related to the new transportation modes, which compared to bus and subway may be strongly affected by the external factors such as availability of the service and road conditions at the time. Close analysis of how uncertainty related to the availability of transportation services, and how different circumstances or travel purpose affect consumer's preference of transportation mode will be necessary to confirm these inferences.

The satiation parameters are all statistically significant for every transportation mode. According to the estimation result, consumers have the highest satiation for CTS, and lowest satiation for taxi. As higher satiation parameter indicates lower satiation effect, it can be inferred that consumer's tendency to increase the usage of CTS is higher compared to the tendency to increase the usage of taxi or public transportations. Consumers show similar satiation parameter for both carpool service and private vehicle, with intentions to increase the usage. Among the three new transportation modes, car sharing service has the lowest satiation parameter, almost similar to the satiation parameter of public transportation. Such result indicates that high satiation occurs for these transportation modes, which may be due to the fact that car sharing service require the consumer to drive, while other services provide separate drivers. The low satiation parameter for traditional transportations may be caused by already optimized usage set up by consumers. Since traditional transportation modes have existed for long time, consumers may have already optimized their usage.

4.5.3.2 Scenario Analysis Setting

Scenario analysis is conducted to analyze the usage pattern changes for different transportation modes with a sequential introduction of new transportation modes. In addition to providing information about the transportation usages of different modes, the scenario result provides quantitative information of fuel use and pollutant-emissions patterns based on the diffusion of new transportation modes. Based on the assumption that individual data on monthly car usage is fixed constraint, the optimal usage of each transportation mode maximizes individual's utility. As a result, consumer i's optimal usage of transportation mode j is denoted by m_{ij}, estimated through equation below.

$$\max \quad \mathrm{U}_i = \sum_{j=1}^{K} \left[\exp\left(\hat{\beta}_i' x_{ij} \right) \right] \left(m_{ij} + \hat{\gamma}_i \right)^{\hat{\alpha}_{ij}}$$

$$s.t \quad \sum_{j=1}^{K} m_{ij} = M_i \qquad \text{Eq. 82}$$

Where $\hat{\beta}_i, \hat{\gamma}_i$, and $\hat{\alpha}_{ij}$ are given values in the scenario analysis, M_i is the monthly travel distance of each individual i, also given in the scenario analysis. Equation (81) estimates the m_{ij} that maximize utility under the given constraint.

To successfully analyze the usage pattern changes of the traditional transportation modes following the introduction of carpool, car sharing, and CTS, four scenarios are considered. The first scenario is the base scenario, reflecting the traditional transportation modes where consumers can choose to drive their private vehicle, taxi, or public transportations like subway and bus. The second scenario introduces car sharing service to the first scenario. CTS is then added to the second scenario in the third scenario, while the carpool service in added to the third scenario in the fourth scenario. The order of the service added to each scenario is based on the actual implementation of new transportation modes in Korea. The setup of the scenarios is depicted in Figure (35).

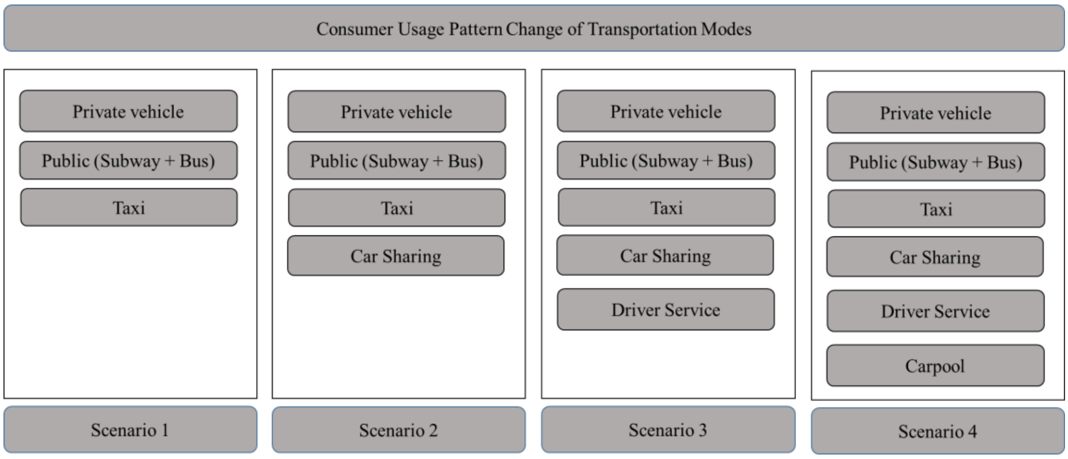


Figure 35. Scenario setup for usage changes of transportation mode

4.5.3.3 Scenario Analysis Result

Table (28) represent the estimated distance travelled (usage) of different transportation modes according to the scenarios. Table (29) represent the changes in usage behavior for different transportation modes according to implementing new modes sequentially.

Table 28. Usage of transportation modes per scenario

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
	Average usage (km/month)	Average usage (km/month)	Average usage (km/month)	Average usage (km/month)
Personal Car	707.8	675.8	662.7	636.8
Public Transit	270.9	270.5	267.9	261.6
Taxi	187.6	154.6	140.1	143.7
Car-sharing Service	-	65.4	48.8	36.7
Driver Service	-	-	46.8	36.4
Carpool	-	-	-	51.1

Table 29. Changes in the usages of transportation modes per scenario

	Scenario 1→2 (Ratio)	Scenario 2→3 (Ratio)	Scenario 3→4 (Ratio)
Personal Car	-32.0 (-4.5%)	-13.1 (-1.9%)	-25.9 (-3.9%)
Public Transit	-0.4 (-0.1%)	-2.6 (-1.0%)	-6.3 (-2.4%)
Taxi	-33.0 (-17.6%)	-14.5 (-9.4%)	+3.6 (+2.6%)
Car-sharing Service	+	-16.6 (-25.4%)	-12.0 (-24.7%)
CTS	-	+	-10.4 (-22.1%)
Carpool	-	-	+

The results show that the introduction of car sharing, and carpool services significantly decrease the use of private vehicle, while the effect is minimal to the use of public transportations. Looking at the changes in the usages of different transportation modes, the use of taxi decreases significantly when car sharing and driving service enter the market but increase weakly after the carpool service enter the market. This increase in usage may indicate the existence of a weak complementary relations between taxi and carpool services. Based on this assumption, the introduction of carpool could alter the consumer transportation mode selection behavior that could have increased the use of taxi. However, further research is necessary to provide the quantitative evidence of this assumption. Among the new transportation modes introduced to the market, carpool service had the most usage, while the usage of car sharing, and CTS were similar to one another. When all of the transportation modes were available in the market, compared to the use of taxi, the use of private vehicle and public transportations were relatively unchanged from scenario 1 to scenario 4. Ratio wise the usage of taxi was transferred to the usage of new transportation modes.

4.5.3.4 Environmental Effect of New Transportation Modes

Based on the usage estimated in the previous section, changes in pollutant-emission in the transportation sector due to new transportation modes entering the market was estimated. To estimate the pollution change, fuel usage for each transportation modes and pollutants statistics were collected. Table (30) denote the ratio of fuel usage for different transportation modes. Taxis are assumed to consist of only LPG vehicles, driver sharing services is assumed to consist of only diesel[14], and the fuel type ratio of car sharing service is based on the actual ratio of commercial vehicles registered in the car sharing market. (Yoo, 2019)

[14] Until 2020, the driver sharing service in Korea majorly consisted of diesel minivans

Table 30. Ratio of fuel usage for transportation modes

Transportation mode	Gasoline	Diesel	LPG	Hybrid	EV
Private vehicle15	56.5%	32.6%	7.8%	2.6%	0.4%
Taxi	0.0%	0.0%	100.0%	0.0%	0.0%
Car Sharing	48.9%	28.2%	6.8%	13.%	2.2%
CTS	0.0%	100.0%	0.0%	0.0%	0.0%
Carpool	56.5%	32.6%	7.8%	2.6%	0.4%

Table (31) denote the pollutant emission (g/km) for different fuel types. The emission factor is based on the statistical report by the Korea' Energy Research Institute. The emission factor of hybrid fuel types is assumed to be partial amount of the emission factor of gasoline[16] This analysis only considered the tank-to-wheel emission, which only accounts for pollutant emission on road side. Thus, the emission factor for EV are considered 0 in this analysis. Further research may include pollution factor of EV by considering the pollutant emitted during the generation of electric fuels.

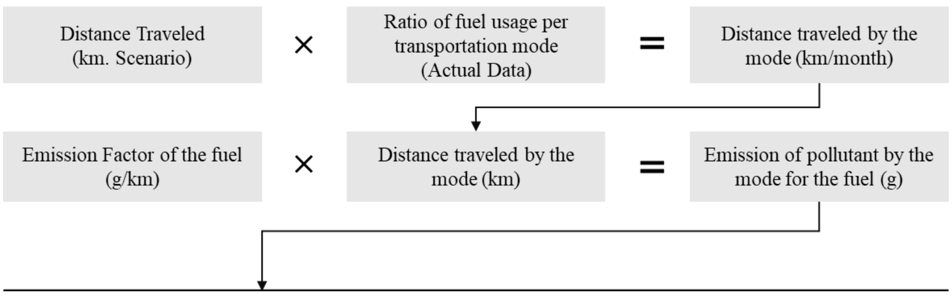
[15] Fuel usage is calculated based on the non-commercial vehicle registered in the Korea Ministry of Land, Infrastructure and Transport in 2020

[16] Ratio of fuel efficiency of gasoline and hybrid for the same type vehicle, Hyundai Grandeur is 69.1%. The emission factor for hybrid vehicle was assumed to follow the same ratio.

Table 31. Emission factor for different pollutant based on fuel types

Pollutant	Gasoline	Diesel	LPG	Hybrid	EV
CO(g/km)	0.1940	0.0590	0.1030	0.1341	0.000
VOC(g/km)	0.0160	0.0050	0.0130	0.0111	0.000
NOx(g/km)	0.0180	0.2010	0.0110	0.0124	0.000
PM(g/km)	0.0018	0.0021	0.0020	0.0012	0.000

Based on the estimated usage (travel distance) for average consumer in scenario analysis of the previous section, the average amount of emission change caused by new transportation entering the market is calculated based on the following procedure. (IMF, 2014) The overall estimation procedure of pollution effect is shown in Figure (36).



Pollutant	Gasoline	Diesel	LPG	Hybrid	EV
CO(g/km)	77.582	13.614	25.009	2.468	0.000
VOC(g/km)	6.399	1.154	3.157	0.204	0.000
NOx(g/km)	7.198	46.379	2.671	0.228	0.000
PM(g/km)	0.720	0.485	0.486	0.022	0.000

Figure 36. Estimation process of pollutant for fuel type based on scenario analysis

Table (32) to Table (35) denote the calculated pollutants for each scenario.

Table 32. Scenario 1 Pollutant emitted by average user per fuel types

Pollutant	Gasoline	Diesel	LPG	Hybrid	EV
CO(g)	77.582	13.614	25.009	2.468	0.000
VOC(g)	6.399	1.154	3.157	0.204	0.000
NOx(g)	7.198	46.379	2.671	0.228	0.000
PM(g)	0.720	0.485	0.486	0.022	0.000
Total	91.899	61.631	31.322	2.922	0.000

Table 33. Scenario 2 Pollutant emitted by average user per fuel types

Pollutant	Gasoline	Diesel	LPG	Hybrid	EV
CO(g)	80.279	14.086	21.811	3.496	0.000
VOC(g)	6.621	1.194	2.753	0.289	0.000
NOx(g)	7.449	47.989	2.329	0.323	0.000
PM(g)	0.745	0.501	0.424	0.031	0.000
Total	95.093	63.771	27.317	4.140	0.000

Table 34. Scenario 3 Pollutant emitted by average user per fuel types

Pollutant	Gasoline	Diesel	LPG	Hybrid	EV
CO(g)	85.929	17.839	20.736	4.753	0
VOC(g)	7.087	1.512	2.617	0.393	0
NOx(g)	7.973	60.772	2.214	0.439	0
PM(g)	0.797	0.635	0.403	0.043	0
Total	101.786	80.757	25.970	5.628	0

Table 35. Scenario 4 Pollutant emitted by average user per fuel types

Pollutant	Gasoline	Diesel	LPG	Hybrid	EV
CO(g)	78.882	15.989	20.585	3.0382232	0.000
VOC(g)	6.506	1.355	2.598	0.251486	0.000
NOx(g)	7.319	54.472	2.198	0.2809394	0.000
PM(g)	0.732	0.569	0.400	0.0271877	0.000
Total	93.439	72.385	25.781	3.5978363	0.000

Since the analysis calculate the average amount of pollutant for an average consumer, the change ratio is more suitable to analyze the changes in pollutant for different scenarios of new transportation modes entering the market. Figure (36) depicts the changes in pollutant emission comparing different scenarios of new transportation modes entering the market.

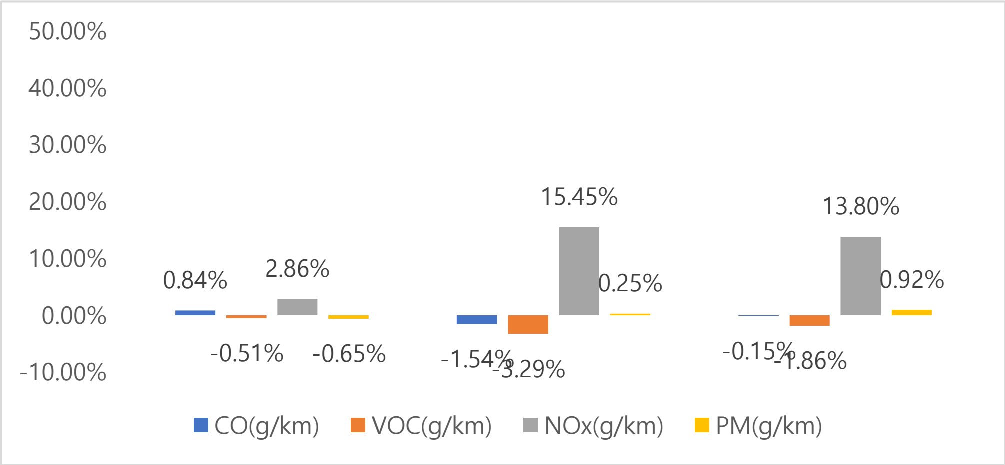


Figure 37. Change in the pollutant emission based on scenario analysis

The result of the analysis shows in the emission of CO and NOx when car sharing enters the market, while the emission of VOC and PM decreases. This reflect the change in the usage of LPG fueled taxi to eco-friendly vehicle consisting the car sharing service. When driver sharing service enter the market, which consists of all diesel vehicle, emission of NOx increases significantly. Lastly, when the carpool service enters the market, NOx emission decreases then emission in scenario 3 due to the decreased usage of CTS, but still remain significantly high compared to scenario 1 and 2. The emission of other pollutants CO, VOC, and PM also increase compared to scenario 3. Based on these results, no promising indicator of significant environmental benefits of adopting the new transportation modes. However, the estimation of the environmental effect in this study does not considers of the different social cost related to the pollutants, which may lead to environmental benefits. Further research should also consider the expansion of environmentally friendly vehicles in the private vehicle sector, as increase in electric vehicle and hydrogen fuel vehicles could significantly decrease the pollutant emission.

Chapter 5. Discussion

This dissertation proposes a research framework that incorporates artificial neural networks and choice models to enhance the analysis of consumer decision making. It analyzes three different types of consumer choices, each representing a different decision stage in the process that leads to the purchase of new products or services. This study assumes that consumers need to make three types of choices prior to adopting new products. These choices include making the decision to buy, deciding among different product options, and determining whether to use the chosen product. The subject of analysis in this study is next-generation transportation services in Korea, which consist of carsharing, carpooling, and centralized taxi services. These services are different from automobile-based transportation, providing consumers new travel alternatives.

Three studies are conducted based on choice types. The first study uses different artificial neural networks to study consumer choice. It consists of two parts. The first part uses various forms of artificial neural networks to predict consumer intentions, yielding superior prediction performance than traditional logistic regression predictions. The second part proposes an attention-based feature identification model that determines the importance of input data in relation to outputs. The identification method proposed in this study shows a marginally superior ability to identify the initial variable than the step-wise variable selection model. Also, the variables with high attention weights show significant model fit. Although limited to this case, the analysis indicates that attention-based neural networks can be applied to identify key variables for analysis with highly uncertain data. In addition, this model has a high potential to identify key variables during the process of designing conjoint surveys regarding new products or services. However, there remains much room for improvement. Since the current model suggests weights for all of the variables in the analysis, constraint measures controlling bias and overfitting in the model should be developed. Such bias might result in the inclusion—based on their weights—of certain variables that are not significant at all. Further research should be conducted to identify a cut-off point for truly selecting the most significant variables using this model. Nevertheless, the proposed model highlights the great potential of using artificial neural networks to analyze consumer choice.

The second and third studies are conducted as in previous research, but are specific to the NGT market in Korea. In particular, the findings in the second and third studies can be used to predict changes in the automobile market and the possible environmental effects of the entrance of NGTs into the market in Korea. The analyses in the second and third empirical studies support the following inferences. First, the results show that consumers who travel outside of normal commuting times have higher tendencies to use the NGT services, indicating that government regulations limiting carpool operation hours are effective in protecting the taxi industry. However, it should also be noted that such regulations might harm social welfare because they limit alternatives for consumers. Secondly, the results show that consumers who are already using NGT services tend to increase their usage, while consumers who use traditional transportation services do not show any intentions to increase their usage of NGT services. This difference in intentions indicates that consumers will likely increase their use of NGT services in the future, which may make the transportation market more oriented toward NGT services.

Third, consumers show a high preference for the CTS in the analysis; however, during the study, the Korean government banned the CTS, stating that it violated taxi industry regulations. Although many acknowledge the fact that the banned CTS had in fact abused the law, consumers show a high preference for this kind of service. Thus, policies should be implemented to strictly define different types of transportation modes, which may allow different types of CTS to grow in the future. Fourth, the analysis regarding the usage of different transportation modes shows that the relationship between the carpool service and taxis may be somewhat complementary. When other types of NGT services entered the market, taxi usage dropped significantly. However, when the carpool service entered the market, there was a slight increase in the usage of taxis, indicating a complementary response. Although the present study does not specifically analyze this dynamic, future studies should analyze the possibility of a complementary relationship between these two services to aid in the development of a sustainable transportation market.

Also, this study's analysis of the environmental effects of NGT services entering the market provides valuable insights regarding environmental regulations in the transportation sector. The promotion of NGT services is designed to boost the sharing economy, which many regards as more sustainable and efficient than traditional transportation services. However, the results of this analysis show that if the NGT services maintain their current shares of fuel type vehicles, their use will not produce environmental benefits. These findings indicate that the use of eco-friendly vehicles should be enforced by regulation or promoted by subsidies to insure that NGT services have beneficial environmental effects.

In sum, this study applies three different models to estimate the market potential of the NGT services by analyzing the consumers' intentions, preferences for different types of services, and usage decisions regarding the NGT services. Although the models or results of each empirical study are not integrated, they provide valuable insights regarding the effects of the NGT services on the transportation service market. The first empirical study predicts consumers' intentions to use NGT services and investigates the possibility of using ANN in modeling consumer decisions. The second empirical study analyzes consumer preferences regarding NGT services using a hierarchical Bayesian model to identify key demographic variables that differentiate consumers' preferences. The final empirical study uses the Mixed MDCEV model to analyze consumer preferences regarding different modes of travel including the NGT services. The analysis results reveal how consumer behavior changes when the NGT services enter the market.

The limitations of this study include that it fails to integrate the results from different models and empirical studies into a single model. As mentioned in Chapter 2, some researchers have tried to combine artificial neural network and discrete choice models. However, many of them have been limited in developing adequate models that successfully incorporate the advantages of ANN and the discrete choice models. Although this dissertation also utilizes both types of models, it doesn't fully combine them.

Furthermore, each empirical study could be improved to provide more detailed explanations regarding the types of consumer decisions under analysis. The interpretative power of the ANN model could be tested by comparing the proposed model to a more advanced choice model. By using different sets of variables sequentially in the ANN and possible choice models, the ANN model's ability to detect non-linear relationships could be tested. Such a procedure may provide valuable insights into methods for overcoming the "black-box" characteristics of ANN. The right combination of the two models in analyzing different consumer decision might provide more accurate (ANN) and richer interpretations (discrete choice model).

Also, this model uses scenario-based analysis as the bases for various inferences. Although scenario analysis is generally accepted as a sound methodology that can be used to draw inferences, it could be improved in many ways by incorporating the abundant information available in the present day. Some of this study's findings reveal the possibility of a complementary relationship between carpool and taxi services. Future studies could dive deeper into this topic to provide more accurate analysis of the NGT services.

Bibliography

Agrawal, D., & Schorling, C. (1996). Market share forecasting: An empirical comparison of artificial neural networks and multinomial logit model. *Journal of Retailing*, 72(4), 383-408.

Abraham, H., Lee, C., Brady, S., Fitzgerald, C., Mehler, B., Reimer, B., & Coughlin, J. F. (2016). Autonomous vehicles, trust, and driving alternatives: A survey of consumer preferences. *Massachusetts Inst. Technol, AgeLab,* Cambridge, 1, 16.

Aghdam, H. H., & Heravi, E. J. (2017). Guide to convolutional neural networks. *New York, NY:* Springer, *10*, 978-973.

Agrawal, D., & Schorling, C. (1996). Market share forecasting: An empirical comparison of artificial neural networks and multinomial logit model. *Journal of Retailing*, 72(4), 383-408.

Ajzen, I. (1991). The Theory of Planned Behavior. Organization Behavior and Human Decision Processes, *Amherst*. Academic Press, Inc. 179-211.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723.

Al-Alawi, B. M., & Bradley, T. H. (2013). Review of hybrid, plug-in hybrid, and electric vehicle market modeling studies. *Renewable and Sustainable Energy Reviews*, 21, 190-203.

Allenby, G. M., & Rossi, P. E. (2006). Hierarchical Bayes models. The Handbook of Marketing Research: Uses, Misuses, and Future Advances, Grover and Vriens. Sage Publications, 418-440.

Alwosheel, A., van Cranenburgh, S., & Chorus, C. G. (2018). Is your dataset big enough? sample size requirements when using artificial neural networks for discrete choice analysis. *Journal of Choice Modelling*, 28, 167-182.

Bagozzi, R. (2007) The Legacy of the Technology Acceptance Model and a Proposal for a Paradigm Shift. *Journal of the Association for Information Systems*, 8, 244-254.

Balas, V. E., Kumar, R., & Srivastava, R. (Eds.). (2020). *Recent Trends and Advances in Artificial Intelligence and Internet of Things*. Cham. Springer.

Barthélemy, J., Dumont, M., & Carletti, T. (2018). Comparison of Discrete Choice Models and Artificial Neural Networks in Presence of Missing Variables. *arXiv preprint arXiv:1811.02284.*

Bhat, C. R. (2005). A multiple discrete–continuous extreme value model: formulation and application to discretionary time-use decisions. *Transportation Research Part B: Methodological*, 39(8), 679-707.

Bhat, C. R. (2008). The multiple discrete-continuous extreme value (MDCEV) model: role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B: Methodological*, 42(3), 274-303.

Bhat, C. R., & Sen, S. (2006). Household vehicle type holdings and usage: an application of the multiple discrete-continuous extreme value (MDCEV) model. *Transportation Research Part B: Methodological*, 40(1), 35-53.

Block, C. E., & Roering, K. J. (1976). *Essentials of Consumer Behavior.* Dryden Press.

Blundell, C., Cornebise, J., Kavukcuoglu, K., & Wierstra, D. (2015). Weight uncertainty in neural networks. *arXiv preprint arXiv:1505.05424.*

Blythe, J. (2008). *Consumer Behaviour*. Cengage Learning EMEA.

Bronner, F., & De Hoog, R. (2010). Consumer-generated versus marketer-generated websites in consumer decision making. *International Journal of Market Research*, 52(2), 231-248.

Brownlee, J. (2017). What is the difference between test and validation datasets. *Machine Learning Mastery*, 14.

Bruner, G. C., & Pomazal, R. J. (1988). Problem recognition: the crucial first stage of the consumer decision process. *The Journal of Services Marketing*, 2(3), 43-53.

Bui, T. D., Hernández-Lobato, J. M., Li, Y., Hernández-Lobato, D., & Turner, R. E. (2015). Training deep Gaussian processes using stochastic expectation propagation and probabilistic backpropagation. *arXiv preprint arXiv:1511.03405.*

Bunn, M. D. (1993). Taxonomy of buying decision approaches. *Journal of Marketing*, *57*(1), 38-56.

Calastri, C., Hess, S., Daly, A., & Carrasco, J. A. (2017). Does the social context help with understanding and predicting the choice of activity type and duration? An application of the Multiple Discrete-Continuous Nested Extreme Value model to activity diary data. *Transportation Research Part A: Policy and Practice*, 104, 1-20.

Cheng, J., Dong, L., & Lapata, M. (2016). Long short-term memory-networks for machine reading. *arXiv preprint arXiv:1601.06733*.

Chiang, C. C., King, B. E., & Nguyen, T. H. (2012). Information searching and the travel behaviours of MICE travellers: A cross-cultural study. *International Journal of Tourism Research*, 14(2), 103-115.

Chiang, W. Y. K., Zhang, D., & Zhou, L. (2006). Predicting and explaining patronage behavior toward web and traditional stores using neural networks: a comparative analysis with logistic regression. *Decision Support Systems*, 41(2), 514-531.

Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., & Xie, W. (2018). Opportunities and obstacles for deep learning in biology and medicine. *Journal of The Royal Society Interface*, *15*(141), 20170387.

Cho, M., Heo, S., & Lee, Y. (2016). Optimal Construction of Renewable Energy Portfolio in use of MDCEV Model. *Korean Energy Economic Review*, 15(2), 55-88.

Choi, J. Y. (2009). *Demand forecasting for new product using ex-ante simulations and market segmentation* (Doctoral dissertation, Ph. D Dissertaion, Seoul National University, South Korea).

Chorus, C. G., Molin, E. J., & van Wee, B. (2006). Travel information as an instrument to change car drivers' travel choices: a literature review. *European Journal of Transport and Infrastructure Research*, 6(4).

Ciresan, D. C., Meier, U., Masci, J., Gambardella, L. M., & Schmidhuber, J. (2011). Flexible, high performance convolutional neural networks for image classification. In *Twenty-Second International Joint Conference on Artificial Intelligence*.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4), 303-314.

Dasgupta, C. G., Dispensa, G. S., & Ghose, S. (1994). Comparing the predictive performance of a neural network model with some traditional market response models. *International Journal of Forecasting*, 10(2), 235-244.

Danaher, P. J. (1997). Using conjoint analysis to determine the relative importance of service attributes measured in customer satisfaction surveys. *Journal of Retailing*, 73(2), 235-260.

Davis, F. (1993). User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *International Journal of Man-Machine Studies*. 38, (3), 475– 487.

Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems: Theory and results* (Doctoral dissertation, Massachusetts Institute of Technology).

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319-340.

Davis, F. D., & Venkatesh, V. (1996). A critical assessment of potential measurement biases in the technology acceptance model: three experiments. *International Journal of Human-Computer Studies*, 45(1), 19-45.

Davis, F., Bogozzi, R. & Warshaw, P. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35, 982-1003.

Delen, D., Walker, G., & Kadam, A. (2005). Predicting breast cancer survivability: a comparison of three data mining methods. *Artificial Intelligence in Medicine*, 34(2), 113-127.

Dellaert, B. G., & Häubl, G. (2012). Searching in choice mode: consumer decision processes in product search with recommendations. *Journal of Marketing Research*, 49(2), 277-288.

Deloitte. (2017). Car sharing in Europe—business models, national variations and upcoming disruptions.

Dewey, J. (1910). How we think. Boston. *MA: DC Heath*.

Dheir, I. M., Mettleq, A. S. A., Elsharif, A. A., & Abu-Naser, S. S. (2020). Classifying Nuts Types Using Convolutional Neural Network. *International Journal of Academic Information Systems Research*. 3(12). 12-18.

Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of Biomedical Informatics*, 35(5-6), 352-359.

Egbue, O., & Long, S. (2012). Barriers to widespread adoption of electric vehicles: An analysis of consumer attitudes and perceptions. *Energy Policy*, 48, 717-729.

Engel, J. F., Kollat, D. T., & Blackwell, R. D. (1973). Consumer behavior. *New York*..Ewing, G., & Sarigöllü, E. (2000). Assessing consumer preferences for clean-fuel vehicles: A discrete choice experiment. *Journal of Public Policy & Marketing*, 19(1), 106-118.

Fanning, K., Cogger, K. O., & Srivastava, R. (1995). Detection of management fraud: a neural network approach. *Intelligent Systems in Accounting, Finance and Management*, 4(2), 113-126.

Ferdous, N., Pinjari, A. R., Bhat, C. R., & Pendyala, R. M. (2010). A comprehensive analysis of household transportation expenditures relative to other goods and services: an application to United States consumer expenditure data. *Transportation*, 37(3), 363-390.

Fiksdal, A. S., Kumbamu, A., Jadhav, A. S., Cocos, C., Nelsen, L. A., Pathak, J., & McCormick, J. B. (2014). Evaluating the process of online health information searching: a qualitative approach to exploring consumer perspectives. *Journal of Medical Internet Research*, 16(10).

Fishbein, M., Ajzen, I. (1975). Belief, attitude, intention, and behavior: An introduction to theory and research. Reading, Mass; Don Mills, Ontario: Addison-Wesley Pub. Co.

Földiák, P. (1991). Learning invariance from transformation sequences. *Neural Computation*, *3*(2), 194-200.

Foxall, G. (2005). *Understanding consumer choice*. *UK*. Springer.

Fukushima, K. (2007). Neocognitron. *Scholarpedia*, *2*(1), 1717.

Gan, C., Limsombunchao, V., Clemes, M. D., & Weng, Y. Y. (2005). Consumer choice prediction: Artificial neural networks versus logistic models. *Commerce Division Discussion Paper*, No. 104

Gaudart, J., Giusiano, B., & Huiart, L. (2004). Comparison of the performance of multi-layer perceptron and linear regression for epidemiological data. *Computational Statistics & Data Analysis*, 44(4), 547-570.

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *In Proceedings of the IEEE conference on computer vision and pattern recognition* pp. 580-587.

Glorot, X., Bordes, A., & Bengio, Y. (2011, June). Deep sparse rectifier neural networks. *In Proceedings of the fourteenth international conference on artificial intelligence and statistics* pp. 315-323.

Goodhue, D. & Thompson, R. (1995). Task technology fit and individual performance. *MIS Quarterly*, 19, 213-236.

Graves, A. (2012). Long short-term memory. *In Supervised sequence labelling with recurrent neural networks*. pp. 37-45.

Hackbarth, A., & Madlener, R. (2013). Consumer preferences for alternative fuel vehicles: A discrete choice analysis. *Transportation Research Part D: Transport and Environment*, 25, 5-17.

Haider, W., & Ewing, G. O. (1990). A model of tourist choices of hypothetical Caribbean destinations. *Leisure Sciences*, 12(1), 33-47.

Han, J., & Moraga, C. (1995, June). The influence of the sigmoid function parameters on the speed of backpropagation learning. *In International Workshop on Artificial Neural Networks.* pp. 195-201.

Hardman, S., Jenn, A., Tal, G., Axsen, J., Beard, G., Daina, N., & Plötz, P. (2018). A review of consumer preferences of and interactions with electric vehicle charging infrastructure. *Transportation Research Part D: Transport and Environment*, 62, 508-523.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *In Proceedings of the IEEE International Conference on Computer Vision.* pp. 1026-1034.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(9), 1904-1916.

Hecht-Nielsen, R. (1992). Theory of the backpropagation neural network. In Neural networks for perception. pp. 65-93. Academic Press.

Hernández-Lobato, J. M., & Adams, R. (2015, June). Probabilistic backpropagation for scalable learning of Bayesian neural networks. In *International Conference on Machine Learning* pp. 1861-1869.

Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(02), 107-116.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.

Hornik, K., Stinchcombe, M., & White, H. (1990). Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks*, 3(5), 551-560.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*(1), 215-243.

Hubel, D. H., & Wiesel, T. N. (2004). *Brain and visual perception: the story of a 25-year collaboration*. Oxford University Press.

Huber, J. & Zwerina, K. (1996). The Importance of Utility Balance in Efficient Choice Set Designs, *Journal of Marketing Research*, 33, 307-317

Hurvich, C. M., & Tsai, C. L. (1995). Model selection for extended quasi-likelihood models in small samples. *Biometrics*, 1077-1084.

Hwang, Y., Gretzel, U., Xiang, Z., & Fesenmaier, D. R. (2006). Information search for travel decisions. *Destination recommendation systems: Behavioral foundations and applications*, 42(4), 357-371.

IBM SPSS Statistics for Windows, Version 23.0.

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167.*

Izenman, A. J. (2008). Modern multivariate statistical techniques. *Regression, Classification and Manifold Learning*, 10, 978-0.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning.* Vol. 112, pp. 3-7. Springer.

Jang, J. (2014). *Analysis of the Social Acceptance in Electric Power Facility Siting* (Doctoral dissertation, Seoul National University).

Jang, W. (2015). Transport Service Status and Prospect in the Era of Sharing Economy. *Monthly KOTI Magazine on Transport*, 5-11.

Jeng, J. M., & Fesenmaier, D. R. (1996). A neural network approach to discrete choice modeling. *Journal of Travel & Tourism Marketing*, 5(1-2), 119-144.

Jian, S., Rashidi, T. H., & Dixit, V. (2017). An analysis of carsharing vehicle choice and utilization patterns using multiple discrete-continuous extreme value (MDCEV) models. *Transportation Research Part A: Policy and Practice*, 103, 362-376.

Johansson, F., Shalit, U., & Sontag, D. (2016, June). Learning representations for counterfactual inference. In *International Conference on Machine Learning*. pp. 3020-3029.

Jylänki, P., Nummenmaa, A., & Vehtari, A. (2014). Expectation propagation for neural networks with sparsity-promoting priors. *The Journal of Machine Learning Research*, 15(1), 1849-1901.

Kim, J., Allenby, G. M., & Rossi, P. E. (2002). Modeling consumer demand for variety. *Marketing Science*, 21(3), 229-250.

Kim, J., Park, K., & Koh, J. (2015). Socioeconomic Impact of Carsharing. *Issue & Analysis*. (183). 1-25.

Kim, J., Lee, J., & Ahn, J. (2016). Reference-dependent preferences on smart phones in South Korea: Focusing on attributes with heterogeneous preference direction. *Computers in Human Behavior*, 64, 393-400.

Koop, G. M. (2003). *Bayesian Econometrics*. John Wiley & Sons Inc.

Krizhevsky, A., & Hinton, G. (2009). Learning multiple layers of features from tiny images.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *In Advances in Neural Information Processing Systems*. pp. 1097-1105.

Kumar, A., Rao, V. R., & Soni, H. (1995). An empirical comparison of neural network and logistic regression models. *Marketing Letters*, 6(4), 251-263.

Lavieri, P. S., Garikapati, V. M., Bhat, C. R., Pendyala, R. M., Astroza, S., & Dias, F. F. (2017). Modeling individual preferences for ownership and sharing of autonomous vehicle technologies. *Transportation Research Record*, 2665(1), 1-10.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.

Lee, D., Derrible, S., & Pereira, F. C. (2018). Comparison of four types of artificial neural network and a multinomial logit model for travel mode choice modeling. *Transportation Research Record*, 2672(49), 101-112.

Lee, D., Kang, S., & Shin, J. (2017). Using deep learning techniques to forecast environmental consumption level. *Sustainability*, 9(10), 1894.

Lenard, M. J., Alam, P., & Madey, G. R. (1995). The application of neural networks and a qualitative response model to the auditor's going concern uncertainty decision. *Decision Sciences*, 26(2), 209-227.

Leong, L. Y., Hew, T. S., Ooi, K. B., Lee, V. H., & Hew, J. J. (2019). A hybrid SEM-neural network analysis of social media addiction. *Expert Systems with Applications*, 133, 296-316.

Li, J., Lu, X., Cheng, K., & Liu, W. (2020). StepReg-package: Stepwise Regression Analysis.

Lindsey, C., & Sheather, S. (2010). Variable selection in linear regression. *The Stata Journal*, 10(4), 650-669.

Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. *In Proceedings.* Vol. 30, No. 1, p. 3.

MacKay, D. J. (1992). A practical Bayesian framework for *backpropagation networks. Neural computation*, 4(3), 448-472.

Malhotra, P., Vig, L., Shroff, G., & Agarwal, P. (2015, April). Long short-term memory networks for anomaly detection in time series. *In Proceedings* Vol. 89, pp. 89-94. Presses Universitaires de Louvain.

Mallows, C. L. (2000). Some comments on Cp. *Technometrics*, 42(1), 87-94.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50(4), 370.

Matsugu, M., Mori, K., Mitari, Y., & Kaneda, Y. (2003). Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Networks*, *16*(5-6), 555-559.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115-133.

Minka, T. P. (2013). Expectation propagation for approximate Bayesian inference. *arXiv preprint arXiv:1301.2294.*

Minsky, M. L., & Papert, S. (1988). Perceptrons: An introduction to computational geometry (Exp. ed.). *The MIT Press:* Cambridge.

MOLIT Statistics System. (2020). Vehicle Registration Statistics, February, 2020. Retrieved from http://stat.molit.go.kr .

Moon, H., Park, S. Y., Jeong, C., & Lee, J. (2018). Forecasting electricity demand of electric vehicles by analyzing consumers' charging patterns. *Transportation Research Part D: Transport and Environment*, 62, 64-79.

Moraes, R., Valiati, J. F., & Neto, W. P. G. (2013). Document-level sentiment classification: An empirical comparison between SVM and ANN. *Expert Systems with Applications*, 40(2), 621-633

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. *In Proceedings of the 27th international conference on machine learning* (ICML-10) (pp. 807-814).

National Air Pollutants Emission Service. (2020). 2016 Pollutant emissions by sector. Retrieved from https://airemiss.nier.go.kr/.

Neal, R. M. (2012). *Bayesian learning for neural networks* (Vol. 118). Springer Science & Business Media.

Nicholas, D., Huntington, P., & Watkinson, A. (2003, March). Digital journals, big deals and online searching behaviour: a pilot study. *In Aslib Proceedings*. MCB UP Ltd.

Paliwal, M., & Kumar, U. A. (2009). Neural networks and statistical techniques: A review of applications. *Expert Systems with Applications*, 36(1), 2-17.

Papachristou, N., Miaskowski, C., Barnaghi, P., Maguire, R., Farajidavar, N., Cooper, B., & Hu, X. (2016, November). Comparing machine learning clustering with latent class analysis on cancer symptoms' data. *In 2016 IEEE Healthcare Innovation Point-Of-Care Technologies Conference* (HI-POCT) (pp. 162-166). IEEE.

Paredes, M., Hemberg, E., O'Reilly, U. M., & Zegras, C. (2017, June). Machine learning or discrete choice models for car ownership demand estimation and prediction? *In 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems* (MT-ITS) (pp. 780-785). IEEE.

Park, J., & Park, J. (2015). Impact of the Car-sharing Service on Transport Demand and Taxi Service. *Journal of Transport Research*, 22(2), 19-34.

Pellegrini, A., & Scagnolari, S. (2019). The relationship between length of stay and land transportation mode in the tourism sector: A discrete–continuous framework applied to Swiss data. *Tourism Economics*. 1-17.

Philip, K., Lane, K. K., Abraham, K., & Mithileshwar, J. (2009). *Marketing Management-A South Asian Perspective*. Prentice Hall Edu.

Razi, M. A., & Athappilly, K. (2005). A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. *Expert Systems with Applications*, 29(1), 65-74.

Reed, S., Lee, H., Anguelov, D., Szegedy, C., Erhan, D., & Rabinovich, A. (2014). Training deep neural networks on noisy labels with bootstrapping. *arXiv preprint arXiv:1412.6596.*

Rezvani, Z., Jansson, J., & Bodin, J. (2015). Advances in consumer electric vehicle adoption research: A review and research agenda. *Transportation Research Part D: Transport and Environment*, 34, 122-136.

Ripley, B. D. (2007). *Pattern recognition and neural networks*. Cambridge University Press.

Rogers, E. M. (2010). *Diffusion of innovations*. Simon and Schuster.

Rosenblatt, F. (1957). *The perceptron, a perceiving and recognizing automaton Project Para*. Cornell Aeronautical Laboratory.

Rosenblatt, F. (*1961). Principles of neurodynamics. perceptrons and the theory of brain mechanisms* (No. VG-1196-G-8). Cornell Aeronautical Lab Inc Buffalo NY.

Rossi, P. E., Allenby, G. M., & McCulloch, R. (2012). *Bayesian statistics and marketing*. John Wiley & Sons.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., & Berg, A. C. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3), 211-252.

Sak, H., Senior, A. W., & Beaufays, F. (2014). Long short-term memory recurrent neural network architectures for large scale acoustic modeling.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461-464.

Seo, J., Sheok, C., & Lee, S. (2017). Analysis of the Effect of Carsharing on Transportation Usage from the Perspective of Sharing Economy: Focused on 2017 Incheon City Carsharing Survey. *Journal of Korea Planning Association*, 53(2), pp.107-121.

Sheather, S. (2009). *A modern approach to regression with R.* Springer Science & Business Media.

Shin, J., Hong, J., Jeong, G., & Lee, J. (2012). Impact of electric vehicles on existing car usage: A mixed multiple discrete–continuous extreme value model approach. *Transportation Research Part D: Transport and Environment*, 17(2), 138-144.

Shin, J., Hwang, W. S., & Choi, H. (2019). Can hydrogen fuel vehicles be a sustainable alternative on vehicle market? : Comparison of electric and hydrogen fuel cell vehicles. *Technological Forecasting and Social Change*, 143, 239-248.

Shin, J., Lim, T., Kim, M. Y., & Choi, J. Y. (2018). Can next-generation vehicles sustainably survive in the automobile market? Evidence from ex-ante market simulation and segmentation. *Sustainability*, 10(3), 607.

Song, J. H., Venkatesh, S. S., Conant, E. A., Arger, P. H., & Sehgal, C. M. (2005). Comparative analysis of logistic regression and artificial neural network for computer-aided diagnosis of breast masses. *Academic Radiology*, 12(4), 487-495.

Spissu, E., Pinjari, A. R., Pendyala, R. M., & Bhat, C. R. (2009). A copula-based joint multinomial discrete–continuous model of vehicle type choice and miles of travel. *Transportation*, 36(4), 403-422.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1), 1929-1958.

Sun, Y., Wang, X., & Tang, X. (2015). Deeply learned face representations are sparse, selective, and robust. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* pp. 2892-2900.

Svozil, D., Kvasnicka, V., & Pospichal, J. (1997). Introduction to multi-layer feed-forward neural networks. *Chemometrics and Intelligent Laboratory Systems*, 39(1), 43-62.

Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research*, 1(1), 77-100.

Thieme, R. J., Song, M., & Calantone, R. J. (2000). Artificial neural network decision support systems for new product development project selection.

Thyne, M., Lawson, R., & Todd, S. (2006). The use of conjoint analysis to assess the impact of the cross-cultural exchange between hosts and guests. *Tourism Management*, 27(2), 201-213.

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

Train, K. E., & Winston, C. (2007). Vehicle choice behavior and the declining market share of US automakers. *International Economic Review*, 48(4), 1469-1496.

Train, K., & Sonnier, G. (2005). *Mixed logit with bounded distributions of correlated partworths. In Applications of simulation methods in environmental and resource economics.* pp. 117-134. Springer.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., & Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems. pp. 5998-6008.

Venkatesan, R., & Li, B. (2017). *Convolutional neural networks in visual computing: a concise guide*. CRC Press.

Venkatesh, V., Bala, H. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, 39 (2), 273-312.

Venkatesh, V., Davis, F. (1996). A model of the antecedents of perceived ease of use: Development and test. *Decision Sciences*, 27(3), 451-481.

Venkatesh, V., Davis, F. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, 46 (2), 186-204.

Wang, H., & Raj, B. (2017). On the origin of deep learning. *arXiv preprint arXiv:1702.07800.*

Wang, N., Li, S., Gupta, A., & Yeung, D. Y. (2015). Transferring rich feature hierarchies for robust visual tracking. *arXiv preprint arXiv:1501.04587*.

Wang, S., & Zhao, J. (2018). Framing discrete choice model as deep neural network with utility interpretation. *arXiv preprint arXiv:1810.10465*.

Wang, S., Fan, J., Zhao, D., Yang, S., & Fu, Y. (2016). Predicting consumers' intention to adopt hybrid electric vehicles: using an extended version of the theory of planned behavior model. *Transportation*, 43(1), 123-143.

Wang, S., Li, J., & Zhao, D. (2017). The impact of policy measures on consumer intention to adopt electric vehicles: Evidence from China. *Transportation Research Part A: Policy and Practice*, 105, 14-26.

Wilson, T. D. (1997). Information behavior: an interdisciplinary perspective. *Information Processing & Management*, 33(4), 551-572.

Wolinetz, M., & Axsen, J. (2017). How policy can build the plug-in electric vehicle market: Insights from the REspondent-based Preference and Constraints (REPAC) model. *Technological Forecasting and Social Change*, Vol. 117, 238-250.

Woo, J., Choi, J. Y., Shin, J., & Lee, J. (2014). The effect of new media on consumer media usage: An empirical study in South Korea. *Technological Forecasting and Social Change*, 89, 3-11.

Xiong, H. Y., Barash, Y., & Frey, B. J. (2011). Bayesian prediction of tissue-regulated splicing using RNA sequence and cellular context. *Bioinformatics*, 27(18), 2554-2562.

Xu, B., Wang, N., Chen, T., & Li, M. (2015). Empirical evaluation of rectified activations in convolutional network. *arXiv preprint arXiv:1505.00853.*

Yoo, C. (2019, July 18). Car sharing industry to introduce 'electric and hybrid cars'… 16% are eco-friendly models. *eToday*.

Yoon, B., Rhee, J., & Kim, J. (2010). A Methodology of Analyzing Effects of an Environmentally Friendly Transportation Service: The Case of Carsharing Service. *IE interfaces*, 23(4), 286-299.

Zeithaml, V. A., Bitner, M. J., Gremler, D. D., & Pandit, A. (2000). *Services marketing: Integrating customer focus across the firm.*

조만석, 허성윤, & 이용길. (2016). 혼합 MDCEV 모형을 이용한 발전사업자의 최적 재생에너지 포트폴리오 구성에 관한 연구. *에너지경제연구*, 15(2), 55-88.

최새로나, 이수일 (2018). 카셰어링 교통안전 대책 마련을 위한 연구. *한국교통안전공단*.

Appendix: Survey used in the analysis

차세대 교통서비스에 대한 국민 선호 및 사용 행동 조사

안녕하십니까?

한국과학기술기획평가원 사회혁신정책센터에서는 차세대 교통서비스에 대한 정책연구를 위하여 일반 국민 여러분의 **미래 교통수요에 대한 인식 및 선호와 사용 행동**을 알아보고 있습니다. 본 질문에는 맞고 틀리는 답이 없으며, 이런 의견을 갖고 있는 사람이 몇 퍼센트 (%)라는 식으로 통계를 내는 데에만 사용됩니다. 그 외의 목적에는 절대로 사용되지 않으니 평소 생각대로 응답해 주시면 됩니다. 또한, 귀하께서 응답해 주신 내용은 통계법 (제33조)에 따라 통계목적으로만 사용되며, 귀하의 의견은 철저히 보호됨을 약속드립니다. 바쁘시겠지만, 조사에 협조해 주시면 대단히 감사드리겠습니다.

2019년 12월

문의: 문형빈 부연구위원 (02-589-2827)

먼저, 응답자 정보에 대한 질문입니다.

SQ1. 귀하의 **성별**은 어떻게 되십니까?:
1) 남자 2) 여자

SQ2. **귀하의 현재 나이**는 만으로 어떻게 되십니까? 출생년도를 기입해주세요. ______년 (대상자: 20-50대)

SQ3. 귀하께서 거주하시는 **지역**을 선택해 주세요. [지도 삽입 / 2개]
SQ4. 귀하의 거주지역을 **읍/면/동 단위를 응답**해 주십시오.

읍/면/동	

SQ5. ① 현재 귀하와 **함께 살고있는 가족은 모두 몇 명**입니까? **응답자 본인을 포함한 가족 수**를 응답해 주시고,
② 귀하와 함께 살고있는 가족 중, **만 60세 이상, 초중고생, 미취학아동과 그 외 가족 수**를 응답해 주십시오.

① 같이 살고 있는 가족 수 (응답자 **본인 포함**) **[range: 1 이상]**	명				
② 같이 살고 있는 가족구성 **(합계가 ①과 같음)**	응답자 본인	만 60세 이상	초중고생	미취학 아동	그 외
	1 명	명	명	명	명

문1. 귀하께서는 **운전면허**를 보유하고 있습니까?
1) 예(보유) 2) 아니오(미보유) ➔ **대상아님**

▣ 질문 응답시 주의사항

1. 질문지는 맨 앞 페이지부터 **순서대로 응답**해 주십시오. 특별한 언급이 없다면, **모든 질문에 빠짐없이 응답**해 주시기 바랍니다.
2. 질문에 응답하시기 전에 질문 앞에 제시된 **설명문을 잘 읽고, 숙지**하신 후 응답해 주시기 바랍니다.

A. 자동차 소유 및 사용행태 조사

다음 질문들은 귀하(댁)의 자동차 보유 현황 및 운전 행동을 묻는 질문입니다.

문2. **(전체 응답자)** 현재 귀하나 귀댁에서는 **자동차를 보유**하고 있습니까?
1. 예 (있다) _____대➔ 문 2-1로 이동
2. 아니오 (없다) ➔ **문3.로 가십시오**

문2-1. **(문1.에서 1. 자동차 보유에 응답한 응답자)** 현재 귀댁에서 **보유하고 있는 자동차의 사양과 사용행태를** 응답해 주십시오.
[PROG: 문2에서 1대는 주사용자동차만 제시 / 2대 이상은 주사용, 부사용 모두 제시]

	주사용 자동차	**부사용 자동차**
01. 신차/중고차 여부	1. 신차 구매 2. 중고차 구매	1.신차 구매 2. 중고차 구매
02. 구매 시점	년	년
03. 자동차 모델명 (구체적으로 응답해 주십시오.)		
04. 차량 제조년도(연식)	년	년
05. 유종	① 휘발유 ② 경유 ③ LPG ④ 하이브리드 ⑤ 전기 ⑥ 수소	① 휘발유 ② 경유 ③ LPG ④ 하이브리드 ⑤ 전기 ⑥ 수소
06. 구입가격(단위 : 만원)	만원	만원
07. 연평균 주행거리	___만 ___천 KM	___만 ___천 KM

문2-2. 귀하께서는 **주로 어떤 상황**에서 차량에 **주유(LPG/전기차일 경우 충전)를 하십니까?** 평소 주유(충전) 패턴을 고려하시어, 아래 문구가 귀하의 주유(충전) 패턴과 얼마나 일치하는지 **응답**해 주시기 바랍니다.

	전혀 그렇지 없다	대체로 그렇지 않다	보통이다	대체로 그런 편이다	항상 그렇다
주유(충전)가 **꼭 필요한 상황은 아니나, 주로 방문하는 주유소(충전소)**를 지날 때 **미리 주유(충전)**한다.	1	2	3	4	5
주유(충전)가 **꼭 필요한 상황은 아니나,** 주로 방문하는 주유소(충전소)보다 **요금이 저렴한 곳을 발견하면 미리 주유(충전)**한다.	1	2	3	4	5
반드시 주유(충전)가 필요한 상황이 되면, **곧바로 가장 가까운 주유소**를 방문하여 주유(충전)한다.	1	2	3	4	5
반드시 주유(충전)가 필요한 상황이 되면, 주변에 있는 주유소(충전소) 중 **요금이 저렴한 곳을 찾아서** 주유(충전)한다	1	2	3	4	5

문2-3. 귀하께서는 연료 잔여량이 얼마 이하로 떨어졌을 때 **주유(충전)가 필요하다고 느끼십니까? (주유소(충전소) 방문을 고려하십니까?)** ※ 연료가 가득 찼을 때를 100%로 고려하고 응답해 주십시오.

_______%

문3. 귀하께서는 **주로 어떤 상황**에서 **스마트폰 배터리를 충전하십니까?** 평소 충전 패턴을 고려하시어, 아래 문구가 귀하의 스마트폰 배터리 충전 패턴과 얼마나 일치하는지 **응답**해 주시기 바랍니다.

※ 스마트폰이 없는 경우 일반 휴대폰, 또는 휴대하는 기타 전자기기(블루투스 이어폰, MP3 등)를 기준으로 응답해 주시기 바랍니다.

	전혀 그렇지 없다	대체로 그렇지 않다	보통이다	대체로 그런 편이다	항상 그렇다
충전이 **꼭 필요한 상황은 아니나, 가능할 때마다 미리 충전한다. (보조배터리를 이용한 충전 포함)**	1	2	3	4	5
일과 시간 동안 **일상적으로 방문하는 장소(직장, 학교 등)**에서 **정기적으로 충전**한다.	1	2	3	4	5
집(숙소)에서 수면(휴식)을 취하는 동안 **정기적으로 충전**한다.	1	2	3	4	5
반드시 충전이 필요한 상황이 되면, 충전이 가능한 장소를 찾아서 충전한다.	1	2	3	4	5

문3-1. 귀하께서는 **배터리 전여량**이 얼마 이하로 떨어졌을 때 충전 **필요하다고 느끼십니까? (충전 가능 장소/방법을 탐색하십니까?)** ※ 배터리가 완충되었을 때를 100%로 고려하고 응답해 주십시오.

		%

(문2-1의 05 유종에서 4~6번 응답이 있는 경우만)

문4. 귀하께서는 현재 **친환경 자동차 (하이브리드, 전기차, 수소연료전지자동차)를 사용**하고 계십니까?

1) 예 (친환경 자동차를 사용한다) 2) 아니오 (친환경 자동차는 사용하지 않는다) —▶ **문5.로 가십시오**

(문4.에서 1. 친환경 자동차를 사용한다에 응답한 응답자)

문4-1. 귀하께서 **친환경 자동차를 구매한 동기**는 무엇 무엇입니까? **모두 응답**해 주십시오. **(순위로 응답)**

1. 환경오염 방지
2. 연료비 절감
3. 구매 보조금
4. 충전 인프라 확보
5. 기타(________________)

➔ **B파트로 가십시오**

(문4.에서 2. 친환경 자동차를 사용하지 않는다에 응답한 응답자)

문5. 귀하께서는 향후 친환경 자동차를 앞으로 **구매할 의향**이 있으십니까?

1. 예 (있다) ➔ **B파트로 가십시오**
2. 아니오 (없다) ➔ 문5-1로 이동

문5-1. 귀하께서 **친환경 자동차를 구매를 고려하지 않는 이유**는 무엇입니까? **모두 응답**해 주십시오.

비구매 요인
1. 높은 구매비용
2. 충전소 인프라 부족
3. 짧은 완충 시 주행거리
4. 원하는 차급(차량 모델) 부족
5. 주변에서 타는 사람이 없어서
6. 기술적 불확실성
7. 내연기관차 대비 다른 주행감
8. 제반 제도 부족(보험, AS 불확실성)
9. 복잡한 정부 지원 프로세스
10. 기타( )

문5-2. 귀하께서 **친환경 자동차를 구매를 고려하지 않는 이유**에 대해 순위를 매겨 주세요.

[선택한 만큼 순위형 / 문5-1 선택 보기만 제시]

비구매 요인
1. 높은 구매비용
2. 충전소 인프라 부족
3. 짧은 완충 시 주행거리
4. 원하는 차급(차량 모델) 부족
5. 주변에서 타는 사람이 없어서
6. 기술적 불확실성
7. 내연기관차 대비 다른 주행감
8. 제반 제도 부족(보험, AS 불확실성)

9. 복잡한 정부 지원 프로세스	
10. 기타(**파이핑**)	

B. 용도별 이동수단 사용행태 조사

▣ 이동수단 사용행태 문항에 대한 설명문

응답요령	뒤이어 제시되는 문항들에서는 귀하의 평소 **이동수단 사용행태**를 알아보고자 합니다. - 이동수단으로는 **① 자가용 차량, ② 카풀, ③ 카셰어링, ④ 기사제공 차량공유 서비스, ⑤ 버스, ⑥ 지하철, ⑦ 택시, ⑧ 기차**를 고려해주시면 되며, 도보나 자전거, 오토바이 이용은 제외하고 응답해주시면 됩니다. - **버스와 환승이 가능한 도시철도, 광역철도 등은 모두 지하철**로 고려하여 주시고, **기차의 경우 버스와 환승이 불가능한 KTX, SRT, 새마을호, 무궁화호, 통근열차 등**이 포함됩니다. - **카풀, 카셰어링 등 공유차량 서비스는 최근 확산되고 있는 새로운 유형의 이동수단**으로, 자세한 설명은 아래와 같습니다.
카셰어링 설명문	- 카셰어링은 **차량 대여 서비스**로서 국내에서는 **쏘카, 그린카 등이 대표 서비스**로 알려져 있습니다. - **렌터카 서비스와 비슷**하지만 아래와 같은 차별적인 특징이 있습니다. · **짧은 기간**만도 대여할 수 있습니다. (10분 단위 대여 가능) · **스마트폰 어플을 통해서 대여**가 가능합니다. · 일반적으로 **대여/반납 과정이 무인화**되어 있습니다. (관리자가 없이 특정 장소(공영주차장, 건물주차장 등)에 주차된 차량에 직접 가서 대여해야 하며, 이용 후 지정된 장소로 차량을 직접 가져다 놓아야 합니다.) · 최초 이용 시 스마트폰 어플에 운전면허증을 등록해두면 **별도 검증절차 없이 차량 대여가 가능**합니다. (타인 계정을 이용한 불법 대여가 사회적 이슈화 되고 있습니다. (청소년 등 무면허자의 차량 대여 문제 발생) · 반납할 때 주유를 할 필요가 없으며, **주유가 필요한 경우 차량 안에 비치된 주유 전용 카드를 사용**합니다. · 일반적으로 **1회 대여에 따른 기본요금 + 주행거리 요금 + 사용시간 요금**으로 정산됩니다. - 더 구체적인 설명은 다음의 영상 링크를 참고해주시기 바랍니다. (링크 클릭)
카풀 설명문	- 카풀은 **승차 공유 서비스**로서 국내에서는 **카카오T 카풀, 풀러스,** 해외에서는 **우버 등이 대표 서비스**로 알려져 있습니다. - 카풀을 기본적으로 **일반 운전자와 탑승자를 연결해주는 서비스**이며, 다음과 같은 특징이 있습니다. · 방향이 비슷하거나 목적지가 같은 **운전자와 탑승자(들)이 승차를 공유**하는 서비스입니다. · **탑승자(들)은 운전자에게 일정 금액을 지불**하게 되며 일반적으로 택시보다는 저렴한 요금제를 유지하고 있습니다. · 택시와 달리 **일반인(전문기사 면허 및 사업자 신고 불필요)도 가능**하며 **일반 승용차로 서비스 제공이**

	가능합니다. - 더 구체적인 설명은 다음의 영상 링크를 참고해주시기 바랍니다. (링크 클릭) 
기사 제공 차량공유 서비스 설명문	- 기사 제공 차량공유 서비스는 **차량과 기사를 함께 제공**하는 서비스로서 국내에서는 **타다**가 대표적입니다. - 기사제공 차량공유 서비스(타다 등)는 차량뿐만 아니라 목적지까지 운전을 해줄 기사를 함께 고용하는 서비스로서, 카셰어링 및 택시와는 다른 특성을 가지고 있습니다. · 기본적으로 **11인승 승합차가 배정**됩니다. · 고객이 목적지를 입력하면 **차량과 함께 기사가 제공된다는 점에서 카셰어링과 차별**됩니다. (스마트폰 어플로 출발지(현재위치 등)와 목적지만 입력하면 별도의 절차없이 차량과 기사가 배정됩니다.) **· 택시와 달리 전문기사 면허가 없는 기사(기사 제공 업체에 고용된 직원)가 운전 서비스를 제공합니다.** - 더 구체적인 설명은 다음의 영상 링크를 참고해주시기 바랍니다. (링크 클릭) 

문1. **(전체 응답자)** 귀하는 평소에 어떠한 목적으로 이동수단을 이용하십니까? **모두 응답**해 주십시오. **(순위형/1순위 필수)**
※ 이동수단에는 자가용 차량, 카풀, 카셰어링, 기사제공 차량공유 서비스(타다 등), 버스, 지하철, 택시, 철도 등이 포함됩니다.
아래 목적을 위해 이와 같은 이동수단을 **적어도 한 달에 1회 이상 이용하는 경우** 체크해주시면 됩니다.
① 출퇴근용 ② 사업용/업무용 ③ 레저 및 장거리 여행 ④ 생활 및 일상(쇼핑 등)

문1-1. 목적에 따라 어떻게 이용하시는지 질문에 응답해 주십시오.
[문1에서 응답한 보기만큼 반복]

항목	보기
목적을 위한 이용 이동수단 (복수 선택)	① 자가용 차량 ② 카풀 ③ 카셰어링(쏘카 등) ④ 기사제공 차량공유 서비스(타다 등) ⑤ 버스 ⑥ 지하철 ⑦ 택시 ⑧ 기차 ⑨ 기타(직접 응답: ______)
이동수단 사용주기	① 매일 ② 1주일에 ( )회 정도 ③ 1개월에 ( )회 정도
1회 이동시 평균 이동거리	☐☐☐ km
이동 시기 (평일/주말)	① 주로 평일에 ② 주로 주말에
이동 시간대 (아래 보기 중 선택)	① 순위: ② 순위: ③ 순위: **◉ 시간대 보기** 1. 6시~9시 2. 9시~12시 3. 12시~15시 4. 15시~18시 5. 18시~21시 6. 21시~24시 7. 24시~6시

문2. 다음은 귀하의 공유 차량 이용 의향에 대한 질문입니다.
이동 유형별로 **카풀, 카셰어링, 기사제공 차량공유 서비스를 이동 수단으로서 이용할 의향**을 5점 척도로 응답해 주시기 바랍니다.

◉ 카풀 이용 의향	전혀 이용하지 않을 것이다	이용을 고려하지 않는 편이다	보통이다	이용을 고려하는 편이다	적극적으로 이용할 것이다
출퇴근용	1	2	3	4	5
사업용/업무용	1	2	3	4	5
레저 및 장거리 여행용	1	2	3	4	5
가정/일상 생활용	1	2	3	4	5

◉ 카셰어링(쏘카 등) 이용 의향	전혀 이용하지 않을 것이다	이용을 고려하지 않는 편이다	보통이다	이용을 고려하는 편이다	적극적으로 이용할 것이다
출퇴근용	1	2	3	4	5
사업용/업무용	1	2	3	4	5
레저 및 장거리 여행용	1	2	3	4	5
가정/일상 생활용	1	2	3	4	5

◉ 기사제공 차량공유 서비스(타다 등) 이용 의향	전혀 이용하지 않을 것이다	이용을 고려하지 않는 편이다	보통이다	이용을 고려하는 편이다	적극적으로 이용할 것이다
출퇴근용	1	2	3	4	5
사업용/업무용	1	2	3	4	5
레저 및 장거리 여행용	1	2	3	4	5
가정/일상 생활용	1	2	3	4	5

문3. 다음은 카셰어링, 카풀 등 공유차량과 관련한 이슈에 대해 귀하의 의견을 묻는 질문입니다.
아래 각 이슈가 발생한 이후 **귀하의 공유차량 이용 의향 변화**에 대해 5점 척도로 응답해 주시기 바랍니다.
※ 현재 일어나지 않은 이슈도 포함되어 있으며, 해당 이슈가 발생하였을 때를 가정하여 응답해주십시오.)
해당 이슈가 귀하의 공유차량 이용 의향에 아무런 영향을 주지 않는다면 **'현재와 동일 수준 유지'**를 응답해주시면 됩니다.

◉ 정책/제도 관련	매우 부정적으로 변화	부정적으로 변화	현재와 동일 수준 유지	긍정적으로 변화	매우 긍정적으로 변화
카풀 서비스의 운행 **시간 제한 완화** 설명) 국내 카풀 서비스 도입 초기에는 평**일 출퇴근 시간 혹은 심야 시간과 같이 택시 탑승이 어려운 시간에만 카풀 서비스 이용이 가능하도록 제한**된 바 있습니다.	1	2	3	4	5
카풀 서비스의 운행 **거리 제한 완화**	1	2	3	4	5

※설명) 카풀 서비스의 경우 정책적으로 **장거리 운행이 제한될** 수 있습니다. 예를 들어 프랑스의 경우 최대 이용요금 규제를 통해 장거리 운행을 규제하고 있으며, 중국의 경우 아예 카풀 이용 시 최대 운행거리를 50km로 제한하는 등의 정책이 시행되고 있습니다.					
카풀 탑승 중 사고에 대비한 **카풀 특약 보험 의무가입제도** 시행 **※설명)** 카풀 탑승 중 동승자의 경우 보험 혜택을 제대로 받지 못할 수 있는데, 카풀 특약 보험 의무가입제도를 시행하면 **동승자도 적절한 보험 혜택**을 받을 수 있습니다.	1	2	3	4	5
카풀 운전자 등록 전에 **범죄경력조회**가 가능하도록 한 제도 시행 **※설명)** 택시의 경우 운전자 등록 전에 범죄경력조회를 시행하는 반면, 카풀의 경우 **관련 법규가 아직 미비**한 측면이 있습니다.	1	2	3	4	5
카셰어링 서비스 이용 시 **운전자(대여자) 인증 절차 강화** 설명) 카셰어링의 경우 차량의 대여가 온라인/모바일 상으로 인증이 이루어지고 대면 거래가 이루어지지 않는 특성상 **인증 수단 도용 등을 통해 허위로 대여**가 이루어질 수 있는 문제가 있습니다 (무면허자, 범죄자 등)	1	2	3	4	5

◉ 사건/사고 관련	매우 부정적으로 변화	부정적으로 변화	현재와 동일 수준 유지	긍정적으로 변화	매우 긍정적으로 변화
카풀, 타다 서비스 도입에 대한 택시업계의 **격렬한 반대** (분신 사고 등) **※설명)** 2018년 말 ~2019년 초 택시업계에서 카풀 서비스 도입을 반대하며 **분신을 시도**하는 등 갈등이 격렬한 상황이 있었습니다.	1	2	3	4	5
운전 미숙자(미성년자, 초보운전자 등)의 카셰어링 서비스 이용으로 인한 **교통사고 발생** **※설명)** 카셰어링 서비스의 경우 운전에 대한 진입 장벽을 낮추고, 인증 절차가 대면 대여보다 빈약한 측면이 있기 때문에 **운전 미숙자의 활용으로 인한 교통사고 발생 가능성**이 있습니다.	1	2	3	4	5
카셰어링 **정비 미흡**으로 인한 사고 발생 **※설명)** 카셰어링의 경우 렌터카 등과 같이 차량 정비가 매 이용마다 이루어지는 것이 아니기 때문에 **정비 미흡으로 인한 사고 발생 가능성**이 있습니다.	1	2	3	4	5
카풀 운전자의 **범죄 행위** 발생 **※설명)** 카풀의 경우 운전 기사의 신분이 보장되지 않기 때문에 여성 승객 성추행 등 범죄 발생 가능성이 있습니다.	1	2	3	4	5

문4. 최근 카셰어링, 카풀 등 공유차량 서비스와 관련하여 **여러 현안(카풀과 택시 업계 간 갈등, 공유차량 관련 교통사고 등)이** 있습니다. 다음은 이과 같은 현안을 둘러싼 **정부개입과 관련한 귀하의 의견**을 묻는 질문입니다. 귀하께서 평소 가지고 계신 의견을 고려하여 각 문항에 대해 응답해주시기 바랍니다.

◉ **정부개입 관련 인식 전반**	전혀 그렇지 않다	그렇지 않은 편이다	보통이다	그런 편이다	매우 그렇다
공유차량 서비스 관련 현안 해결을 위한 **정부개입은 타당하다.**	1	2	3	4	5
공유차량 서비스 관련 현안 해결을 위한 **정부개입은 시급하다.**	1	2	3	4	5
정부개입을 통해 공유차량 서비스 관련 **현안이 해결**될 수 있다.	1	2	3	4	5
정부가 나서서 공유차량 서비스 **시장 생태계를 정리하지 않으면** 산업·경제적 손실, 사회적 갈등 등 **사회적 혼란이 빚어질 것이다.**	1	2	3	4	5
정부가 나서지 않더라도 가격 및 서비스 경쟁 등 **시장 논리에 의해 공유차량 서비스 시장은 안정화** 될 것이다.	1	2	3	4	5

◉ **세부 이슈별 인식**	전혀 그렇지 않다	그렇지 않은 편이다	보통이다	그런 편이다	매우 그렇다
정부가 법·제도 재정을 통해 **택시와 카풀의 서비스 영역을 구분**해야 한다. ※ 카풀 운행 시간대 규제, 단거리 이동 시 카풀 이용 금지 등	1	2	3	4	5
공유차량 서비스 등장에 따른 **기존 사업자(택시, 렌트카 등)의 손실**을 정**부가 보조**해야 한다.	1	2	3	4	5
공유차량 업계와 기존 사업자 간 **갈등 해소를 위한 정부 정책 추진 시** 사회적 분위기를 따르기보다는 **객관적인 분석이 우선**되어야 한다.	1	2	3	4	5
정부가 공유차량 **서비스 이용자의 안전을 위한 법제도 마련**이 **시장 안정화를 위한 선제조건**이다.	1	2	3	4	5

문5. 귀하께서는 카풀, 카셰어링 등 공유차량 서비스가 현재보다 더욱 활성화되고 제도 보완, 이해관계자 간 갈등 해소 등 시장이 안정화될 경우 현재 귀하의 용도별 이동거리를 증가시킬 의향이 있으십니까?

1) 예 (있다) ➔ 문5-1로 이동　　2) 아니오 (없다) ➔ **C로 가십시오**

문5-1. 안정화된 카풀, 카셰어링, 기사제공 차량공유 서비스 등 서비스를 도입(첫 고객의 경우)하거나 추가로 활용(기존의 고객)하여 어떤 용도로, 얼마나 사용할 것인지 응답해 주십시오.

※ 사용주기와 평균 이동거리의 경우 문1의 응답과 비교하여 사용량의 변화를 고려하여 응답해 주시기 바랍니다.

	출퇴근용	사업용/업무용	레저 및 장거리 여행용	가정/일상 생활용 (쇼핑 등)
목적을 위한 이용 이동수단 (복수 선택)	① 카풀 ② 카셰어링 (쏘카 등) ③ 기사제공 차량공유 서비스 (타다 등)	① 카풀 ② 카셰어링 (쏘카 등) ③ 기사제공 차량공유 서비스 (타다 등)	① 카풀 ② 카셰어링 (쏘카 등) ③ 기사제공 차량공유 서비스 (타다 등)	① 카풀 ② 카셰어링 (쏘카 등) ③ 기사제공 차량공유 서비스 (타다 등)
신규 서비스 사용주기	① 매일 ② 1주일에 ( )회 정도 ③ 1개월에 ( )회 정도	① 매일 ② 1주일에 ( )회 정도 ③ 1개월에 ( )회 정도	① 매일 ② 1주일에 ( )회 정도 ③ 1개월에 ( )회 정도	① 매일 ② 1주일에 ( )회 정도 ③ 1개월에 ( )회 정도
1회 이동시 평균 이동거리	☐☐☐ km	☐☐☐ km	☐☐☐ km	☐☐☐ km
이동 시기 (평일/주말)	① 주로 평일에 ② 주로 주말에	① 주로 평일에 ② 주로 주말에	① 주로 평일에 ② 주로 주말에	① 주로 평일에 ② 주로 주말에
이동 시간대 (아래 보기 중 선택)	① 순위: ② 순위: ③ 순위:	① 순위: ② 순위: ③ 순위:	① 순위: ② 순위: ③ 순위:	① 순위: ② 순위: ③ 순위:
	◉ **시간대 보기** 1. 6시~9시 2. 9시~12시 3. 12시~15시 4. 15시~18시 5. 18시~21시 6. 21시~24시 7. 24시~6시			

C. 차량 기반 교통서비스 선호 조사

다음은 귀하께서 정기적인 이동(출퇴근, 등하교, 주기적인 업무를 위한 이동 등)을 위해 이용하실 수 있는 카풀, 카셰어링, 택시 등의 차량 기반 교통서비스에 대한 선호도를 측정하는 질문입니다.
아래는 차량 기반 교통수단과 관련된 속성 수준에 대한 설명입니다. 제시된 속성 설명을 잘 숙지하고 응답해 주시기 바랍니다.

▣ 차량 기반 교통서비스 속성 설명문

속성		속성 설명 및 수준
1. 주행거리당 소요 비용	설명	**교통서비스 이용 시 발생하는 모든 비용을 합산하여 km당 요금으로 환산함** - 카셰어링: 대여요금 + 주행요금(주행거리에 비례한 요금) - 카풀: 호출비용(기본료) + 주행요금(주행거리에 비례한 요금) - 택시: 기본료 + 주행요금(주행거리에 비례한 요금)
	수준 (3개)	① 200원/km ② 500원/km ③ 800원/km
2. 연료종류	설명	**차량의 연료 종류로서, 휘발유, 경유, LPG, 전기, 하이브리드로 구분됨**
	수준 (5개)	① 휘발유 ② 경유 ③ LPG ④ 전기 ⑤ 하이브리드
3. 차종	설명	**차량의 외형적 특성으로서, 경차·소형차, 준중형차·중형차, 대형차, SUV·RV로 구분됨**
	수준 (3개)	①경차·소형차·준중형차 ②중형차·대형차 ③SUV·RV
4. 대기 및 이동 시간	설명	**차량을 이용하기 위해 대기하는 평균 시간과 이동하는 시간을 나타냄** - 카셰어링 서비스는 혼잡한 시간대에 배차예약이 몰려 경우 차량을 당장 사용하지 못할 수 있음 - 카셰어링 이용을 위해서는 지정된 서비스 센터까지 차를 픽업하기 위해 직접 이동해야 함 - 택시 및 카풀의 경우 혼잡한 시간대에 배차 대기시간이 존재할 수 있음
	수준 (4개)	① 5분 ② 10분 ③ 15분 ④ 20분

5. 운전자 및 차량 유형	**5-1. 운전자**	**설명**	**특정 수송 모드를 이용할 때 자동차의 운전자가 누군지를 나타냄** - 카셰어링의 경우 직접 운전하는 수송 모드임 - 카풀의 경우 본인과 전문기사가 아닌 일반 운전자가 운행하는 수송 모드임 - 택시의 경우 전문기사가 운전하는 수송 모드임
		수준 (3개)	① 본인 (직접 운전) ② 일반 운전자 (전문기사 및 본인 이외 운전자) ③ 전문기사 (신분이 검증된 전문 운전자)
	5-2. 차량	**설명**	**차량의 소유 및 등록된 타입을 나타냄**
		수준 (2개)	① 영업용 (택시, 카셰어링 등 특정 수송서비스를 제공하기 위한 차량) ② 비영업용 (일반 운전자가 소유한 자차)

[문항당 설명 고정 제시]

1. 다음부터 설명 드린 6개의 속성을 조합하여 구성한 **차량 기반 교통서비스에 대한 선호를 묻는 질문이 제시됩니다.**
   ☞ 제시한 가상의 대안은 **차량 기반 교통서비스**로서 **카셰어링 서비스, 카풀 서비스, 택시 등이 포함됩니다.**
2. 귀하께서는 유형별 자동차 속성 수준을 잘 확인하시고,
   ① 제시된 **4**가지 가상의 대안에 선호하는 순서대로 순위를 매겨 주시고,
   ② 비선택을 포함한 5가지 대안 중 가장 선호하는 것에 ○표 해주십시오.
3. 각 유형에 제시된 6개의 속성 이외의 **다른 모든 서비스 속성은 서로 동일**한 것으로 가정하고 응답해주십시오.

▣ 이용가능 운송수단 서비스 선호도 질문 1

속성		유형 A	유형 B	유형 C	유형 D	비선택
1. 주행거리당 소요 비용		800원/km	200원/km	200원/km	200원/km	
2. 연료종류		LPG	휘발유	경유	전기	
3. 차종		중형차·대형차	경차·소형차·준중형차	SUV·RV	SUV·RV	
4. 대기 및 이동 시간		5분	5분	5분	10분	
5. 운전자 및 차량 유형	**5-1. 운전자**	일반 운전자가	본인이 직접	일반 운전자가	일반 운전자가	
	5-2. 차량	비영업용차를 운전	영업용 차를 운전	비영업용 차를 운전	영업용 차를 운전	
순위 응답란(4순위 필수)						
선호 유형 응답란[SA]						

▣ 이용가능 운송수단 서비스 선호도 질문 2

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 주행거리당 소요 비용	200원/km	200원/km	500원/km	200원/km	
2. 연료종류	휘발유	LPG	전기	하이브리드	
3. 차종	중형차·대형차	경차·소형차·준중형차	경차·소형차·준중형차	경차·소형차·준중형차	

4. 대기 및 이동 시간		10분	10분	5분	10분	
5. 운전자 및 차량 유형	5-1. 운전자	본인이 직접	전문기사가	본인이 직접	일반 운전자가	
	5-2. 차량	영업용 차를 운전	영업용 차를 운전	영업용 차를 운전	비영업용 차를 운전	
순위 응답란(4순위 필수)						
선호 유형 응답란[SA]						

▣ 이용가능 운송수단 서비스 선호도 질문 3

속성		유형 A	유형 B	유형 C	유형 D	비선택
1. 주행거리당 소요 비용		200원/km	200원/km	200원/km	500원/km	
2. 연료종류		경유	LPG	휘발유	하이브리드	
3. 차종		경차·소형차·준중형차	중형차·대형차	SUV·RV	SUV·RV	
4. 대기 및 이동 시간		5분	5분	15분	5분	
5. 운전자 및 차량 유형	5-1. 운전자	일반 운전자가	전문기사가	일반 운전자가	전문기사가	
	5-2. 차량	영업용 차를 운전	영업용 차를 운전	비영업용 차를 운전	영업용 차를 운전	
순위 응답란(4순위 필수)						
선호 유형 응답란[SA[						

▣ 이용가능 운송수단 서비스 선호도 질문 4

속성		유형 A	유형 B	유형 C	유형 D	비선택
1. 주행거리당 소요 비용		200원/km	200원/km	500원/km	500원/km	
2. 연료종류		하이브리드	전기	휘발유	LPG	
3. 차종		경차·소형차·준중형차	중형차·대형차	중형차·대형차	경차·소형차·준중형차	
4. 대기 및 이동 시간		15분	15분	10분	10분	
5. 운전자 및 차량 유형	5-1. 운전자	본인이 직접	전문기사가	일반 운전자가	일반 운전자가	
	5-2. 차량	영업용 차를 운전	영업용 차를 운전	영업용 차를 운전	비영업용 차를 운전	
순위 응답란(4순위 필수)						
선호 유형 응답란[SA]						

▣ 이용가능 운송수단 서비스 선호도 질문 5

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 주행거리당 소요 비용	200원/km	800원/km	200원/km	200원/km	
2. 연료종류	LPG	휘발유	경유	경유	

3. 차종	경차·소형차·준중형차	경차·소형차·준중형차	중형차·대형차	경차·소형차·준중형차	
4. 대기 및 이동 시간	15분	5분	20분	20분	
5. 운전자 및 차량 유형 – 5-1. 운전자	일반 운전자가	일반 운전자가	본인이 직접	전문기사가	
5. 운전자 및 차량 유형 – 5-2. 차량	영업용 차를 운전	영업용 차를 운전	영업용 차를 운전	영업용 차를 운전	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

▣ 이용가능 운송수단 서비스 선호도 질문 6

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 주행거리당 소요 비용	200원/km	200원/km	800원/km	500원/km	
2. 연료종류	휘발유	LPG	전기	경유	
3. 차종	경차·소형차·준중형차	SUV·RV	경차·소형차·준중형차	중형차·대형차	
4. 대기 및 이동 시간	20분	20분	20분	15분	
5. 운전자 및 차량 유형 – 5-1. 운전자	일반 운전자가	일반 운전자가	일반 운전자가	일반 운전자가	
5. 운전자 및 차량 유형 – 5-2. 차량	비영업용 차를 운전	영업용 차를 운전	비영업용 차를 운전	비영업용 차를 운전	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

▣ 이용가능 운송수단 서비스 선호도 질문 7

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 주행거리당 소요 비용	800원/km	800원/km	500원/km	500원/km	
2. 연료종류	경유	경유	LPG	휘발유	
3. 차종	경차·소형차·준중형차	SUV·RV	SUV·RV	경차·소형차·준중형차	
4. 대기 및 이동 시간	10분	10분	20분	20분	
5. 운전자 및 차량 유형 – 5-1. 운전자	전문기사가	본인이 직접	본인이 직접	전문기사가	
5. 운전자 및 차량 유형 – 5-2. 차량	영업용 차를 운전	영업용 차를 운전	영업용 차를 운전	영업용 차를 운전	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

▣ 이용가능 운송수단 서비스 선호도 질문 8

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 주행거리당 소요 비용	800원/km	800원/km	500원/km	800원/km	
2. 연료종류	LPG	휘발유	경유	하이브리드	

3. 차종		경차·소형차·준중형차	SUV·RV	경차·소형차·준중형차	중형차·대형차	
4. 대기 및 이동 시간		15분	15분	15분	20분	
5. 운전자 및 차량 유형	5-1. 운전자	본인이 직접	전문기사가	일반 운전자가	일반 운전자가	
	5-2. 차량	영업용 차를 운전	영업용 차를 운전	영업용 차를 운전	영업용 차를 운전	
순위 응답란(4순위 필수)						
선호 유형 응답란[SA]						

문1-1. **(전체 응답자)** 귀하께서는 앞서 제시된 운송수단 서비스 중에서 **비선택**을 제외하고 **실제로 무엇을 이용**하시겠습니까? **운송수단 서비스 선호도 질문1~6의 ①선호 순위 응답란에서 1위를 선택한 운송수단 서비스 중 하나에 응답**해 주십시오. **(단답)**

1. 질문 1의 선호 1위 운송수단 서비스
2. 질문 2의 선호 1위 운송수단 서비스
3. 질문 3의 선호 1위 운송수단 서비스
4. 질문 1의 선호 1위 운송수단 서비스
5. 질문 2의 선호 1위 운송수단 서비스
6. 질문 3의 선호 1위 운송수단 서비스

문1-2. **(문1-1.에서 응답한 운송수단 서비스에 대해) 귀하께서는 위에 선택하신 운송수단 서비스를 실제로 이용하실 의향**이 있으십니까?

1. 예 (있다) ➔ 문1-3으로 이동
2. 아니오(없다) —▶ **문2.로 이동하십시오.**

문1-3. **(문1-2.에서 응답한 운송수단 서비스에 대해) 선택하신 운송수단 서비스를 언제 이용**하시겠습니까?
※ 제시된 운송서비스가 현재부터 제공된다고 가정하고 응답해 주시기 바랍니다.

1. 1개월 안에 사용
2. 1~3개월 안에 사용
3. 3~6개월 안에 사용
4. 6~9개월 안에 사용
5. 9~12개월 안에 사용
6. 12~15개월 안에 사용
7. 15~18개월 안에 사용
8. 18~24개월 안에 사용
9. 24개월 이후 사용

문2. 다음은 귀하께서 위의 선호도 질문에 응답하면서 비교한 각 속성별 고려 수준에 대한 질문입니다.
위 질문에 응답하면서 귀하께서 각 속성에 대해 고려한 수준을 5점 척도로 응답해 주시기 바랍니다.

	전혀 고려하지 않았다	고려하지 않은 편이다	보통이다	대체로 고려했다	매우 고려했다
1. 주행거리당 소요 비용	1	2	3	4	5
2. 연료종류	1	2	3	4	5
3. 차종	1	2	3	4	5
4. 대기 및 이동시간	1	2	3	4	5
5. 운전자	1	2	3	4	5
6. 차량 종류	1	2	3	4	5

E. 자동차 구매행동 조사

[문1-1~문1-5 / 문1-6~문1-9 한페이지씩 두페이지로 제시]

다음 질문들은 귀하(댁)의 자동차 구해 행동과 구매시 고려 사항에 대해 묻는 질문입니다.

문1. 귀하께서 오늘 차를 구매한다고 가정할 경우 귀하께서 고려하실 차량의 속성범위를 식별해 주시기 바랍니다.
※ 모든 질문은 복수 선택이 가능합니다.

1-1. 자동차 **구매가격(초기 구매 비용)**을 선택해 주세요.
① 1,000만원 이상~2,000만원 미만　② 2,000만원 이상~3,000만원 미만
③ 3,000만원 이상~4,000만원 미만　④ 4,000만원 이상~5,000만원 미만
⑤ 5,000만원 이상~6,000만원 미만　⑥ 6,000만원 이상~7,000만원 미만
⑦ 7,000만원 이상~8,000만원 미만　⑧ 8,000만원 이상

1-2. 자동차의 **유종**을 선택해 주세요.
① 가솔린, 휘발유　② 경유(디젤)　③ 하이브리드
④ LPG　⑤ 전기자동차(EV)　⑥ 수소연료

1-3. 자동차의 **차종 및 크기**를 선택해 주세요.
① 경차(예: 모닝, 스파크)　② 준중형(예: K3, 아반떼, SM3)
③ 중형차(예: 소나타, K5, SM6)　④ 대형(예: 제네시스 G90, K9, 에쿠스)
⑤ SUV(예: 산타페, 투싼, QM6)　⑥ RV(예: 카니발, 스타렉스)

1-4. **이동거리당 이용비용**을 선택해 주세요.
※ 2018년 연평균 가솔린평균가 1581원, 디젤평균가 1392원/L, LPG가 875원/L, 전기가 313원/kWh
① 50원/km 이하　(가: 32km/L 이상, 디: 28km/L 이상, 전: 6.3 km/kWh 이상, LPG: 17.5km/L 이상)
② 50~100원/km 이하　(가: 16~32km/L, 디: 14~28km/L, 전: 3.1~6.3 km/kWh, LPG: 8.8~17.5km/L)
③ 100~150원/km 이하　(가: 11~16km/L, 디: 9~14km/L, 전: 2.1~3.1 km/kWh, LPG: 5.8~8.8km/L)
④ 150~200원/km 이하　(가: 8~11km/L, 디: 7~9km/L, 전: 1.6~2.1 km/KWh, LPG: 4.4~5.8km/L)
⑤ 200~250원/km 이하　(가: 6~8km/L, 디: 6~7km/L, 전: 1.3~1.6 km/kWh, LPG: 3.5~4.4km/L)
⑥ 250원/km 이상　(가: 6km/L 미만, 디: 6km/L 미만, 전: 1.3km/kWh 미만, LPG: 3.5km/L 미만)

1-5. 자동차의 **배기량**을 선택해 주세요.
※ 현대 모닝, 기아 스파크와 같은 경차는 1000cc 이하, 현대 아반떼와 같은 준중형은 1500cc
현대 소나타 기아 K5과 같은 중형은 2000~2500cc, 그랜져 및 제네시스는 2500~3000cc 수준
① 1000cc 이하　② 1,000cc 이상~2,000cc 미만
③ 2,000cc 이상~3,000cc 미만　④ 3,000cc 이상~4,000cc 미만　⑤ 4,000cc 이상

1-6. **자동차 유지비**(구매가격을 제외하고 연간 지출되는 비용)를 선택해 주세요.
※ 보험금 및 주유비를 제외하고 주차비, 세금, 점검비 등 자동차를 보유하고 유지함으로써 발생하는 부수적인 비용
(예시) 경차 및 친환경 자동차의 경우 다양한 과세지원이 가능하며, 공공주차장 이용시 일반 승용차에 비해 50% 저렴한 비용이 부과됨
(예시) 국내 자동차의 경우 외제차 대비 점검 비용이 저렴함: 현대 제네시스 엔진오일 교체비용 7~8만원 수준 BMW 5시리즈 엔진오일 교체비용 20만원 수준

① 월 10만원 수준 ② 월 20만원 수준
③ 월 30만원 수준 ④ 월 40만원 수준
⑤ 월 50만원 수준 ⑥ 월 60만원 수준

1-7. 자동차의 **온실가스 배출량 수준** (차량의 km당 CO2 배출량)
① 에너지 소비 효율 1등급 (휘발유: CO2 배출량 100g/km 이하, 경유: CO2 배출량 120g/km 이하)
② 에너지 소비 효율 2등급 (1등급 대비 배출량 10% 증가)
③ 에너지 소비 효율 3등급 (1등급 대비 배출량 30% 증가)
④ 에너지 소비 효율 4등급 (1등급 대비 배출량 70% 증가)
⑤ 에너지 소비 효율 5등급 (1등급 대비 배출량 100% 이상 증가)

1-8 **충전소 접근성**을 선택해 주세요.
※ 소비자가 위차한 곳에서 가장 가까운 주유/충전소까지의 평균 거리를 의미하며 2019년 기준 국내 휘발유/경유를 판매하는 일반 주유소까지의 평균거리는 약 2km임
① 2km
② 4km
③ 6km
④ 8km
⑤ 10km

1-9 자동차 **브랜드**를 선택해 주세요.
① 국산차 (**현대자동차, 기아자동차, 르노삼성, 쌍용 등)**
② 수입차 (**벤츠, BMW, 아우디, 렉서스, 토요타, 혼다, 지프, 랜드로바, 크라이슬러 등)**

F. 자동차 유형별 선호도

다음은 자동차의 여러 속성과 속성별 수준에 대한 설명입니다.
본 조사에서는 현재 시장에서 본격적으로 판매되고 있는 차량과 그렇지 않은 대체에너지 자동차를 포함하여, 신차구입에 대한 선호도를 알아보고 있습니다. 다음 제시한 속성 설명을 숙지하시고 응답해 주시기 바랍니다.

▣ 자동차 속성 및 수준 설명문

속성		속성 설명 및 수준
1. 연료종류	설명	수송 모드의 연료종류는 휘발유, 경유, LPG, 전기로 구분됨
	수준 (5개)	① 휘발유 ② 경유 ③ LPG ④ 전기 ⑤ 하이브리드
2. 차종	설명	수송 모드의 차종은 경차·소형차, 준중형차·중형차, 대형차, SUV·RV로 구분됨
	수준 (3개)	①경차·소형차·준중형차 ②중형차·대형차 ③SUV·RV
3. 연료비용 (원/km)	설명	연료비용은 1Km 주행 시 드는 비용(원)임.
	수준 (4개)	①250원/km (가솔린 및 디젤: 6km/L 미만, 전기: 1.3km/kWh 미만, LPG: 3.5km/L 미만) ② 200원/km (가솔린 및 디젤: 8km/L, 전기: 1.6 km/kWh, LPG: 4.4km/L) ②150원/km (가솔린 및 디젤: 11km/L, 전기: 2.1 km/KWh, LPG: 5.8km/L) ④100원/km (가솔린 및 디젤: 16km/L, 전기: 3.1 km/kWh, LPG: 8.8km/L)
4. 차량가격 (만원)	설명	보험이나 세금 등을 제외한, 차량 구입비용
	수준 (6개)	① 1,500만원 ②3,000만원 ③4,500만원 ④6,000만원 ⑤ 7,500만원 ⑥ 9,000만원
5. 제조사 브랜드	설명	자동차 제조사의 국적 국내: 현대자동차, 기아자동차, 르노삼성 국외: 벤츠, BMW, 아우디, 렉서스, 토요타, 혼다, 지프, 랜드로바, 크라이슬러 등
	수준 (2개)	① 국내 제조사 ② 외국 제조사

[문항당 설명 고정 제시]

1. 다음부터 설명 드린 5개의 속성을 조합하여 구성한 귀하께서 구매할 수 있는 신규 자동차에 **대한 선호를 묻는 질문입니다.**
   ☞ 제시한 가상의 대안은 **신규 차량의**로서 **휘발유, 경유, LPG, 전기, 수소 자동차 등 현재 판매되는 자동차가 포함됩니다.**
2. 귀하께서는 유형별 자동차 속성 수준을 잘 확인하시고,
   ① 제시된 **4**가지 가상의 대안에 선호하는 순서대로 순위를 매겨 주시고,
   ② 비선택을 포함한 5가지 대안 중 가장 선호하는 것에 ○표 해주십시오.
3. 각 유형에 제시된 5개의 속성 이외의 **다른 모든 서비스 속성은 서로 동일**한 것으로 가정하고 응답해주십시오.

▣ 이용가능 운송수단 서비스 선호도 질문 1

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 연료종류	하이브리드	전기	경유	경유	
2. 차종	SUV·RV	중형차·대형차	경차·소형차·준중형차	SUV·RV	
3. 연료비용 (원/km)	150원/km (11km/L)	150원/km (2.1km/kWh)	100원/km (16km/L)	100원/km (16km/L)	
4. 차량가격 (만원)	1,500만원	1,500만원	1,500만원	3,000만원	
5. 제조사 브랜드	외제	외제	국산	국산	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

▣ 이용가능 운송수단 서비스 선호도 질문 2

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 연료종류	휘발유	하이브리드	휘발유	전기	
2. 차종	경차·소형차·준중형차	중형차·대형차	중형차·대형차	SUV·RV	
3. 연료비용 (원/km)	150원/km (11km/L)	100원/km (16km/L)	150원/km (11km/L)	200원/km (1.6km/kWh)	
4. 차량가격 (만원)	1,500만원	1,500만원	3,000만원	1,500만원	
5. 제조사 브랜드	외제	외제	국산	국산	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

▣ 이용가능 운송수단 서비스 선호도 질문 3

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 연료종류	휘발유	휘발유	경유	경유	
2. 차종	경차·소형차·준중	경차·소형차·준중	SUV·RV	중형차·대형차	

	형차	형차			
3. 연료비용 (원/km)	250원/km (6km/L)	200원/km (8km/L)	250원/km (6km/L)	150원/km (11km/L)	
4. 차량가격 (만원)	1,500만원	1,500만원	1,500만원	3,000만원	
5. 제조사 브랜드	국산	외제	국산	외제	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

▣ 이용가능 운송수단 서비스 선호도 질문 4

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 연료종류	LPG	휘발유	하이브리드	전기	
2. 차종	경차·소형차·준중형차	SUV·RV	경차·소형차·준중형차	경차·소형차·준중형차	
3. 연료비용 (원/km)	150원/km (5.8km/L)	250원/km (6km/L)	200원/km (8km/L)	250원/km (1.3km/kWh)	
4. 차량가격 (만원)	1,500만원	1,500만원	3,000만원	3,000만원	
5. 제조사 브랜드	국산	국산	국산	국산	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

▣ 이용가능 운송수단 서비스 선호도 질문 5

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 연료종류	경유	경유	LPG	휘발유	
2. 차종	중형차·대형차	경차·소형차·준중형차	중형차·대형차	경차·소형차·준중형차	
3. 연료비용 (원/km)	250원/km (6km/L)	200원/km (8km/L)	200원/km (4.4km/L)	200원/km (8km/L)	
4. 차량가격 (만원)	1,500만원	1,500만원	1,500만원	3,000만원	
5. 제조사 브랜드	외제	국산	국산	외제	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

▣ 이용가능 운송수단 서비스 선호도 질문 6

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 연료종류	LPG	경유	경유	전기	
2. 차종	SUV·RV	SUV·RV	경차·소형차·준중형차	경차·소형차·준중형차	

3. 연료비용 (원/km)	250원/km (3.5km/L)	200원/km (8km/L)	150원/km (11km/L)	100원/km (3.1km/kWh)	
4. 차량가격 (만원)	3,000만원	4,500만워	4,500만원	4,500만워	
5. 제조사 브랜드	외제	외제	국산	외제	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

▣ 이용가능 운송수단 서비스 선호도 질문 7

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 연료종류	휘발유	LPG	하이브리드	휘발유	
2. 차종	SUV·RV	경차·소형차·준중형차	중형차·대형차	경차·소형차·준중형차	
3. 연료비용 (원/km)	200원/km (8km/L)	150원/km (5.8km/L)	250원/km (6km/L)	250원/km (6km/L)	
4. 차량가격 (만원)	4,500만워	4,500만워	4,500만워	6,000만원	
5. 제조사 브랜드	외제	국산	국산	외제	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

▣ 이용가능 운송수단 서비스 선호도 질문 8

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 연료종류	휘발유	휘발유	전기	경유	
2. 차종	중형차·대형차	SUV·RV	중형차·대형차	중형차·대형차	
3. 연료비용 (원/km)	250원/km (6km/L)	100원/km (16km/L)	150원/km (2.1km/kWh)	250원/km (6km/L)	
4. 차량가격 (만원)	4,500만워	6,000만원	6,000만원	7,500만원	
5. 제조사 브랜드	국산	국산	국산	외제	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

▣ 이용가능 운송수단 서비스 선호도 질문 9

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 연료종류	하이브리드	경유	하이브리드	휘발유	
2. 차종	SUV·RV	경차·소형차·준중형차	경차·소형차·준중형차	경차·소형차·준중형차	
3. 연료비용 (원/km)	150원/km (11km/L)	200원/km (8km/L)	200원/km (8km/L)	150원/km (11km/L)	

4. 차량가격 (만원)	6,000만원	6,000만원	7,500만원	7,500만원	
5. 제조사 브랜드	외제	외제	국산	외제	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

▣ 이용가능 운송수단 서비스 선호도 질문 10

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 연료종류	경유	LPG	경유	휘발유	
2. 차종	경차·소형차·준중형차	중형차·대형차	중형차·대형차	SUV·RV	
3. 연료비용 (원/km)	250원/km (6km/L)	200원/km (4.4km/L)	200원/km (8km/L)	150원/km (11km/L)	
4. 차량가격 (만원)	6,000만원	6,000만원	9,000만원	9,000만원	
5. 제조사 브랜드	국산	국산	외제	국산	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

▣ 이용가능 운송수단 서비스 선호도 질문 11

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 연료종류	경유	전기	경유	휘발유	
2. 차종	경차·소형차·준중형차	SUV·RV	SUV·RV	중형차·대형차	
3. 연료비용 (원/km)	150원/km (11km/L)	200원/km (1.6km/kWh)	150원/km (11km/L)	200원/km (8km/L)	
4. 차량가격 (만원)	7,500만원	7,500만원	9,000만원	9,000만원	
5. 제조사 브랜드	국산	국산	국산	국산	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

▣ 이용가능 운송수단 서비스 선호도 질문 12

속성	유형 A	유형 B	유형 C	유형 D	비선택
1. 연료종류	LPG	하이브리드	휘발유	전기	
2. 차종	SUV·RV	경차·소형차·준중형차	중형차·대형차	경차·소형차·준중형차	
3. 연료비용 (원/km)	250원/km (3.5km/L)	250원/km (6km/L)	100원/km (16km/L)	250원/km (1.3km/kWh)	
4. 차량가격 (만원)	7,500만원	9,000만원	7,500만원	9,000만원	

5. 제조사 브랜드	외제	국산	국산	외제	
순위 응답란(4순위 필수)					
선호 유형 응답란[SA]					

G. 수송모드 관련 인식 조사

문1. **(전체 응답자)** 평소 귀하께서 교통수단을 사용할 때, **느껴지는 감정이나 생각을 아래 제시한 6개 항목별로 응답**해 주십시오.

	전혀 아니다	아닌 편이다	보통이다	그런 편이다	매우 그렇다
운전은 즐거운 것이다	1	2	3	4	5
다른 사람이 운전하는 것 보다 **내가 직접 운전하는 것을 선호**한다	1	2	3	4	5
내 차는 **나를 나타내는 상징**과 같다	1	2	3	4	5
차는 **운전자가 어떤 사람인지를 대변**한다	1	2	3	4	5
모르는 사람들과 운송수단을 **함께 사용하는 것은 불편**하다	1	2	3	4	5
개인적 운송수단(승용차, 택시, 렌트카)을 선호한다	1	2	3	4	5
직접 주차하는 것을 선호한다	1	2	3	4	5
주차 자리를 **찾는 것이 불편하다**	1	2	3	4	5
편안하게 이동하는 것을 선호한다	1	2	3	4	5
이동시 **안전하다는** 느낌을 받는 것을 선호한다	1	2	3	4	5

문2. **(전체 응답자)** 새로운 자동차를 구매할 경우, **아래 9개의 자동차 속성을 얼마나 중요하게 고려**할지 응답해 주십시오.

	전혀 중요하지 않다	중요하지 않은 편이다	보통이다	중요한 편이다	매우 중요하다
주행 시 발생하는 **소음**	1	2	3	4	5
주유 및 충전 비용 (연료비용)	1	2	3	4	5
연료 충전에 소요되는 **시간**	1	2	3	4	5
주유 또는 충전을 위해 대기하는 **시간**	1	2	3	4	5
트렁크 여유 **공간**	1	2	3	4	5
부품 결함으로 인한 **자동차 고장 확률**	1	2	3	4	5
최대 주행 가능 거리	1	2	3	4	5
자율주행시스템	1	2	3	4	5
오염물질 (CO2, NOx, SOx 미세먼지 등) 배출량	1	2	3	4	5
	1	2	3	4	5

문3. **(전체 응답자) 차세대 자동차(전기차, 수소차) 확산 정책(4가지 항목), 위험성 및 환경문제 인식(6가지 항목)**에 대해 귀하의 생각과 가장 가까운 번호에 표시해 주십시오.

◉ 차세대 자동차 확산 정책	전혀 아니다	아닌 편이다	보통이다	그런 편이다	매우 그렇다
우리나라의 친환경차 보급 정책은 홍보가 잘 되고 있다	1	2	3	4	5
우리나라의 친환경차 보급 정책이 원활하게 이루어지고 있다	1	2	3	4	5
친환경차의 확산은 온실가스 및 미세먼지 배출 감축에 기여 한다	1	2	3	4	5
환경개선을 위해 친환경차 확산은 반드시 필요하다	1	2	3	4	5

◉ 차세대 자동차 위험성 및 환경문제 인식	전혀 아니다	아닌 편이다	보통이다	그런 편이다	매우 그렇다
차세대자동차를 운전하는 것은 **잠재적인 위험이 따를 것**이다	1	2	3	4	5
전반적으로 **차세대자동차의 안정성이 내연기관차보다 낮**다	1	2	3	4	5
차세대자동차를 운전하는 것은 **불확실성이 따를 것**이다	1	2	3	4	5
차세대자동차는 홍보와 달리 **성능이 기대 이하일 것**이다	1	2	3	4	5
소비자는 **구매하는 제품의 환경적 영향에 관심**을 가져야 한다	1	2	3	4	5
소비자는 **환경을 오염시키는 제품에 더 높은 가격을 지불**해야 한다	1	2	3	4	5

F. 자료 분류용 질문

마지막으로 응답자 분류를 위한 질문입니다.

문1. 다음 질문은 **정부의 정책수행 프로세스**와 관련한 귀하께서 가지고 있는 인식을 묻는 질문입니다.

◉ 정책 수행 전반	전혀 그렇지 않다	그렇지 않은 편이다	보통이다	그런 편이다	매우 그렇다
우리나라의 정책 입안 과정은 투명하게 진행되고 있다	1	2	3	4	5
우리나라의 정책 입안 과정은 효율적으로 진행되고 있다	1	2	3	4	5
나는 우리나라의 정책 설계 및 입안 현황에 대해 긍정적으로 평가한다	1	2	3	4	5
나는 우리나라 정부 정책이 사회 현안을 해결하는 데 효과	1	2	3	4	5

적이라고 생각한다					
나는 우리나라 정부를 신뢰한다	1	2	3	4	5

◉ **정책프로세스 참여 관련 인식**	전혀 그렇지 않다	그렇지 않은 편이다	보통이다	그런 편이다	매우 그렇다
나는 정부의 정책 입안 과정에 적극적으로 참여하고 싶다····	1	2	3	4	5
우리나라의 정책 입안 과정에 국민이 참여할 수 있는 체계가 효과적으로 구축되어 있다	1	2	3	4	5
누구나 쉽게 정책 입안 과정에 참여할 수 있다	1	2	3	4	5

문2. 귀하의 **직업**은 무엇입니까?
1. 자영업 (종업원 9명이하 소규모업소 주인/가족종사자)
2. 판매/서비스직 (상점점원, 세일즈맨 등)
3. 기능/숙련공 (운전사, 선반/목공, 숙련공 등)
4. 일반작업직 (토목 현장작업/청소/수위/육체노동 등)
5. 사무/기술직 (일반회사 사무직/기술직, 교사 등)
6. 경영/관리직 (5급 이상 공무원/기업체 부장 이상 등)
7. 전문/자유직 (대학교수/의사/변호사/예술가/종교가 등)
8. 전업주부
9. 학생
10. 무직
11. 기타(구체적으로 응답해 주십시오 : _______________)

문3. 귀하의 **최종학력**은 어떻게 됩니까? 전문대 또는 대학교 재학 중인 경우, 1. 중/고등학교 졸업에 해당합니다.
1. 중/고등학교 졸업　　2. 전문대 졸업
3. 대학교 졸업　　4. 대학원 졸업

문4. 그럼, 귀하께서 살고 계신 **주택 유형**은 다음 중 어디에 해당됩니까?
1. 단독주택
2. 다가구 주택(연립주택/다세대주택)
3. 공동주택(오피스텔, 5개동 미만의 아파트/주상복합)
4. 아파트 단지(5개동 이상의 아파트/주상복합)
5. 기타( )

문5. 그럼, 귀하께서 살고 계신 **주택에 주차장**이 있습니까?
1. 예(있다)　　2. 아니오(없다)

문6. 그럼, 귀댁에서 가장 가까운 **대중교통(버스정류장 또는 지하철역 등)까지의 거리**는 걸어서 얼마나 걸립니까?
1. 5분 이내　　2. 10분 이내　　3. 15분 이내　　4. 20분 이내
5. 20분 이상

문7. 현재 귀 댁의 **월 평균 소득 수준**은 얼마나 됩니까? 세금은 제외한 **보너스, 이자수입 등 모든 수입**을 합해서 응답해 주십

시오. ___________만원 (RANGE: 50만원 이상)

Appendix: Transformer

The transformer neural network is one of the most recently developed form of ANN that is generally applied in dealing with natural language processing. Although similar to the RNN and LSTM in that it deals well with sequences of data, Transformers don’t require the sequences to be computed in order. Thus, unlike the RNN or LSTM, Transformer allow a parallelization of learning, or simultaneous training of the sequence data. (Polosukhin et al., 2017)

The main difference between the Transformer neural network (TNN), and traditional sequence ANN such as RNN and LSTM is that it is solely based on the attention mechanism. TNN propose that attention mechanism alone, is strong enough to achieve high performance compared to recurrent sequential processing. As mentioned earlier, both RNN and LSTM have inherent problem regarding long-term dependencies. Although LSTM is much more improved compared to general form of RNN, all models that rely on recurrent processing are prone to such threat. On the other hand, by using the attention mechanism, which lets the model to directly learn and draw the latent state of the data.

The transformer neural network is solely based on the attention mechanism. Similar to other neural networks, the Transformer network consist of stacked self-attention and point-wise, fully connected layers for both the encoder and decoder. The encoder maps the input sequence of symbol representations $x_1, ..., x_n$ to the sequence of continuous representations $z = z_1, ..., z_n$. Based on z , the decoder produces output sequence

$y_1, ..., y_n$ for each element at a time. Each step of the model is auto-regressive, meaning that it consumes the previously generated symbols as an additional input when generating the next symbol. The general construction of the Transformer is shown in Figure (38).

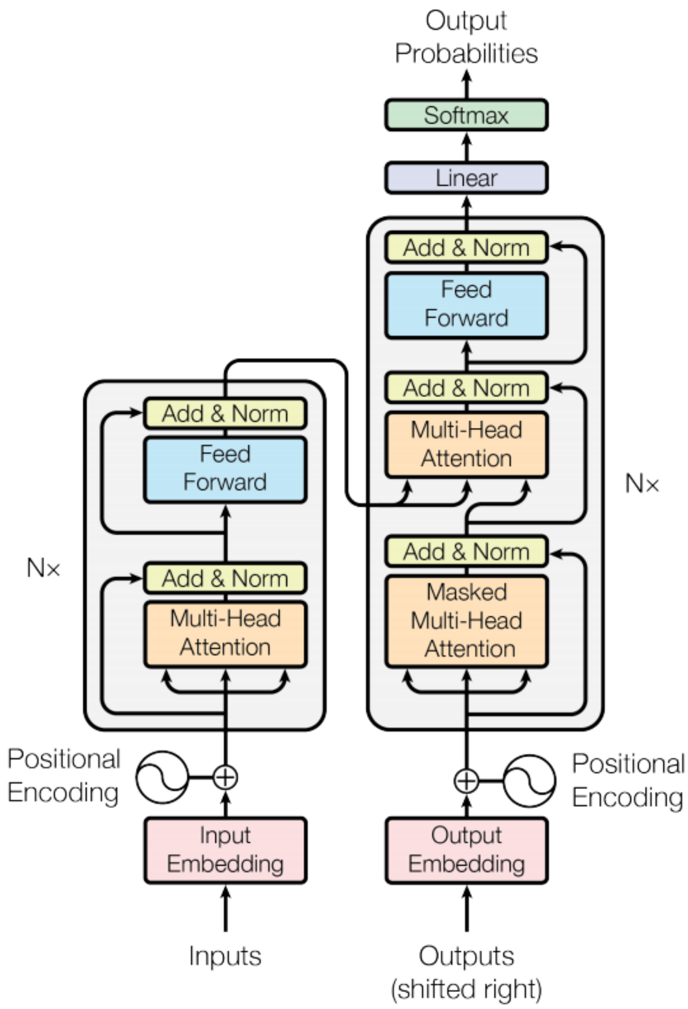


Figure 38. General structure of Transformer

The encoder is composed of N identical layers of stack, each with two sub-layers. The first sub-layer consist of multi-head self-attention mechanism and the second sub-layer consists of position-wise fully connected feed-forward network. The sub-layers are connected by residual connection that performs the layer normalization. In other words, the output of each sub-layer is defined by the equation below.

$$\mathrm{LayerNorm}\left(x + \text{Sub-layer}\left(x\right)\right)$$

Where $\text{Sub-layer}(x)$ is the function implemented through the sub-layer. The decoder is also composed of N identical layers of stack, each with three sub-layers. The third sub-layer performs the multi-head attention over the output of the encoder stack. Similar to the encoder component, the sub-layers in decoder is also connected by residual connections with layer normalization. The self-attention sub-layer in decoder is modified in to prevent positions from attending to subsequent positions. This structure ensure that the predictions for position i only depend on the outputs at positions less than i.

Attention is a function that pairs a vector of query, keys, values, and outputs. The output is computed as a weighted sum of the values, where the weight for each value is estimated through a compatibility function of the query with the corresponding key. The general Transformer proposed by Polosukhin et al. 2019, utilizes the 'Scaled Dot-production Attention', which consists of queries and keys of dimension d_k, and values of d_v. The model computed the dot products of the query with all keys, obtaining the weights on the values by dividing the values by $\sqrt{d_k}$ and applying the Softmax function. In particular, attention function is computed simultaneously by matrix Q, with K keys and V values. Each attention unit of the Transformer model learns three weight matrices: W_Q

is the query weight, W_K denote the key weight, and W_V denote the value weights. Each input x_i is multiplied with the matrices to produce the query vector $q_i = x_i W_Q$, a key vector $k_i = x_i W_K$ and the value vector $v_i = x_i W_V$. The attention weights are computed by combining the query and key vector, which as mentioned above, are divided by the square root of the dimension of the key vectors $\sqrt{d_k}$. The attention weights are used to stabilize the gradients during training, and passed through the Softmax function that normalizes the weights to sum to 1. The output of the matrix is computed according to the equation below.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

The multi-head attention consists of a set of matrices called an attention head, which consist of a set of matrices (W_Q, W_K, W_V). By combining multiple attentions, the model can learn the relevance relations that are transparent to humans. In other words, multiple attention heads have the potential to capture many levels and types of relevance relations, which are later passed into the feed-forward network layer. The formulation of the multi-head attention is expressed in the equation below

$$\begin{gathered}\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \ldots\ldots, \text{head}_h)W^O \\ \text{where head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)\end{gathered}$$

Transformer uses the multi-head attention in many different ways. The encoder-decoder attention layers receive the queries from the previous decoder layer, the memory keys and the values come from the output of the encoder, allowing every position in the decoder to attend over all positions in the input sequence. The encoder contains self-attention layers, which holds of all the keys, values, and queries from the previous layer. Thus, each position in the encoder can attend to all positions in the previous layer of the encoder. Likewise, self-attention layers in the decoder allow each position in the decoder to attend to all positions in the previous and the next decoder. The Transformer are trained based on semi-supervised learning involving unsupervised pre=training followed by supervised tuning.

Abstract (Korean)

본 연구는 기계학습 기반의 인공지능망과 기존의 통계적 마케팅 선택모형을 통합적으로 활용하여 제품 및 서비스 수용 이론으로 정의된 소비자들의 제품 수용 행위를 분석하였다. 기존의 제품 수용 이론들은 소비자들의 선택에 끼치는 영향을 단계별로 정의하였지만, 대부분의 이론은 제품 특성이 소비자 선택에 미치는 영향을 분석하기 보다는 소비자들의 의향, 제품의 대한 의견, 지각 수준과 소비자 선택의 관계 분석에 집중하였다. 따라서 본 연구는 소비자의 제품 수용 의향, 대안 평가 그리고 제품 및 사용량 선택을 포함하여 더욱 포괄적인 측면에서 소비자 제품 수용 행위를 분석하였다.

본 연구에서는 소비자의 제품 수용 관련 선택을 총 세 단계로 분류하였다. 첫 번째는 소비자의 제품 사용 의향을 결정하는 단계, 두 번째는 제품들의 대안을 평가하는 단계, 세 번째는 제품의 사용량을 선택하는 단계로, 각 단계를 분석하기 위해서 본 연구는 인공지능망과 통계적 마케팅 선택모형을 활용하였다. 인공지능망은 예측과 분류하는 작업에서 월등한 성능을 가진 모형으로 소비자들의 제품 수용 의향을 예측하고, 의향 선택에 영향을 주는 주요 변수들을 식별하는 데 활용되었다. 본 연구에서 제안한 주요 변수 식별을 위한 인공지능망은 기존의 변수 선택 기법 보다 모형 추정 적합도 측면에서 높은 성능을 보였다. 본 모형은 향후 빅데이터와 같이 많은 양의 소비자 관련 데이터를 처리하는데 활용될 가능성이 클 뿐만 아니라, 기존의

설문 설계 기법을 개선하는데 용이한 방법론으로 판단된다. 소비자 선호를 기반으로 한 대안 평가 및 사용량을 분석하기 위해서 통계적 선택 모형 중 계층적 베이지안 모형과 혼합 MDCEV 모형을 활용하였다. 계층적 베이지안 모형은개별적인 소비자 선호를 추정할 수 있는 장점이 있고, 혼합 MDCEV 모형의 경우 소비자들의 선호를 기반하여 선택된 대안들로 다양한 포트폴리오를 구성할 수 있고, 각 대안에 대한 사용량을 분석할 수 있다.

제안된 모형들의 실증 연구를 위해 차세대 자동차 수송 서비스에 대한 소비자들의 사용 의향, 서비스 대안에 대한 선호, 수송 서비스별 사용량을 분석하였다. 실증 연구에서는 차세대 자동차 수송 서비스를 수용하기까지 소비자들이 경험하는 단계별 선택 상황을 반영하였으며, 각 단계에서 도출된 결과를 통해 향후 차세대 자동차 수송 서비스의 성장 가능성과 소비자들의 이동 행위 변화에 대해 예측하였다. 본 연구를 통해 인공지능망이 소비자 관련 연구에서 유용하게 활용될 수 있음을 보였으며, 인공지능망과 통계적 마케팅 선택모형이 결합될 경우 소비자들의 제품 선택 행위뿐만 아니라, 제품 선택 의사결정 과정 전반에 걸쳐 소비자 선호를 포괄적으로 분석할 수 있음을 확인하였다.

주요어 : 인공지능망, 차세대 수송 서비스, 선택모형, 소비자 선택, 제품 수용 이론
학 번 : 2015-31038