INCLUSION OF THE LATENT PERSONALITY VARIABLE IN MULTINOMIAL LOGIT MODELS USING THE 16PF PSYCHOMETRIC TEST

INCLUSIÓN DE LA VARIABLE LATENTE PERSONALIDAD EN MODELOS LOGIT MULTINOMIAL UTILIZANDO PRUEBA PSICOMÉTRICA 16PF

JORGE E. CÓRDOBA MAQUILÓN

PhD. Professor, National University of Colombia, Medellin, jecordob@unal.edu.co

G. PATRICIA JARAMILLO ÁLVAREZ

PhD. Professor, National University of Colombia, Medellin, gpjarami@unal.edu.co

ABSTRACT: Travel demand models typically use modal attributes and socioeconomic characteristics as explanatory variables.It has been established that attitudes and perceptions as well as individual psychological variables influencea user's behavior. In this study, the latent personality variable was included in the estimation of a hybrid discrete choice model to incorporate the effects of subjective factors. The latent personality variable was assessed with the 16PF psychometric test,which has been widely use by researchersworldwide. The paper analyzes the results of applying this model to a sample of employees and university professors and proposes a way in which the psychometric tests can be used in hybrid discrete choice models. Our results show that hybrid models that include latent psychological variables are superior to traditional models that ignore the effects of user'sbehavior.

KEY WORDS: Latent psychologicalvariable, Discrete choice model, Multinomial logit model, Personality, psychometric test

RESUMEN: Los modelos de demanda de viajes utilizan principalmente los atributos modales y las características socioeconómicas como variables explicativas. También se ha establecido que las actitudes y percepciones influyen en el comportamiento de los usuarios. Sin embargo, las variables psicológicas del individuo condicionan la conducta del usuario. En este estudio se incluyó la variable latente personalidad, en la estimación del modelo híbrido de elección discreta, el cual constituye una buena alternativa para incorporar los efectos de los factores subjetivos. La variable latente personalidad se evaluó con la prueba psicométrica 16PF de validez internacional. El artículo analiza los resultados de la aplicación de este modelo a una población de empleados y docentes universitarios, y también propone un camino para la utilización de pruebas psicométricas en los modelos híbridos de elección discreta. Nuestros resultados muestran que los modelos híbridos que incluyen variables latentes psicológicas son superiores a los modelos tradicionales que ignoran los efectos de la conducta de los usuarios.

PALABRAS CLAVE: Variable latente psicológica, Modelo de elección discreta, Modelo logit multinomial, Personalidad, Prueba psicométrica

1. INTRODUCTION

Travel demand models typically use modal attributes and socioeconomic characteristics as explanatory variables of the choice modal. It has been established that attitudes and perceptions influence auser'sbehavior in the choice modal, and in the last decade, hybrid discrete choice models, which include a latent variable model and a discrete choice model, have been developed that can account for attitudes and perceptions as well as modal attributes and socioeconomic characteristics [1].

Furthermore, the inclusion of latent variables improves the fit of these choice models as seen in [2]. Traditional discrete choice models have been enriched with the construction of latent variables by [3-9], but the research carried out by [10] released the first complete methodology for the inclusion of latent variables in discrete choice models. However, there have been no studies thus far that use a personality variable, such as a psychological aspect,for modal choice. Furthermore, psychometric tests have not been used to measure the latent indicators of latent variables.

To estimate the hybrid discrete choice model with the latent personality variable, a sequential approach was used in which the latent variable was first constructed previous to the estimation by the multiple indicator multiple cause (MIMIC) model and then included in the discrete choice model as a regular variable [8].

The main objectives of this research were to estimate a hybrid discrete choice model to include psychological issues, such as personality, and psychometric tests to measure latent variables. The results of an application of this model to a population of employees and professors in auniversity of Medellin (Colombia) were reviewed, and a methodology for the use of psychometric tests in the hybrid discrete choice models was proposed. Our results show that hybrid models that include psychological latent variables are superior to traditional models that ignore the effects of behavior.

The paper is structured as follows: Section 2 summarizes the theoretical framework for estimating discrete choice econometric models, latent variables models, hybrid discrete choice models, the theory of personality, and 16PF psychological testing. Section 3 presents the proposed model. Section 4 shows the results of the model application to a population of employees and university professors, and Section 5 presents the conclusions and main findings of the research.

2. THEORETICAL FRAMEWORK

A model is the simplified representation of reality with a mathematical framework because it takes the most representative variables of a system and evaluates theirimpacts on the systemby testing several alternatives. In this section, econometric discrete choice models and latent variables models are presented.

2.1. Econometric discrete choice models

There is a microeconomic analysis of consumer behavior based on the fundamental assumption that the rational consumer will always choose the combination of alternatives more useful for him among those belonging to the set of feasible alternatives. This analysis includes psychological variables such as personality.

The set of feasible alternatives for the set of all combinations that consumers can choose, $p = (p_1,$

$p_2, ..., p_k)$, is the vector of prices of all goods X and the income Iavailable to consumer q. The set of possible combinations is given by equation (1).

$$A(q) = \{x \in X: p\ x \leq I\} \qquad (1)$$

Thus, the problem facing the consumer can be expressed as equation (2).

$$Max\ U(x)s.t.\ p\ x \leq I \qquad x \in X \qquad (2)$$

The random utility theory [11] was used, for estimate discrete choice modelwhich statesthat individuals belonging to a certain homogeneous population Q act rationally and have perfect information.

There must be a set $A = \{A_1, A_i, A_j\}$ of

alternatives available that meet three characteristics. First, they must be mutually exclusive from the perspective of the decision makeruser's; thus,selecting an alternative does not necessarily imply the selection of any other alternative.Second, the set must be exhaustive, such that all of the possible alternatives are included and the individual must choose one. Third, the number of alternatives must be finite because discrete choice models cannot be applied otherwise.

The set of alternatives available to an individual q is A(q); this sethas an associated set of attributes x ∈ X. The set of alternatives available to each individual is assumed to incorporate the effect of their restrictions.

Each alternative $A_i \in A$ has an associated utility U_{iq} for individual q. As an observer, the modelerhas no full information on all of the factors considered by individuals when making their choice. Therefore, it is assumed that this utility can be represented by two components:

A deterministic component, called the systematic or representative utility V_{iq}, which is a function of the measured attributes X. A linear additive function in the parameters $V_{iq} = \sum \theta_{ikq} * X_{ikq}$is typically used,where X_{ikq} represents the attribute value k of alternative A_i for individual q. Parameters θ are assumed to be constant for all individuals and may vary between alternatives. These parametersare obtained through an estimation process, such as the maximum likelihood method,

where the observations of the choices made by a sample of individuals are consistent with the model.

A random component ε_{iq} that reflects the idiosyncrasies and preferences of each individual, as well as measurement errors and observation errors by the modeler. The errors ε aretypically assumed to be random variables with zero mean and a specified probability distribution [12]. Thus, this random component is expressed as Eq.(3).

$$U_{iq} = V_{iq} + \varepsilon_{iq} \tag{3}$$

Individual q chooses the most useful alternative, i.e., choiceA_i if and only if equation (4) is satisfied.

$$U_{iq} \geq U_{jq}\,\forall A_j \in A(q) \tag{4}$$

This expressioncan be rewritten in terms of components expressed asEq. (5)

$$V_{iq} - V_{jq} \geq \varepsilon_{jq} - \varepsilon_{iq}\,\forall A_j \in A(q) \tag{5}$$

Becauseε_{jq}-ε_{iq} is unknown, it is not possible to determine whether the above relationship is true; therefore,probabilities are assigned. The probability that individual q chooses alternative i is given by equation (6).

$$P_{iq} = Prob\{\varepsilon_{jq} \leq \varepsilon_{iq} + (V_{iq} - V_{jq}), \forall A_j \in A(q)\} \tag{6}$$

If $f(\varepsilon) = f(\varepsilon_1, ..., \varepsilon_N)$is the distribution function of the random variables,

$$P_{iq} = \int_{\varepsilon_{iq}=-\infty}^{\infty} \int_{\varepsilon_{1q}=-\infty}^{V_{iq}-V_{1q}+\varepsilon_{iq}} .... \int_{\varepsilon_{Nq}=-\infty}^{V_{iq}-V_{Nq}+\varepsilon_{iq}} f(\varepsilon_{1q}, \varepsilon_{2q}, ..., \varepsilon_{Nq}) d\varepsilon_{1q}.....d\varepsilon_{Nq} \tag{7}$$

Therefore, the probability of selecting a certain alternative is a multidimensional integral over the density of the unobserved portion of utility. Different models can be obtained based on the assumptions made about the distribution of ε.

Multinomial Logit Model (MNL)

The MNL is used in this research and is the most widely used discrete choice model becauseEq. (7) of the integral of the choice probability formula is closed.

This model is obtained by assuming that the error terms are independently and identically distributed (i.i.d.) with a Gumbel distribution. This distribution It is also known as extreme value, type I extreme value, and Weibull, with zero mean and variance σ^2. Therefore, the terms are uncorrelated and have the same variance for each alternative and each individual. Thus, the probability that an individual $q$selects alternative i is given by Eq.(8).

$$P_{iq} = \frac{\exp(\mu \cdot V_{iq})}{\displaystyle\sum_{A_j \in A(q)} \exp(\mu \cdot V_{jq})} \tag{8}$$

Whereμ is a scaling factor related to the variance of the error term, $\mu = \pi/(\sqrt{6}\sigma)$.

This factor is typically unidentifiable, so it is necessary to set this value (scaling of the variance covariance matrix). It is typically assumed that$\mu = 1$, implying that $\sigma^2 = \pi^2/6$.

2.2. Modeling with latent variables

Latent variables are abstract variables representing the subjective elements in the choice conduct; they cannot be measured directly, so they are expressed by only the individual through latent indicators. The methodology developed by [1] is presented here to incorporate latent variables as explanatory factors in discrete choice models.

2.2.1. *Structural equations for the latent variable model*

The distribution of the latent variables given the observed variables is expressed asEq. (11)

$$f_1\left(X_n^* \mid X_n; \lambda, \Sigma_\omega\right) \tag{11}$$

Where:

X_n= Observable variable

X_n^*= Latent variable

λ=Unknown parameter

Σ_ω= Covariance of the error term

Thus, the equation derived from this function for each latent variable can be expressed asEq. (12)

$$X_n^* = h\left(X_n;\lambda\right)+\omega_n \quad \text{and} \quad \omega_n \sim D\left(0,\Sigma_\omega\right) \tag{2}$$

Where h is a function to be defined and is linear in its parameters. The other variables are defined above. The distribution of the error term ω must also be specified.

2.2.2. *Structural equations for the discrete choice model*

The distribution of utilities is expressed asEq. (13)

$$f_2\left(U_n \mid X_n, X_n^*;\beta,\Sigma_\varepsilon\right) \tag{13}$$

The equation that is derived from this function can be expressed as

$$U_n = V_n\left(X_n, X_n^*;\beta\right)+\varepsilon_n \quad \text{and} \quad \varepsilon_n \sim D\left(0,\Sigma_\varepsilon\right) \tag{14}$$

In Eq. (14), the random utility consists of systematic utility and random errors. Similarly,V is a function to be defined and is linear in its parameters. The other variables are defined above. The error term distribution must also be specified.

2.2.3. *Measurement equations for the latent variable model*

The conditional distribution of the indicators for the latent variables is expressed asEq. (15)

$$f_3\left(I_n \mid X_n, X_n^*;\alpha,\Sigma_\upsilon\right) \tag{15}$$

Where:

I_n=Indicator of X_n^*

α=Unknown parameter

Σ_υ= Covariance of the error term

Thus, the equation derived from this function for each survey question (indicator) can be expressed asEq. (16)

$$I_n = m\left(X_n, X_n^*;\alpha\right)+\upsilon_n \quad \text{and} \quad \upsilon_n \sim D\left(0,\Sigma_\upsilon\right) \tag{16}$$

Whereυ_n is the error term.

Similarly, m is a function to be defined and is linear in its parameters. The other variables are defined above. The error term distribution must also be specified.

Finally, the following expression Eq. (17) provides the choice based on the utilities (assuming utility maximization):

$$Y_{in} = \begin{cases} 1, & \text{if } U_{in} = \max_j\left\{U_{jn}\right\} \\ 0, & \text{otherwise} \end{cases} \tag{17}$$

2.3. **Hybrid discrete choice modeling**

The latent variable model consists of equations (12) and (16), and the choice model consists of the equations (14) and (17). With these last two equations and an assumed error (ε_n) distribution, the derived conditional choice probability with observable and latent variables is expressed asEq. (18)

$$P\left(Y_n \mid X_n, X_n^*;\beta,\Sigma_\varepsilon\right) \tag{18}$$

The maximum likelihood method is used to estimate the unknown parameter. The easiest method for creating the likelihood function for the integrated model is to start with the probability of a choice model without latent variables, as shown in Eq. (19)

$$P\left(Y_n \mid X_n;\beta,\Sigma_\varepsilon\right) \tag{19}$$

The latent variables are then added to the choice model, resulting inEq. (20)

$$\begin{aligned} &P\left(Y_n \mid X_n;\beta,\lambda,\Sigma_\omega,\Sigma_\varepsilon\right)= \\ &\int_{X_n^*} P\left(Y_n \mid X_n, X_n^*;\beta,\Sigma_\varepsilon\right)f_1\left(X_n^* \mid X_n;\lambda,\Sigma_\omega\right)dX_n^* \end{aligned} \tag{20}$$

Equation (20) is the integrated model's probability function, which is equal to the integral of the choice model for the distribution function of the latent constructs.

Assuming linearity in the parameters and normally distributed errors for the choice of alternative i, the choice model of the probability function is a standard choice model, except that utility is a function of latent constructs, as shown in Eq. (21) and (22).

$$U_{in} = V_{in}+\varepsilon_{in} \quad \text{And} \quad V_{in} = V_{in}\left(X_n, X_n^*;\beta\right)$$

$$i \in A(n), \ A(n) \text{ choice set} \tag{21}$$

$$P\left(Y_{in} = 1 \mid X_n, X_n^*;\beta\right)=\frac{e^{V_{in}}}{\sum_{j\in C_n} e^{V_{jn}}} \tag{22}$$

2.4. Theory of the Eysenck personality

A study on the theory of personality [13] through the factorial model seeks intermediate variables that explain differences in behavior in similar situations, as well as the consequences of such behavior. The theory defines personality as the sum of the behavior patterns and potential of the organism, both of which are determined by heredity and the social environment in which the organism originated and developed through the functional interaction of four main factors: (a) cognitive sector (intelligence), (b) conative sector (character), (c) affective sector (temperament), and (d) somatic sector (constitution).

Using the theory of Eysenck, Cattell [14] made extensive use of the factorial analysis method and isolated 16 personality factors, which he brought together in a psychometric test called 16PF. The most relevant aspects of this test are presented below.

2.5. 16PF Psychometric Testing

This test consists of 187 questions evaluating 16 factors, each of which is measured in decatypes (a score of 1 to 10).These factors are described below [14].

Intellectual area (B)

Personal area (A - E - H - I - M - N-O)

Emotional area (C - G - Q3 - Q1 - Q4)

Social area (F - L - N - Q2)

FactorA (reserved–open) measures the individual's gregarious nature, defined as the degree to which the person seeks to establish contact with other people because they find satisfying and rewarding relationships through them. **Factor B** (concrete thinking–abstract thinking) measures intelligence based on the predominance of abstract or concrete thinking, where abstract thinkingis characteristic of a person of higher intelligence and concrete thinking is an indicator of lower intelligence. **Factor C** (emotional instability–emotional stability) is related to the emotional stability of the person and the way in which he adapts to his environment. This factor specifically determines the strength of the ego.

Factor E (submissive–dominant) measures the degree of control that the person tends to hold in their relationships with other human beings and is determined in terms of whether the person is dominant or submissive. **Factor F** (prudent-impulsive) is related to the level of enthusiasm evident in social contexts. **Factor G** (carefree-scrupulous) measures the internalization of moral values. It structurally explores the superego. **Factor H** (shy-spontaneous) measures the reactivity of the nervous system based on the parasympathetic or sympathetic dominance trends of the person.

Factor I (rational-emotional) is used to measure the prevalence of either feeling or rational thought in making decisions for behaving in everyday life.

Factor L (trusting-suspicious) explores the social identity of the individual and specifically measures the degree to which the person is identified or linked to the human race in general.

Factor M (practical-dreamer) is based on the observation that humans can perceive things in two ways. The first way is to receive feed from direct contact between the senses and the environment. The other way is composed mostly of a subliminal connection of thoughts and speculations through which information is organized.

Factor N (single-sly) is related to social masks and describes the extent to which people are hidden, showing only those features that generate the answers you want from others.

Factor O (safe-unsafe) explores the self-esteem of those trends based on experience, guilt, or insecurities. This factor is not intended to categorize people by high and low self-esteem, as the level at the time of the test may be a transient because it is influenced by recent events.

Finally, **Factor Q1** (traditionalist-innovative) explores the psychological orientation toward change. **Factor Q2** (dependence on the groupself-sufficient) measures the degree of dependence on the person. **Factor Q3** (uninhibited-controlled) explores the efforts of the individual to maintain congruence between their ideal and real selves, molding according to standards

established and approved by society. **Factor Q4** (calm-stressed) measures the unpleasant sensations that tend to accompany the excitation of the autonomic nervous system, commonly known as stress.

3. PROPOSED MODEL

Córdoba et al. [15] developed a discrete choice model using psychological variables. The authors found that anxiety affects the choice of urban transportation mode and show that physiological alterations, as well as problems in perception and beliefs, can affect the decision-making process. However, they did not include the personality variable in their model. Based on that, this research is about building a more realistic choice behavior model that incorporates latent constructs such as personality.

The responses to the questions to the 16PF personality psychometric test are used as indicators of the latent psychological aspects (see Figure 1).
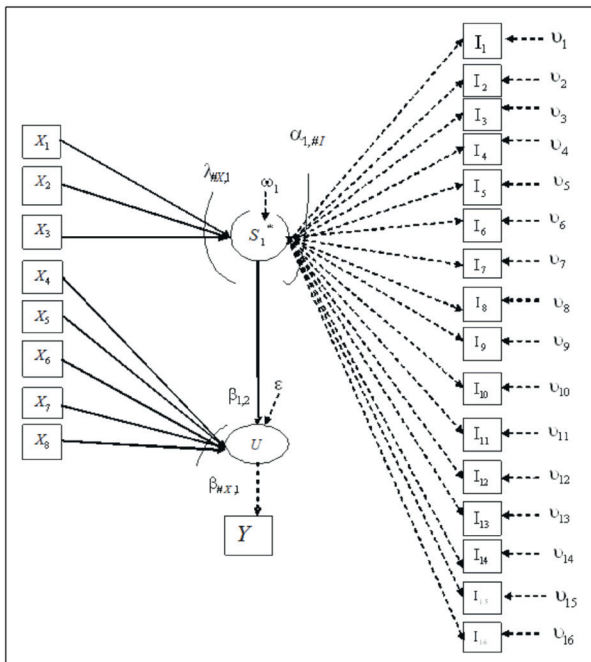

Figure 1. Integrated choice model with latent variable Personality

There are 3 exogenous explanatory variables and 16 indicators for the latent personality variable from the 16PF psychometric test assessment. The model equations are given by Eq. (23).

Structural equation model:

$$S_{ln}^{*} = X_n \lambda_l + \omega_{ln} \; {}_{l=1} \; \omega_n \sim N\left(0, \Sigma_{\omega} diagonal\right) \; (23)$$

$$S_{1n}^{*} = X_n \lambda_1 + \omega_{1n} \; \omega_n \sim N\left(0, \Sigma_{\omega} diagonal\right)$$

$$X_n = X_1, X_2, X_3$$

X_1= Sex

X_2= Age

X_3= Education

Where:

S_{1n}^{*} =Latent personality of individual n

X_n =Observed variables, including socioeconomic characteristics of individual n and attributes of alternative i and individual n.

λ_1= Unknown personality parameter

ω_{1n} = Error term in the personality equation the resulting utility equation is given by equation (24).

$$U_n = X_n \beta_1 + S_{1n}^{*} \beta_2 + \varepsilon_n \qquad \varepsilon_n \sim N\left(0,1\right) \qquad (24)$$

$$X_n = X_4, X_5, X_6, X_7, X_8$$

With:

X_4= Cost

X_5= Travel time

X_6= Cost by Income

X_7= Cost by Sex

X_8= Walking time

U_n = Utility vector

Where

S_{1n}^{*} = Latent personality of the individual n

β_1 =Unknown parameter of the utility that has to do with the exogenous variables

β_2 = Unknown parameter of the utility that has to do with the latent variable of personality

ε_n = Error term in the utility equation

The 16 equations (one per indicator) for measuring the latent variables through the indicators are as follows (Eq.(25)):

$$I_{rn} = S_{1n}^{*}\alpha_r + \upsilon_{rn} \quad r = 1,2,3,...,16$$
$$\upsilon_{rn} \sim N\left(0, \Sigma_{r\upsilon} diagonal\right) \tag{25}$$

Where

I_{rn} = Latent personality variable indicator for r indicators for individual n

α_r = Unknown parameter of the indicator regarding the latent personality variable

υ_{rn} = Error term in the indicator equation

The utility equation for personality is then given by
$$V_{in} = V_{in}(X_n, X_n^{*}; \beta) \quad i \in A(n), \, A(n) \text{ choice set}$$

Where

X_n = Observable variable

X_n^{*} = Latent variable

β = Unknown parameters

The resulting utility functions are shown in equations (26) and (27).

$$V_{in} = \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8$$
$$+ \beta_2 S_{1n}^{*} \tag{26}$$

$$U_{in} = V_{in} + \varepsilon_{in} \; U_{in} = V_{in}\left(X_n, X_n^{*}; \beta\right) + \varepsilon_{in}$$
$$\varepsilon_n \sim D\left(0, \Sigma_\varepsilon\right)$$

$$U_{in} = \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8$$
$$+ \beta_2 S_{1n}^{*} + \varepsilon_{in} \tag{27}$$

We can obtain the choice probability as equation (28),

$$P\left(Y_{in} = 1 \mid X_n, X_n^{*}; \beta\right) = \frac{e^{V_{in}}}{\displaystyle\sum_{j \in C_n} e^{V_{jn}}} \quad \forall_j \in A(n) \tag{28}$$

With
$$Y_{in} = \begin{cases} 1, & if \; U_{in} = \max_j \{U_{jn}\} \\ 0, & otherwise \end{cases} \tag{29}$$

Where

Y_{in} = Indicator of choice

U_{in} = Utility of alternative i for individual n

4. MODEL APPLICATION

The proposed model was applied to asample population of 218 people, 85% of whom are employed and 15% of whom are professors at the National University of Colombia at Medellin. Of the sample population, 53% are women; 43% men; 59% are over 35 years old; 52% had graduated level education; 48% have college degree, and 23% have incomes over $2,000,000. The modal share of the sample was 24.8% auto;45.9% bus;6% taxi;16.5% motorcycle; 3.2% walkingand 3.6% Metro. The data were collected using revealed preference surveys.

4.1. Basic discrete choice model

This analysis corresponds to the results of a basic discrete choice model that does not include latent variables (estimated with the software BIOGEME); the model contains only alternative attributes and socio-economic characteristics of the elector. To assess the discrete model, auto, bus, taxi, motorcycle,walking and metro were used as alternatives.

4.2. Hybrid model with the personality variable

In this study, sex, age, and education were used as exogenous variables in the structural equations that determine the latent personalitybecause these variables, especially sex and age, influence an individual's personality.

After applying equations (23) and (25) (using AMOS in SPSS software), the results shown in Table 1 were obtained

Table1. Parameters Personality traits

Factor		Personality traits	Parameter	t-test value
1	A	Reserved-Open	-0.944	(-2.018)
2	B	Concrete Thinking-Abstract Thinking	-1.422	(-2.618)
3	C	Emotional Instability-Emotional Stability	-3.993	(-9.224)

Factor		Personality traits	Parameter	t-test value
4	E	Submissive-Dominant	0.171	(0.375)
5	F	Prudent-Impulsive	-1.905	(-3.907)
6	G	Carefree-Scrupulous	-1.445	(-2.861)
7	H	Shy-Spontaneous	-2.269	(-5.765)
8	I	Rational-Emotional	-0.208	(-0.435)
9	L	Trusting-Suspicious	2.129	(4.611)
10	M	Practical-Dreamer	-0.105	(-0.211)
11	N	Single-Sly	-0.198	(-0.374)
12	O	Safe-Unsafe	3.463	(8.023)
13	Q1	Traditionalist-Innovative	0.696	(1.448)
14	Q2	Dependence on the group-Self-sufficient	0.360	(0.773)
15	Q3	Uninhibited-Controlled	-2.148	(-4.486)
16	Q4	Calm-Stressed	3.820	(9.661)

Overall, the sample population is a reserved community that poses concrete thinking and has a significant level of emotional instability. According to the parameters of the personality traits and their respective t-test, the community is shy, unsafe, and very stressed (see Table1).

Table 2 shows the values of the parameter estimates (β) and their respective t-test, log likelihood and ρ^2.

Table 2. Model Comparison

Variables	Basic MNL	Personality MNL
ASC1 auto	fixed	fixed
ASC2 bus	fixed	fixed
ASC3 taxi	-2.51	-1.79
	(-4.81)	(-3.23)
ASC4 motorcycle	fixed	fixed
ASC5 walk	-0.53	-0.595
	(-0.79)	(-0.85)
ASC6 metro	-2.44	-3.04
	(-2.5)	(-3.16)
(β_4)cost	-0.00016	-0.000166
	(-2.19)	(-2.03)
(β_8)walking time	-0.171	-0.0834
	(-4.36)	(-2.11)

Variables	Basic MNL	Personality MNL
(β_6) cost by income	0.000117	0.000102
	1.49	(+1.13)
(β_7) cost by sex	-0.00014	-0.000216
	(-1.59)	(-1.69)
(β_5) travel time	-0.0962	-0.0706
	(-5.5)	(-4.37)
(β_{per}) personality	-----	76.9
	-----	(+3.1)
l(β)	-110.439	-92.677
ρ^2	0.657	0.712

This model provides an excellent fit to the data, as all of the parameters have the correct signs (are conceptually valid) and are statistically appropriate. The most important variable is the latent personality variable, which is highly significant. Accordingly, for the sample population studied, personality is an important variable to take into account when modeling an individual's mode of transport.

Comparing the two models, theMNL Basic and the hybrid model with a personality variable (see Table 2), we found that the model considering the personality variable has better fit l(β); (-92.677 > -110.439)and a higher ρ^2 (0.712 > 0.657). Furthermore, the latent personality variable is significant at the 95% confidence level (t-test is 3.1 > 1.96); thus, the hybrid model is superior to the model that does not consider latent variables.

5. CONCLUSIONS

By integrating the latent personality variable into a discrete choice model, this study has presented a model that more accurately explains the decision process and thus hasa smallererror term than that of the basic model.

Econometric and psychometric discrete choice models have a better fit and are more explanatory than those that do not consider these factors. Thus, the models should be estimated using these two disciplines in a synergistic fashion.

Hybrid discrete choice models that consider psychometric tests for the construction of latent

variables provide more accurate results than models that do not do consider these tests.

The impacts of including the latent personality variable in a hybrid discrete choice model or using psychometric tests for the construction of latent variable indicators, which are then introduced into the hybrid model by sequential estimation, are unknown. This study has shown that a significant amount of research must be performedto incorporatethe latent personality variable into the hybrid model by sequential estimation. More research should be conducted in this field to improve this type of model and thus advance the discrete choice model theory including personality variables.

REFERENCES

[1] Ben-Akiva, M.E., Walker, J.L., Bernardino, A.T., Gopinath, D.A., Morikawa, T. y Polydoropoulou, A., Integration of choiceand latent variable models. En H.S. Mahmassani (ed.), In Perpetual Motion: Travel Behaviour Research Opportunities and Challenges, Pergamon, Amsterdam. pp. 431-470, 2002.

[2] Raveau, S., Alvarez, R., Yáñez, M.F., Bolduc, D. and Ortúzar J.de D., Sequential and simultaneous estimation of hybrid discrete choice models: some new findings. 89th Annual Meeting of the Transportation Research Board. Washington, USA. 2010.

[3] Ashok, K., Dillon, W. y Yuan, S., Extending discrete choice models to incorporate attitudinal and other latent variables. Journal of Marketing Research,39, pp. 31-46, 2002.

[4] McFadden, D. The choice theory approach to marketing research. Marketing Science 5, 275-97, 1986.

[5] Morikawa,T., Sasaki, K., Discrete choice models with latent variables using subjective data. In: Ortúzar, J. de D., Hensher, D.A., Jara-Diaz, S. (Eds.), Travel Behavior. Research: Updating the State of Play. Pergamon, Oxford, pp. 435–455, 1998.

[6] Morikawa, T., Ben-Akiva, M., McFadden, D., Discrete Choice Models Incorporating Revealed Preferences and Psychometric Data. Econometric Models in Marketing Advances in Econometrics: A Research Annual, vol. 16. Elsevier Science Ltd. 2002.

[7] Pendleton, L.H., Shonkwiler, J.S., Valuing bundled attributes: a latent characteristics approach. Land Economics 77 (1), pp. 118–129, 2001.

[8] Vredin Johansson, M., Heldt, T. y Johansson, P., Latent variables in a travel mode choice model: Attitudinal and behavioural indicator variables. Working Paper, Department of Economics, Uppsala University. 2005.

[9] Yánez, M.F., Raveau, S., Ortúzar, J. de D., Inclusion of latent variables in Mixed logit models: Modeling and forecasting. Transportation Research Part A 44, pp. 744-753, 2010.

[10] Walker, J., Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures and Latent Variables. PhD. Thesis, MIT, 2001.

[11] Domencich, T. and McFadden, D., Urban Travel Demand: A Behavioral Analysis, Ed. North-Holland, New York. 1975.

[12] Williams, H. y Ortúzar, J., Behavioral theories of dispersion and the misspecification of travel demand models, Transportation Research, 16B (3), pp. 167-219, 1982.

[13] Cattell, R. y Eysenck, H., The Biological Basis of Personality, 1967.

[14] Cattell, R. B., Eber, H. W. y Tatsuoka, M. M., Handbook for the sixteen personality factor questionnaire (16PF), Institute for Personality and Ability Testing, Champaing, IL, 1970.

[15] Córdoba, J.E., González, C., Posada, J.J., The selection of a mode of urban transportation: Integrating psychological variables to discrete choice models. Dyna, 78, Nro. 170, pp. 98-105, 2011.