

# Mode Choice Behavior of Various Airport User Groups for Ground Airport Access

Hyuk-Jae Roh\*

City of Regina, Old Fire Hall #1, 1640 11th Avenue, Regina, Saskatchewan, S4P 3C8, Canada

**Abstract:** In this research, we used a multinomial logit (MNL) discrete choice analysis technique to deepen the understanding of the mode choice behavior of various airport user groups categorized by trip purpose and trip distance for ground airport access. We used revealed preference (RP) data collected by an on-site-survey administrated by the Korea Transport Institute (KOTI) at the Kimpo International Airport passenger terminal in South Korea. Initially, four basic models were selected from a variety of model specifications, and these were analyzed to address general preferences in mode choice. The models were then evaluated in terms of the resulting estimation. The best-fitting model specification among four models was chosen for further study. Both trip distance models (standard-distance (SD) and long-distance (LD)) and trip purpose models (departing (D) and non-departing (ND)) were estimated. The results analyzed in this study encompass an unambiguous spectrum of mode choice behaviors associated with distinct airport user groups. The fundamental information, either revealed or reflected by modeling ground airport access for various airport user groups, could be essential not only to transportation planners -especially at the first phase of airport planning- but also to airport authorities faced with difficulties in managing ground transportation facilities to effectively serve airport users.

**Keywords:** Trip distance model, trip purpose model, airport user groups, discrete choice analysis, data input system.

## 1. INTRODUCTION

Most airport authorities and transportation planners who deal with airport access transportation systems would readily agree that there is a need for a deeper understanding of various airport user groups' preferences for ground access transportation modes. This interest might lead to efforts to improve the understanding of human behavior in mode choice situations, in order to spur more effective airport access transportation planning. However, until recently, a paucity of literature has addressed research on mode choice problems in airport ground access transportation [1, 2]. Furthermore, to the extent of the author's knowledge and experience, no prior research has dealt with mode choice behaviors for a variety of airport user grouped according to their trip distance and trip purpose.

In the present paper, we examine the effect of travel distance and purpose of a variety of airport user groups on their mode choice behaviors to determine the factors affecting their preferences by using a discrete choice analysis technique. It should be noted that in this research, "ground access" means a one-way trip headed to the airport and excludes any trips originating at the airport. Four basic models are presented that all have intuitive appeal in terms of parameter signs and relative magnitudes and that also show reasonable robustness in statistical terms. Subsequently, the best-fitted model specifications are prepared for the trip distance models, standard-distance (SD) and long-distance (LD), and the trip purpose models,

departing (D) and non-departing (ND). The implications of the estimation results are discussed from a behavioral point of view as well as from a statistical point of view under the circumstances of the airport ground access transportation system in South Korea. The following section presents a literature review mainly focused on examples of the application of discrete choice analysis to transportation studies as well as to airport-related studies. The next section briefly describes the data used for this research. An introduction to important elements used in operational discrete choice analysis follows, along with the procedure used for deriving a multinomial logit (MNL) model. The results obtained from the model estimations are presented next, followed by a conclusion. All steps in this research process are summarized in Fig. (1).

## 2. STATE-OF-THE-ART

### 2.1. Literature Review

The forecasting of the behavior of decision-makers in the transportation market heavily depends on econometric modeling methods such as discrete choice analysis. This technique has been widely applied in disciplines such as economics, engineering, psychology, and even management to analyze a given choice situation and to gain a deeper awareness of all factors affecting the choice situation in which the decision-makers are involved [3]. A large amount of literature exists in which this method is employed. For example, several studies have aimed to evaluate projects by calculating the monetary value of competing proposed projects in terms of the value of time (VOT) [4, 5]. Some of these studies focused on the description of each specific choice situation that the decision-maker faced [6-10]. Other studies applied the method to a policy-sensitive analysis in

\*Address correspondence to this author at the City of Regina, Old Fire Hall #1, 1640 11th Avenue, Regina, Saskatchewan, S4P 3C8, Canada;  
Tel: +1 (306) 751-4381; Fax: +1 (306) 777-6700; E-mail: [jroh@regina.ca](mailto:jroh@regina.ca)

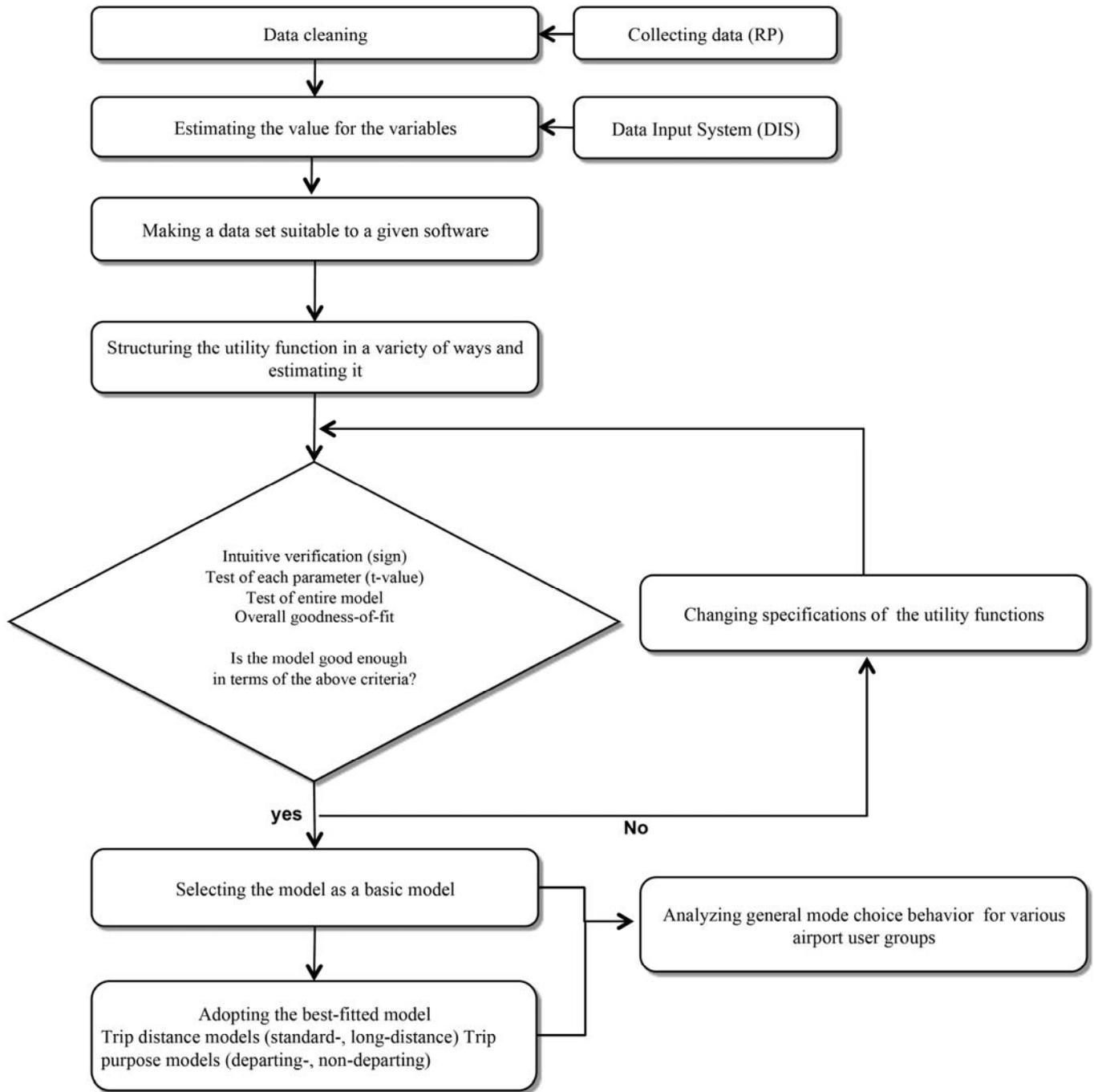


Fig. (1). Research flow.

order to predict mode share shifting by changing one of the policy variables included in the discrete choice model [4, 5].

The popularity of discrete choice analysis in a variety of research fields has been accelerated not only by its well-defined theoretical background, which involves the process of driving an operational discrete choice model function, but also by the ease with which its results can be interpreted and analyzed. However, in spite of its general application as a popular modeling tool in various fields, until recently, only a few studies have used discrete choice models in an airport-

related analysis. Moreover, almost all such research has involved either an analysis of airport choice for air cargo trans-shipment [6] or a study of airport choice for air transport passengers in multi-airport regions [11, 12]. Another important topic associated with airport-related research is a study on the ground access mode choice behavior of travelers arriving at the airport for a flight [1, 2, 13]. Another important topic is a study on the estimation of airport passengers' willingness to pay for access time savings in the framework of both probit discrete choice and regression methods [14].

## 2.2. Elements of Discrete Choice Analysis

One of the assumptions about decision-makers that are required in order to conduct the estimation of the operational discrete choice model is that they act as rational economic consumers in determining their choices. This is required in the framework of utility maximization concepts (or utility decision rules) that are used in discrete choice analysis [15, 16]. Rational decision-makers must follow two underpinning rules, “consistency” and “transitivity”, in their choice of behavior. More details on the required assumptions can be found in Train [15]. In this research, airport users are considered to satisfy all of the conditions mentioned above and, thus, are expected to behave accordingly.

The choice set from which decision-makers choose one alternative is another critical aspect to be considered in the process of discrete choice analysis. Theoretically, two choice sets can be conceived according to the decision-making context, the universal choice set and the feasible choice set [16]. For the sake of simplicity in modeling, it is assumed that the two choice sets are the same in this research, which means that all alternatives are accessible for all decision-makers in the spatial scope of this research. This is not an unreasonable assumption given the present knowledge and experience with transportation involving airport ground access in South Korea. In discrete choice modeling, three compulsory rules must be satisfied when setting the choice set of alternatives [15, 17]. First, the alternatives should be mutually exclusive, which means that the choice of one alternative does not permit the choice of another alternative. Second, the choice set should be exhaustive, which is accomplished by including all possible alternatives in a specific choice modeling activity. Finally, the number of alternatives should be finite, and the counting of alternatives should be complete. With these points in mind, all five alternatives (i.e., car, taxi, subway, bus, and limousine) considered in this research are classified based on the service characteristics that they provide in the airport access transportation market in Korea and are assumed to satisfy all of the conditions described above.

The mathematical form of the utility function consists of two parts, a systematic part and a random part. As will be addressed in the following section, the systematic part is typically built as a mathematical relationship between the surveyed variables. On the other hand, the random part assumes that the variables are each distributed in terms of a certain probability distribution. The mathematical relationship of the selected variables can be formulated as either “linear-in-parameters” or “non-linear-in-parameters”, depending heavily on the judgment of the researchers. However, in a case in which the researchers choose a non-linear-in-parameters specification, a certain problem may arise: a commercial software cannot handle this type of specification, and thus, the researcher might be forced to write a specific computer code dedicated only for that work. For this reason, a linear-in-parameters relationship is generally preferred [17]. In this research, the mathematical form of the representative utility is assumed to follow a linear-in-parameters relationship for all five alternatives.

The alternatives are weighted by their attribute values, thereby measuring their relative attractiveness, with the

attribute variables being observed by a researcher [16]. Additionally, a decision rule that involves a general random utility maximization concept has been continuously employed in making operational discrete choice models, due to its strong theoretical basis, extensive application history, and amenability in mathematical and statistical terms [17]. It is notable that the behavior of decision-makers can be represented or formulated in the decision rule, regardless of whether the decision-makers make a choice decision within a structured decision process or not [16]. Two type of data set representing attributes of alternatives have been known as revealed preference (RP) and stated preference (SP) data. RP data reflects situation where mode choice is really made in transportation market place. Conversely, SP data reflects situation where mode choice is hypothetically made to consider new alternatives that are not observed in RP data [18]. In this research, RP data were solely used to develop models.

## 3. METHODOLOGY

### 3.1. Multinomial Logit Model

The multinomial logit (MNL) model is the most widely used discrete choice model to date, due to its simple probabilistic choice function, the ease of estimation result interpretation, and the clarity in the algebraic manipulations related to the derivation of the final probabilistic choice function.

The probability that any alternative  $i$  in  $C_n$  is chosen by decision-maker  $n$  can be formulated as:

$$\begin{aligned} Prob_n(i) &= Pr(U_{ni} > U_{nj}, \forall j \in C_n, i \neq j) \\ &= Pr(S_{ni} + R_{ni} > S_{nj} + R_{nj}, \forall j \in C_n, i \neq j) \\ &= Pr(R_{nj} < R_{ni} + S_{ni} - S_{nj}, \forall j \in C_n, i \neq j). \end{aligned} \quad (1)$$

To accommodate unexplained variations of choice behavior in a reasonable way, as well as to account for a lack of information originating from a variety of sources (e.g., measurement errors, omission of alternative attributes, and omission of the characteristics of the decision-maker [16]), the total utility  $(U_{ni}, U_{nj})$  can be divided into the systematic  $(S_{ni}, S_{nj})$  and random  $(R_{ni}, R_{nj})$  parts. Different types of discrete choice models can be derived from Equation (1) by adopting a specific assumption regarding the distribution of the random parts,  $R_{ni}$  and  $R_{nj}$ . To obtain a multinomial logit probabilistic choice function,  $R_{ni}$  and  $R_{nj}$  are assumed to have a Weibull distribution. This distribution is also known as Gumbel or Type-I Extreme Value distribution. After some algebraic manipulations, the final form of the choice function for the MNL model is:

$$Prob_n(i) = \frac{\exp^{S_{ni}}}{\sum_j \exp^{S_{nj}}}. \quad (2)$$

For further details of the derivation procedures, the reader may refer to Train [15] and Ben Akiva and Lerman [17]. The final logit choice function is composed of two

parts, a denominator, which is the sum of the systematic utility of all alternatives in the choice set, and a numerator, which is the systematic utility of the alternative chosen by the decision-makers.

**3.2. Utility specification, Variable, and Software**

Based on preliminary knowledge of the discrete choice model described up to this point, the explanatory variables are combined in a variety of ways in structuring the utility function. The model parameters are then estimated using a discrete choice model provided in LIMDEP Ver. 7.0., Econometric Software, Inc. [19]. Four types of utility specifications are selected in the initial stage of research to investigate mode choice preferences when people access the airport by ground transportation. The variables and their mathematical relations in the utility functions are summarized in Tables 2 and 3, respectively. Included are two generic variables of cost and travel time ( $TC_{C,T,S,B,L}$  one way total cost, and  $TT_{C,T,S,B,L}$  one way total travel time), one alternative-specific socioeconomic variable ( $Accom$ ), two alternative-specific dummy variables  $Sexm$  (if male=1) and  $Age340$  (if age range of age 30-40 = 1) and four alternative-specific constants.  $Tasc$  represents Taxi alternative-specific constant ( $Tasc$ ). Likewise, other three acronyms (i.e.,  $Sasc, Basc, and Lasc$ ) are devised by employing the same naming system for Subway, Bus, and Limousine.

As presented in Table 3, the Car mode is designated to be the reference mode. Therefore, except the function for the car mode, other functions belonging to other modes have an alternative-specific constant. By incorporating alternative-specific constant, a modeler can capture relative utility of

each mode as compared to the Car mode, which is not explained by attributes in the model specification.

**4. DATA**

The data used for this research were collected in the year of 2003 during an on-site-survey managed by the Korea Transport Institute (KOTI) at the Kimpo International Airport passenger terminal in South Korea. This survey was carried out for 6 days and conducted using a two-part questionnaire. The first part of the form was designed to collect information regarding the socio-economic traits of airport users (sex, age, job, family size, etc.), and the second part was composed of questions designed to collect information about the characteristics of both the chosen travel modes and the travel itself (travel purpose, origin, travel time, travel cost, number of companions, etc.). It should be noted that all airport users answering this questionnaire arrived at the airport *via* ground access. In this research, for the purposes of either reducing the effort of preparing data sets suitable for a given software or maximizing the efficiency of data management and minimizing errors from data handling and formatting, a data input system (DIS) was developed using the Visual Basic Application (VBA) provided in MS EXCEL. The developed system is shown in Fig. (2). The data input system (DIS) proved to function appropriately in both ways, saving time and enhancing the accuracy of data entry. With the help of the data input system (DIS) and data cleaning processes, a final data set consisting of 540 observations (out of 800 questionnaires) was prepared for modelling purpose. The sample frequencies for the chosen mode from this data set are given in Table 1.

**Table 1. The Sample Frequencies of the Chosen Mode**

Alternative	Car	Taxi	Subway	Bus	Limousine
Choice (%)	247 (46)	70 (13)	112 (21)	73 (13)	38 (7)

**Table 2. Variables Included in the Four Basic Models**

Model	Variables Included in the Utility Functions
Model 1	Two generic variables, $TCost$ and $TTime$ , for car, taxi, subway, bus, and limousine
	Alternative-specific constants, $Tasc, Sasc, Basc, Lasc$ , for taxi, subway, bus, and limousine
Model 2	Mode-specific variable, $Cost$ , for car
	Mode-specific variable, $Fare$ , for taxi, subway, bus, and limousine
	Generic variable, $TTime$ , for car, taxi, subway, bus, and limousine
	Alternative-specific constants, $Tasc, Sasc, Basc, Lasc$ , for taxi, subway, bus, and limousine
Model 3	Two generic variables, $TCost$ and $TTime$ , for car, taxi, subway, bus, and limousine
	Mode-specific variable, $Accom$ , for car
	Alternative-specific constants, $Tasc, Sasc, Basc, Lasc$ , for taxi, subway, bus, and limousine
Model 4	Two generic variables, $TCost$ and $TTime$ , for car, taxi, subway, bus, and limousine
	Mode-specific variable, $Accom$ , for car
	Two socio-economic dummy variables, $Sexm$ and $Age340$ , for car
	Alternative-specific constants, $Tasc, Sasc, Basc, Lasc$ , for taxi, subway, bus, and limousine

**Table 3. Mathematical Relations of Variables in the Utility Function for Each of the Four Models**

<b>Model 1</b>	$S(car) = \beta_0 TCost + \beta_3 TTime$
	$S(taxi) = \beta_0 TCost + \beta_3 TTime + \beta_7 Tasc (= 1)$
	$S(subway) = \beta_0 TCost + \beta_3 TTime + \beta_8 Sasc (= 1)$
	$S(bus) = \beta_0 TCost + \beta_3 TTime + \beta_9 Basc (= 1)$
	$S(limousine) = \beta_0 TCost + \beta_3 TTime + \beta_{10} Lasc (= 1)$
<b>Model 2</b>	$S(car) = \beta_1 Cost + \beta_3 TTime$
	$S(taxi) = \beta_2 Fare + \beta_3 TTime + \beta_7 Tasc (= 1)$
	$S(subway) = \beta_2 Fare + \beta_3 TTime + \beta_8 Sasc (= 1)$
	$S(bus) = \beta_2 Fare + \beta_3 TTime + \beta_9 Basc (= 1)$
	$S(limousine) = \beta_2 Fare + \beta_3 TTime + \beta_{10} Lasc (= 1)$
<b>Model 3</b>	$S(car) = \beta_0 TCost + \beta_3 TTime + \beta_4 Accom$
	$S(taxi) = \beta_0 TCost + \beta_3 TTime + \beta_7 Tasc (= 1)$
	$S(subway) = \beta_0 TCost + \beta_3 TTime + \beta_8 Sasc (= 1)$
	$S(bus) = \beta_0 TCost + \beta_3 TTime + \beta_9 Basc (= 1)$
	$S(limousine) = \beta_0 TCost + \beta_3 TTime + \beta_{10} Lasc (= 1)$
<b>Model 4</b>	$S(car) = \beta_0 TCost + \beta_3 TTime + \beta_4 Accom + \beta_5 Sexm + \beta_6 Age340$
	$S(taxi) = \beta_0 TCost + \beta_3 TTime + \beta_7 Tasc (= 1)$
	$S(subway) = \beta_0 TCost + \beta_3 TTime + \beta_8 Sasc (= 1)$
	$S(bus) = \beta_0 TCost + \beta_3 TTime + \beta_9 Basc (= 1)$
	$S(limousine) = \beta_0 TCost + \beta_3 TTime + \beta_{10} Lasc (= 1)$

## 5. RESULTS

### 5.1. Four Basic Models

The results of estimation for the four models are summarized in Table 4. The table includes the values of parameter estimates for all representative utilities specified in the four models. In addition, the t-statistics used for the hypothesis test for each of the estimated parameters are also shown. The values of the log-likelihood for each of the four models are included to demonstrate the change in log-likelihood values due to different specifications and to test reasonableness of the model specification by calculating statistics such as  $-2[LL(0) - LL(\beta)]$ . Here,  $LL(0)$  is a log-likelihood value for the novice model that does not contain any explanatory variables in the utility function, and  $LL(\beta)$  is a log-likelihood value for the full model containing all explanatory variables considered in the model specifications, representing characteristics of each of the alternatives and demographics of the decision-makers.

Finally, researchers can obtain test values such as  $\rho_0^2$  and  $-\rho_0^{-2}$  to determine the overall goodness-of-fit. It is helpful to know that the value of time (VOT) can be used to approximate an answer to “How much of the amount of money are individuals willing to forfeit in order to obtain some benefits by choosing a specific mode for carrying out different purposes (i.e., actions or tasks)? This is a common objective in the use of discrete choice models. In simple linear models, VOT is calculated as the ration of two parameter estimates (i.e.,  $\beta_{TTime}, \beta_{TCost}$ ). For example, the estimated VOT (won/hr) of the first model among four basic models is calculated in Table 4 with following formula:

$$VOT = (\beta_{TTime} / \beta_{TCost}) \times 60 = (-0.020532 / -0.000067) \times 60 = 18,387 \text{ won / hr.}$$

As an intuitive verification, all parameters for all four models have signs and relative magnitudes corresponding to our prior knowledge. For example, the sign of the one-way total cost variable,  $TCost$  is negative (-), which means that this variable has a negative effect on mode choice. In the same line of thinking, the variables,  $Cost$  and  $Fare$ , included in Model 2 make sense in their signs. In other words, it is preferable for decision-makers to experience a decrease in their values for the chosen mode. As we can see by comparing the magnitude of two estimated parameters describing time and cost variables, travelers going to the airport are more sensitive to travel time than travel cost because their primary concern is to arrive at the airport without delay. A person who travels to the airport to board a plane has a tendency to choose the car mode in proportion to their number of companions. Males between the ages of 30 and 40 show a preference for using a car for travelling to the airport. Model 2 uses two variables,  $Cost$  (which is alternative-specific to the car mode) and  $Fare$  (which is alternative-specific to the other modes) instead of commonly using  $TCost$  as a generic variable for all five alternatives. This is done to reflect the fact that decision-makers might perceive monetary value differently according to the mode they choose, even though the money is paid for the same purpose. The results obtained from Models 1 to 3, if other conditions are equal, show that the car mode is the most attractive for people traveling to the airport, followed by taxi, bus, subway, and limousine in that order, but in the case of Model 4, the taxi mode is the most attractive. Considering the cumbersome nature of transfers between transit modes in the context of the Korean transit infrastructure, these results reflect a choice pattern that occurs in the real-world choice situation.

Statistical tests for each parameter of the model can be used to evaluate the formal hypothesis about each individual parameter [16]. The test statistic used for this evaluation takes the form of  $t_{statistics} = (\beta_k - \beta_k^*) / S_k$ , where  $\beta_k$  is the estimate for the  $K^{th}$  parameter,  $\beta_k^*$  is the hypothesized value for the  $K^{th}$  parameter, and  $S_k$  is the standard error of the  $K^{th}$  parameter. By comparing  $t_{statistics}$  calculated above with the critical absolute  $t$ -values calculated at a certain

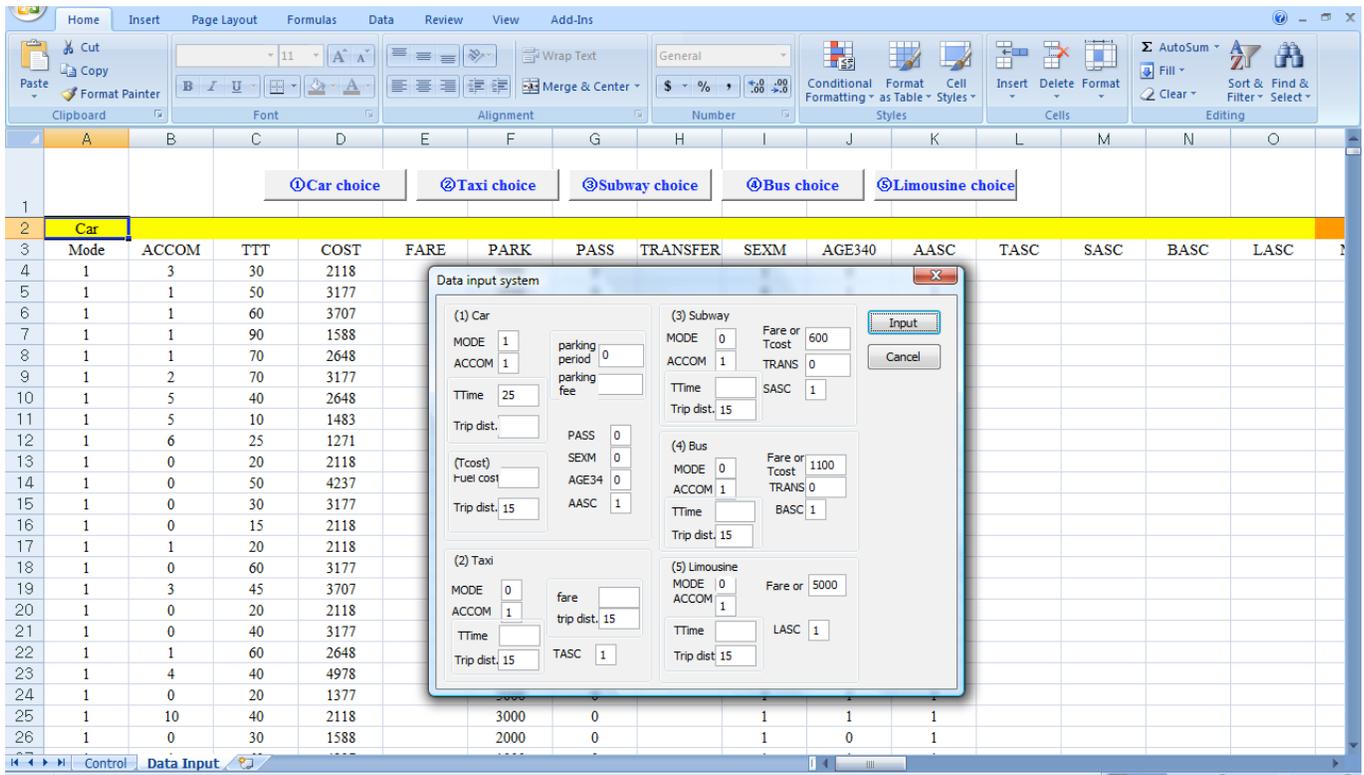


Fig. (2). Data input system developed using VBA.

level of confidence, a researcher can decide whether to reject the null hypothesis that a particular variable has no influence, in the context of this research, on mode choice, or equivalently, a researcher can determine whether the true parameter related to a particular variable is zero. Based on a 99% level of confidence, if the variable's absolute *t-values* is greater than 2.576, a modeler can reject the null hypothesis. At a 95% level of confidence, the corresponding *t-values* is 1.960. For example, as shown in Table 4, we can decide that the parameters ( $\beta_0, \beta_1, \beta_2, \beta_3$ ) significantly influence the utility function of each of the five modes at a 99% level of confidence, and thus, the variables should be retained in the utility function.

Statistical tests used for evaluating multiple hypotheses that certain parameters are zero or that pairs of parameters are equal to one another can be conducted by using  $-2 \times [LL(R) - LL(U)]$  [16], where  $LL(R)$  is the log-likelihood value in the restricted model, and  $LL(U)$  is the log-likelihood value in the unrestricted model. The model that exists before a restriction is imposed is considered as the unrestricted model. After imposing the restrictions of the multiple hypotheses mentioned above, the model becomes a restricted model. In the context of this research,  $LL(R)$  is  $LL(0)$  and  $LL(U)$  is  $LL(\beta)$ ; using the results from Table 4, the test statistics can be calculated to decide whether the pre-determined hypothesis should be rejected. The hypothesis for all four models is that all variables included in the model provide no additional information in explaining

the mode choice behavior of decision-makers. For example, in the case of Model 2, the restrictions given as a null hypothesis are  $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_7 = \beta_8 = \beta_9 = \beta_{10} = 0$ . The test statistics calculated are  $-2 \times [LL(0) - LL(\beta)] = 311.6804$ , and by referring to a value for specific conditions (i.e., the number of restrictions is 7, critical value is 24.32), a researcher can reject the null hypothesis at a confidence level higher than 99.9%. The result suggests that the true parameter value of these seven variables is not zero, and it makes sense to include all seven variables in the model specification. Other models can be interpreted in the same manner.

The overall goodness-of-fit of the discrete choice models can be described by  $\rho^2$ , known as the rho-squared value. As shown in Fig. (3) [16], by using the scalar relationship among the log-likelihood values, a researcher can use four different values for evaluating the model. In this research, as shown in Table 4, two different log-likelihood values are obtained. For a zero parameter model, we obtain  $LL(0) = -869.0965$ , and this value is the same for all four models, due to the fact it is determined by the same calculation, e.g.,  $\ln(1/5) \times 540$ , in all four models. For an estimated model, the log-likelihood shows variation in its value as given in Table 4. Two other log-likelihood values given in Fig. (3) are for a constant only model and a perfect prediction model, represented by  $LL(C)$  and  $LL(\beta_{true})$ , respectively. The numeric value for the perfect model should always be 0, due to the fact that  $\ln(1) \times 540 = 0$ .

Table 4. The Estimation Results for the Four Basic Models

Alternatives <sup>a</sup>	Variable	Model 1		Model 2		Model 3		Model 4	
		Parameter Estimates	t-Value	Parameter Estimates	t-Value	Parameter Estimates	t-Value	Parameter Estimates	t-Value
C, T, S, B, L	TCost ( $\beta_0$ )	-0.000067	-4.056**	-	-	-0.000059	-3.652**	-0.000054	-3.433**
C	Cost ( $\beta_1$ )	-	-	-0.000075	-2.634**	-	-	-	-
T, S, B, L	Fare ( $\beta_2$ )	-	-	-0.000064	-3.487**	-	-	-	-
C, T, S, B, L	TTime ( $\beta_3$ )	-0.020532	-6.410**	-0.020250	-6.167**	-0.021409	-6.798**	-0.021253	-6.659**
C	Accom ( $\beta_4$ )	-	-	-	-	0.344466	4.911**	0.396669	5.375**
C	Sexm ( $\beta_5$ )	-	-	-	-	-	-	0.712224	3.514**
C	Age340 ( $\beta_6$ )	-	-	-	-	-	-	0.760922	3.882**
T	Tasc ( $\beta_7$ )	-0.672839	-3.722**	-0.775256	-2.382***	-0.352171	-1.825	0.505511	1.949
S	Sasc ( $\beta_8$ )	-1.633653	-9.998**	-1.691339	-7.519**	-1.211261	-7.519**	-0.270681	-1.054
B	Basc ( $\beta_9$ )	-1.479586	-9.009**	-1.545750	-6.407**	-1.056059	-5.760**	-0.130696	-0.506
L	Lasc ( $\beta_{10}$ )	-1.836485	-10.153**	-1.914775	-6.943**	-1.438421	-7.276**	-0.530279	-1.984***
<b>Summary Statistics</b>									
Number of observation		540		540		540		540	
Number of cases		2160		2160		2160		2160	
Log-likelihood at starting, $LL(0)$		-869.0965		-869.0965		-869.0965		-869.0965	
Log-likelihood at convergence, $LL(\beta)$		-713.3277		-713.2563		-698.4410		-682.9926	
Test of entire model, $-2[LL(0) - LL(\beta)]$		311.5376		311.6804		341.311		372.207	
Rho-squared with respect to zero, $\rho_0^2$		0.1792		0.1793		0.1964		0.2141	
Adjusted Rho-squared with respect to zero, $\bar{\rho}_0^2$		0.1769		0.1766		0.1937		0.2108	
value of time (VOT)		18,387 won/hour		16,200 won/hour		21,772 won/hour		23,614 won/hour	

<sup>a</sup>Indicates alternatives containing variables in their utility function; C (Car), T (Taxi), S (Subway), B (Bus), L (Limousine), \*\* Indicates 99 % level of confidence, \*\*\* indicates 95 % level of confidence.

In this research, two different rho-squared values can be calculated by using the equations  $\rho_0^2 = 1 - LL(\beta) / LL(0)$  and  $\bar{\rho}_0^2 = 1 - (LL(\beta) - K) / LL(0)$ , where  $LL(0)$  is the log-likelihood, with  $\beta_k = 0$ ,  $LL(\beta)$  is the log-likelihood for the estimated model and  $K$  is the number of degrees of freedom (i.e., parameters) used in the model. The rho-squared value is widely used for evaluating the goodness-of-fit for discrete choice models and generally has a value between 0 and 1. In

addition, when it closely approaches the value of 1, the model is considered to have a better goodness-of-fit. In addition, if  $\rho^2$  is positioned around or over at the value of 0.1, the model might be appropriate for use [16]. However, there are no guidelines for a “good” rho-squared value, and it may be more useful to compare the log-likelihood values calculated from each unique model specification in order to select a better model specification [16]. It is notable that the four models shown in Table 4 are estimated appropriately in terms of rho-squared values ranging from 0.1792 to 0.2141



Fig. (3). The scalar relationship of log-likelihood functions.

and other statistics calculated from the results. For further research, Model 4 was selected, because it is robust in terms of its log-likelihood and rho-squared values. In the following sections, by adopting the model specification for Model 4, trip distance and trip purpose models are estimated for each different airport user group, which are then divided corresponding to trip distance to the airport and original trip purpose for coming to the airport.

**5.2. Trip Distance Model**

In this section, the effects of the distance traveled by airport users on the mode choice behavior are explored. Initially, three types of distance models (e.g., short-, medium-, and long-distance models) were planned for estimation. However, the number of observations relating to short-distance travel was not sufficient to build a model, so only two models, standard (short + medium) and long distance, were estimated. The data are divided into different groups according to the travel distance. The travel distance covered by travelers was obtained by either referring to a survey response or by calculating the distance, using a web-provided map of origin-destination pairs. It should be noted that one criterion by which the travelers were divided into two distance groups is the distance of 40 km. Viewed in the context of the Korean ground transportation market for

airport access, the area within a 40 km radius from the location of the airport indicates the Seoul Metropolitan area inside of which travelers can more conveniently access transportation services than travellers coming from other parts of the country.

As can be seen from the results shown in Table 5, all parameters estimated for a standard-distance model appear to make sense in their sign and relative magnitudes, except for *Sasc* and *Basc*, and are important in the model specification at a 99% confidence level. The negative (-) sign for variables *TCost* and *TTime* suggests that the variables contribute to devaluing the utility of mode choice. In addition, the absolute value of the estimated parameter travel time is relatively larger, as compared to that of travel cost, and the result, if we appropriately understand the importance of the flight schedule, describes the situation faced by airport users quite well. Consequently, it appears to be true that travelers attach great significance to travel time. Three mode-specific variables specified only for the car mode (i.e., *Accom*, *Sexm*, and *Age340*) are estimated with a positive (+) sign. This result implies that car might be the first-choice alternative for a group of people traveling to the airport together. It is also shown that men are most likely to choose the car mode over other alternatives; especially if the

**Table 5. Trip Distance Models**

Trip Distance Models					
Alternatives <sup>a</sup>	Variable	Standard-Distance (SD)		Long-Distance (LD)	
		Parameter Estimates	t-Value	Parameter Estimates	t-Value
C, T, S, B, L	<i>TCost</i> ( $\beta_0$ )	-0.000086	-4.578**	-0.000029	-0.948
C, T, S, B, L	<i>TTime</i> ( $\beta_3$ )	-0.009868	-2.750**	-0.065600	-5.378**
C	<i>Accom</i> ( $\beta_4$ )	0.327949	4.377**	1.401363	2.544***
C	<i>Sexm</i> ( $\beta_5$ )	0.752958	3.615**	1.160811	0.806
C	<i>Age340</i> ( $\beta_6$ )	0.798452	3.981**	-0.896289	-0.671
T	<i>Tasc</i> ( $\beta_7$ )	0.709575	2.666**	0.096459	0.066
S	<i>Sasc</i> ( $\beta_8$ )	-0.247276	-0.912	-1.308978	-0.928
B	<i>Basc</i> ( $\beta_9$ )	-0.444178	-1.630	1.266745	0.906
L	<i>Lasc</i> ( $\beta_{10}$ )	-0.750766	-2.671**	0.185195	0.123
Summary Statistics					
Number of observation		489		51	
Number of cases		1956		204	
<i>LL</i> (0)		-787.0151		-83.6908	
<i>LL</i> ( $\beta$ )		-629.3104		-25.7664	
$-2[LL(0) - LL(\beta)]$		315.4094		115.8488	
$\rho_0^2$	$\bar{\rho}_0^2$	0.2004	0.1967	0.6921	0.6782
Value of time (VOT)		6,885 won/hour		135,724 won/hour	

<sup>a</sup>Indicates alternatives containing variables in their utility function; C (Car), T (Taxi), S (Subway), B (Bus), L (Limousine), \*\* Indicates 99 % level of confidence, \*\*\* indicates 95 % level of confidence.

decision-makers are between the ages of 30 and 40. If all other conditions are equal, the taxi and car modes are preferable to any other modes. In the context of the complicated transportation network within the Seoul Metropolitan area, the effect of the difficulty in transferring between modes appeared to be quite strong.

On the other hand, only two variables, *TTime* and *Accom*, estimated for the long-distance (LD) model appear to be statistically significant at confidence level greater than 95%. The reason for such poor results could possibly be the failure to satisfy the condition for discrete choice analysis that the number of observations must be more than 200, in order to obtain a reasonable estimation [16]. However, in spite of the disadvantage caused by the observation number, the overall goodness-of-fit  $\rho^2$  appeared to be the highest value among all the models. The two statistics mentioned above do not immediately make sense, and although this phenomenon must have cause, we could not find a clear explanation in this research. As travel distance increases, people may find that their travel becomes increasingly difficult, and the strongest evidence to support this tendency can be obtained in the results of this research by comparing the absolute value of the parameter of travel time for the two distance models (i.e., LD and SD models). Further proof in support of this phenomenon can be also found in a study on travel time value [20], which stated that long-distance travelers attach much more significance to the amount of travel time than do short-distance travelers. In a similar vein, as shown in Fig. (4), value of travel time (VOT) estimated from the long-distance model has a much larger value than that obtained from the standard-distance (SD) model. The result can be considered as a support for the assumption made for travel time from an experimental point of view. In comparison with the values obtained from the other models studied in the previous section, a VOT for the long-distance (LD) model might be overestimated; in contrast, a VOT for the standard-distance (SD) model might be underestimated. Other noteworthy aspects of the results shown in Table 5 from the long-distance model are that, if all conditions are equal, either a bus or limousine might be the first-choice alternative for long-distance travelers, due to their

competitive pricing in the Korea airport access transportation market, and a car is not a preferred mode, regardless of whether the decision-makers are between the ages of 30 and 40.

### 5.3. Trip Purpose Model

In this section, the effects of the airport visitation purpose of several airport user groups on mode choice behavior are explored. The results of the trip purpose model, consisting of departing (D) and non-departing (ND) models, are summarized in Tables 6 and 7, respectively. In addition, the departing model is divided into two different models, e.g., business (B) and tourism (T) models, according to their departure purpose. The non-departing model is also divided into two different models, e.g., see-off (SO) and welcome (W) models, according to their airport visitation purpose.

As can be seen from the results shown in Table 6, all parameters estimated for the departing (D) model and its sub-models, business (B) and tourism (T), appear to make sense in their sign and relative magnitudes; however, only a few parameters for each model should be included in the model specification at a confidence level higher than 95%. The meaning of a negative (-) sign and the relative magnitudes of both generic variables could be interpreted in the same manner discussed previously, except that business travelers attach more importance to travel time than do other travelers. It can be seen from a business (B) model that business travelers are more likely to choose a taxi if all other conditions are equal. Those who are between the ages 30 and 40 would not prefer a car as an access mode. As shown in Fig. (5), the value of travel time (VOT) for business travelers is approximately 2.5 times as large as the value for other travelers. However, since these results might be inadequate from a statistical viewpoint, i.e., low goodness-of-fit measures, it is advisable to use the results with caution, only for outlining the mode choice behavior of airport user groups traveling to the airport for departure.

In contrast, as shown in Table 7, the non-departing (ND) model and its sub-models (i.e., the see-off (SO) and welcome (W) models) seem to fit well with the given data in terms of the overall goodness-of-fit measure  $\rho^2$ . In addition, all parameters except for the alternative specific constants of

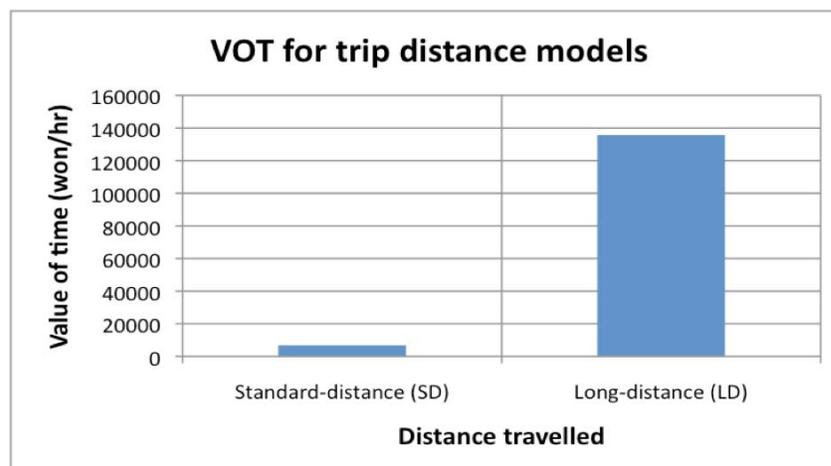


Fig. (4). Value-of-time estimated for the trip distance models

**Table 6. Trip Purpose Model (Departing)**

Trip Purpose Model (Departing)							
Alternatives <sup>a</sup>	Variable	Departing (D)		Business (B)		Tourism (T)	
		Parameter Estimates	t-Value	Parameter Estimates	t-Value	Parameter Estimates	t-Value
C, T, S, B, L	$TCost (\beta_0)$	-0.000022	-1.599	-0.000017	-0.919	-0.000026	-1.611
C, T, S, B, L	$TTime (\beta_1)$	-0.014078	-3.384 **	-0.015882	-2.824 **	-0.010108	-1.614
C	$Accom (\beta_4)$	0.080655	1.362	0.012194	0.268	0.840127	3.845 **
C	$Sexm (\beta_5)$	0.130941	0.430	0.510831	0.924	0.284490	0.585
C	$Age340 (\beta_6)$	0.176536	0.594	-0.288138	-0.674	0.831026	1.702
T	$Tasc (\beta_7)$	0.113784	0.352	0.042501	0.071	1.452544	2.845 **
S	$Sasc (\beta_8)$	-0.792709	-2.321 ***	-0.688231	-1.109	0.423816	0.790
B	$Basc (\beta_9)$	-0.394785	-1.212	-0.048204	-0.082	0.328959	0.598
L	$Lasc (\beta_{10})$	-0.823197	-2.417 ***	-0.404916	-0.677	-0.173527	-0.293
Summary Statistics							
Number of observation		231		126		105	
Number of cases		924		504		420	
$LL(0)$		-371.7802		-202.7892		-168.9910	
$LL(\beta)$		-346.8068		-192.7448		-137.8918	
$-2[LL(0) - LL(\beta)]$		49.9468		20.0888		62.1984	
$\rho_0^2$	$\bar{\rho}_0^2$	0.0672	0.0580	0.0495	0.0322	0.1840	0.1662
Value of time (VOT)		38,395 won/hour		56,054 won/hour		23,326 won/hour	

<sup>a</sup>Indicates alternatives containing variables in their utility function; C (Car), T (Taxi), S (Subway), B (Bus), L (Limousine), \*\* Indicates 99 % level of confidence, \*\*\* indicates 95 % level of confidence.

each model should be considered as variables that significantly affect model specifications at a confidence level higher than 95%. The signs of the estimated model and the relative magnitudes between variables are similar to those in the previous section, and thus, the same interpretation can be applied using a similar line of reasoning. The results presented in Figs. (5, 6) show that the departing (D) group attaches more importance to travel time than the non-departing (ND) group. Based on the result, the value of time for the departing model (D) is approximately 6.9 times as large as the value of time for the non-departing model (ND). If all conditions are equal, a taxi would be the first choice for the non-departing airport user group.

## 6. CONCLUSION

As stated in the introduction, the central purpose of this research is to explore the effect of various airport users grouped according to their travel distance and purpose on mode choice behavior in the context of airport access travel at a passenger terminal of the Gimpo International Airport in South Korea. For this purpose, we used revealed preference (RP) data gathered during on-site-surveys conducted by the Korea Transport Institute (KOTI) and targeted at people who traveled to the airport using ground access. As an analysis

tool, we employed a multinomial logit (MNL) discrete choice analysis technique with the help of the econometric software package LIMDEP ver. 7.0. After trying to structure a wide variety of model specifications formulated with explanatory variables in a linear-in-parameters relationship, four model specifications were selected based on their goodness-of-fit measures, in order to conduct a basic analysis of the general characteristics of airport users and to discuss the statistics involved in the estimation results. Only the best model was adopted to estimate the trip distance models (standard-distance (SD) and long-distance (LD)) and the trip purpose models (departing (D) and non-departing (ND)), in order to identify the different characteristics of mode choice behaviour occurring in each specific airport user group.

The behavioral characteristics of the various airport user groups in terms of mode choice were identified from the results of the four basic models for ground airport access mode choice. These characteristics are as follows: (a) People who travel to the airport show a high sensitivity to travel time. This result reflects the fact that the first priority of these travelers is to arrive at the airport in a timely manner and to then proceed with their intended tasks such as catching a flight, waiting for a person, and so forth, making travel time a critical factor of concern. (b) Cars are the

**Table 7. Trip Purpose Model (Non-Departing)**

Trip Purpose Model (Non-Departing)							
Alternatives <sup>a</sup>	Variable	Non-Departing (ND)		See-Off (SO)		Welcome (W)	
		Parameter Estimates	t-Value	Parameter Estimates	t-value	Parameter Estimates	t-Value
C, T, S, B, L	<i>TCost</i> ( $\beta_0$ )	-0.000260	-5.348 **	-0.000260	-4.351 **	-0.000286	-3.176 **
C, T, S, B, L	<i>TTime</i> ( $\beta_1$ )	-0.023996	-3.787 **	-0.020370	-2.380 ***	-0.026764	-2.739 **
C	<i>Accom</i> ( $\beta_4$ )	0.791389	5.191 **	0.777093	4.141 **	0.961117	3.166 **
C	<i>Sexm</i> ( $\beta_5$ )	1.430175	4.214 **	1.317745	2.903 **	1.500997	2.819 **
C	<i>Age340</i> ( $\beta_6$ )	1.402384	4.270 **	1.469986	3.267 **	1.260741	2.526 ***
T	<i>Tasc</i> ( $\beta_7$ )	1.570179	3.221 **	1.907826	3.039 **	1.066819	1.296
S	<i>Sasc</i> ( $\beta_8$ )	-0.497986	-1.046	-0.438279	-0.709	-0.676201	-0.837
B	<i>Basc</i> ( $\beta_9$ )	-0.849301	-1.709	-0.558052	-0.902	-1.437662	-1.631
L	<i>Lasc</i> ( $\beta_{10}$ )	-0.188564	-0.405	-0.572400	-0.864	0.086031	0.121
Summary Statistics							
Number of observation		278		155		123	
Number of cases		1112		620		492	
<i>LL</i> (0)		-447.4237		-249.4629		-197.9609	
<i>LL</i> ( $\beta$ )		-236.1401		-130.8487		-101.5985	
-2[ <i>LL</i> (0) - <i>LL</i> ( $\beta$ ) ]		422.5672		237.2284		192.7248	
$\rho_0^2$	$\bar{\rho}_0^2$	0.4722	0.4679	0.4755	0.4678	0.4868	0.4772
Value of time (VOT)		5,538 won/hour		4,700 won/hour		5,615 won/hour	

<sup>a</sup>Indicates alternatives containing variables in their utility function; C (Car), T (Taxi), S (Subway), B (Bus), L (Limousine), \*\* Indicates 99 % level of confidence, \*\*\* indicates 95 % level of confidence.

preferred mode for males between the ages of 30 and 40. In addition, cars are also preferable for people accompanied by a number of companions. (c) The value of time (VOT) estimated from the four basic models ranged from 16,000 to 24,000 (won/hour). This value is approximately 1.6 to 2.4 times as large as the value of time obtained for people who travel with a purpose other than going to the airport in a different transportation context. This result implies that travelers who attach strong significance to their travel, such as people heading to the airport with a specific purpose, are willing to pay more money to save travel time.

The trip distance model used to increase the understanding of the effects of travel distance on mode choice behavior was explored by sorting data according to travel distance and by then estimating two models, a standard distance (SD) and a long distance model (LD). The results obtained from the distance models are as follows: (a) People who travel long distances attach more importance to travel time than do standard-distance travelers. This result reflects the fact that people who start their travel within the Seoul metropolitan area are probably exposed to many possible transportation options, and thus, they are more likely to reach the airport in a timely manner. In contrast, people who travel long distances, coming from regions outside of the Seoul metropolitan area can be assumed to

experience difficulty in either transferring or waiting, due to the complexity of the transportation networks required to arrive at the airport. As a consequence of the ground airport access mechanism mentioned above, the situation faced by long-distance travelers would make them more sensitive to travel time. (b) The results of the value of time analysis showed that long-distance travelers are willing to pay more to save travel time. However, it should be noted that the value of time for the standard-distance (SD) model is underestimated and that the value of time for the long-distance (LD) model is overestimated when compared to the values of the four basic models.

The trip purpose model used for deepening the understanding of the effects of the trip purpose of a variety of airport user groups on mode choice behaviors was explored by sorting data according to travel purpose and by then estimating two models, a departing (D) and a non-departing model (ND). The results obtained from these two models are as follows: (a) People generally prefer cars and taxis to other modes, irrespective of their trip purpose. (b) People who are between the ages of 30 and 40 and who travel to the airport for business trips are not likely to use a car as their access mode. One possible reason for this result is that parking is too costly. (c) As for the value of time, the results showed that the value for the departing (D) model is

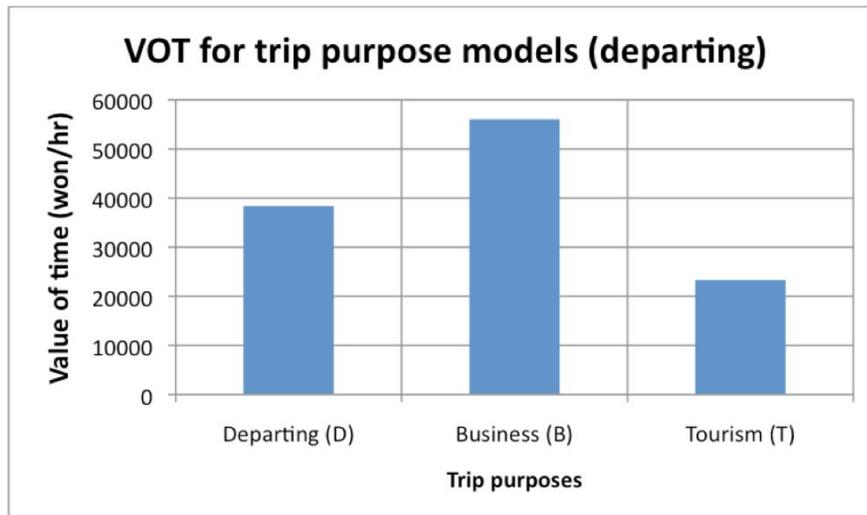


Fig. (5). Value-of-time estimated for the trip purpose models (departing)

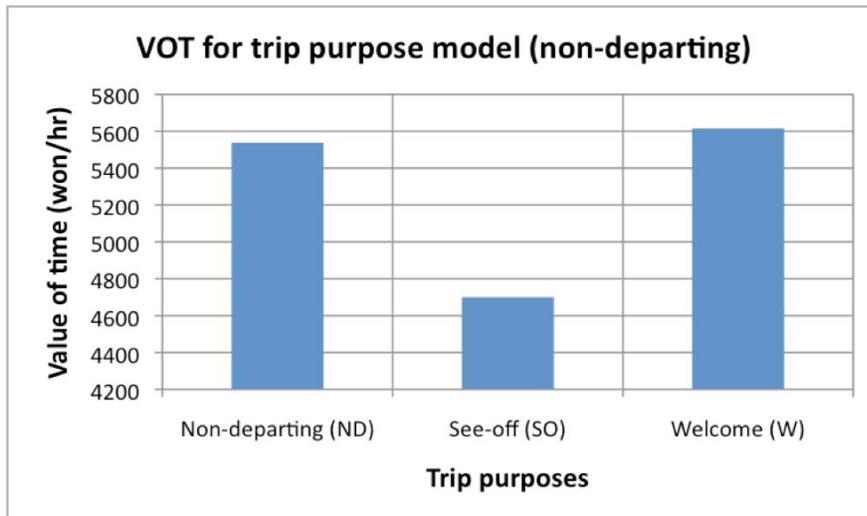


Fig. (6). Value-of-time estimated for the trip purpose models (non-departing).

approximately 6.9 times larger than that of the non-departing (ND) model. More importantly, the value of the business (B) model is approximately 2.4 times as large as that of the tourism (T) model.

### CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

### ACKNOWLEDGEMENTS

The author is grateful to the Korea Transport Institute (KOTI) for providing related data for this research. The opinions and views presented here are solely those of the author.

### REFERENCES

- [1] V. Psaraki, C. Abacoumkin, "Access mode choice for relocated airports: the new Athens international airport", *Journal of Air Transport Management*, vol. 8, no. 2, pp. 89-98, 2002.
- [2] G. Harvey, "Study of airport access mode choice", *Journal of Transportation Engineering*, vol. 112, no. 5, pp. 525-545, 1986.
- [3] M. Ben Akiva, D. McFadden, K. Train, J. Walker, C. R. Bhat, M. Bierlaire, D. Bolduc, A. Boersch Supan, D. Brownstone, D. Bunch, A. Daly, A. de Palma, D. Gopinath, A. Karlstrom, M. A. Munizaga, "Hybrid choice models: progress and challenges", *Marketing Letters*, vol. 13, no. 3, pp.163-175, 2002.
- [4] C. R. Bhat, "Accommodating flexible substitution patterns in multidimensional choice modelling: formulation and application to travel mode and departure time choice", *Transportation Research Part B*, vol. 32, no. 7, pp. 455-466, 1998a.
- [5] C. R. Bhat, "Accommodating variations in responsiveness to level-of-service measure in travel mode choice modelling", *Transportation Research Part A*, vol. 32, no. 7, pp. 495-507, 1998b.
- [6] H. Ohashi, T.-S. Kim, T. H. Oum, C.Yu, "Choice of air cargo transshipment airport; an application to air cargo traffic to/from Northeast Asia", *Journal of Air Transportation Management*, vol. 11, no. 3, pp. 148-159, 2005.
- [7] J. D. Hunt, J.E. Abraham, *Stated response examination of factors influencing commercial movement route choice behavior, Applications of Advanced Technologies in Transportation Engineering*, 2004, pp. 72-77.
- [8] D. Revelt, K. Train, "Mixed logit with repeated choice: households' choice of appliance efficiency level", *Review of Economics and Statistics*, vol. 80, no.4, pp. 647-657, 1998.
- [9] D. Brownstone, K. Train, "Forecasting new product penetration with flexible substitution patterns", *Journal of Econometrics*, vol. 89, no. 1-2, pp. 109-129, 1999.
- [10] M.R. Tayyaran, A.M. Khan, "Telecommuting and residential location choice decisions: combined stated and revealed preference model", *Canadian Journal of Civil Engineering*, vol. 34, no. 10, pp. 1324- 1333, 2007.

- [11] G. Harvey, "Airport choice in a multiple airport region", *Transportation Research Part A*, vol. 21A, no. 6, pp. 439-449, 1987.
- [12] S. Hess, J. W. Polak, "Mixed logit modeling of airport choice in multi-airport regions", *Journal of Air Transportation Management*, vol. 11, no. 2, pp. 59-68, 2005.
- [13] Mei Ling T., Mei Lam T., William H. K. L., "Analysis of airport access mode choice: A case study in Hong Kong", *Journal of the Eastern Asia Society for Transportation Studies*, vol. 6, pp. 708 – 723, 2005.
- [14] D. A. Tsamboulas, A. Nikoleris, "Passengers' willingness to pay for airport ground access time savings", *Transportation Research A*, vol. 42, no 10, pp. 1274-1282, 2008.
- [15] K. Train, *Discrete Choice Methods with Simulation*. Cambridge University Press: Cambridge, 2003.
- [16] F. S. Koppelman, C. R. Bhat, *A Self Instructing Course in Mode Choice Modelling: Multinomial and Nested Logit Models*. U.S. Department of Transportation: Federal Transit Administration, 2006.
- [17] M. Ben Akiva, S. Lermans, *Discrete Choice Analysis: Theory and Application to Travel Demand*. The MIT Press: Cambridge, 1985.
- [18] D. A. Hensher, J. M. Rose, W. H. Greene, *Applied Choice Analysis: A Primer*. Cambridge University Press: NY, 2005.
- [19] William H. Greene, *LIMDEP User's Manual and Reference Guide*. Econometric Software, Inc. Ver. 7.0., 1997.
- [20] C.H. Oglesby, R.G. Hicks, *Highway Engineering. 4<sup>th</sup> ed.*, John Wiley and Sons: New York, 1982.

---

Received: August 26, 2013

Revised: September 16, 2013

Accepted: September 25, 2013

© Hyuk-Jae Roh; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.