Estimating discrete choice experiments : theoretical fundamentals

Benoit Chèze, Charles Collet, Anthony Paris

ÉNERGIE ET INNOVATION : RECHERCHE / ÉTUDE & SYNTHÈSE

LES CAHIERS DE L'ÉCONOMIE

IFP SCHOOL - IFPEN

N° 142

AVRIL • 2021

RECHERCHE

ESTIMATING DISCRETE CHOICE EXPERIMENTS: THEORETICAL FUNDAMENTALS

This working paper overviews theoretical foundations and estimators derived from econometric models used to analyze stated choices proposed in Discrete Choice Experiment (DCE) surveys.

Benoît Chèze
Charles Collet
Anthony Paris

Pour toute information complémentaire
For any additional information

Victor Court
IFP School
Centre Economie et Management de l'Energie
Energy Economics and Management Center
victor.court@ifpen.fr
Tél +33 1 47 52 73 17

Estimating discrete choice experiments: theoretical fundamentals

Benoît Chèze *[a,b], Charles Collet[c], and Anthony Paris[d,b]

[a]IFP Energies nouvelles, 1 et 4 avenue de Bois-Préau, 92852 Rueil-Malmaison, France.
[b]EconomiX-CNRS, University Paris Nanterre, France.
[c]CIRED-CNRS, 45 bis avenue de La Belle Gabrielle, 94736 Nogent sur Marne Cedex, France.
[d]LEO, University of Orléans, Rue de Blois, 45067 Orléans, France.

April 12^{th}, 2021

Abstract

This working paper overviews theoretical foundations and estimators derived from econometric models used to analyze stated choices proposed in Discrete Choice Experiment (DCE) surveys. Discrete Choice Modelling is adapted to the case where the variable to be explained is a qualitative variable which cannot be ranked in relation to each other. A situation which occurs in many cases in everyday life as people often have to choose only one alternative among a proposed set of different ones in many fields (early in the morning, just think about how you pick clothes for instance). DCE is a Stated Preference method in which preferences are elicited through repeated fictional choices, proposed to a sample of respondents. Compared to Revealed Preference methods, DCEs allow for an *ex ante* evaluation of public policies that do not yet exists.

Keywords: Revealed preference theory; Stated Preference / Stated Choice methods; Discrete Choice Modelling; Discrete Choice Experiment.

*Corresponding author: benoit.cheze@ifpen.fr .

1 Introduction

Discrete Choice Experiments (DCEs) have gained popularity among environmental economists in the last ten years, as shown by its considerable recent use in many fields. DCE is a Stated Preference (SP) / Stated Choice (SC) method in which preferences are elicited through repeated fictional choices made by respondents (Hoyos, 2010; Louviere et al., 2000). In comparison to Revealed Preference (RP) approaches, it presents the advantage of (i) allowing an *ex ante* evaluation of public policy scenarios and (ii) capturing the non-use value.[1]

Compared to other SP methods, such as the contingent valuation approaches, DCEs take into account several characteristics, called *attributes* in the literature, of the considered issue. It is particularly useful in examining the trade-offs that occur in decision-making. As DCE framework aims at considering several attributes of an issue, it delivers more detailed information than other SP methods. Especially, it makes it possible to estimate the marginal rates of substitution (MRS) between different attributes. When one of these attributes corresponds to a monetary contribution, these MRS can be interpreted as the willingness To pay (WTP), or To Accept (WTA), for changes in the attributes' levels. It thus procures a WTP/WTA for each of these attributes, rather than a global WTP/WTA, as a contingent valuation would do.

This multi-dimensional framework is especially useful for the valuation of hypothetical public policies as it not only allows to values respondent' preferences globally but also every attribute of the problem. It thus can help in setting priorities for public action by identifying the most important factors for a given population. It exhibits the trade-offs at stake for people involved. It can also identify the specific reason a public program is either successful or unsuccessful.

DCE approaches rely on the economic theories of consumer choice and non-market valuation. In a DCE survey, respondents must choose from several options defined by their attributes (*i.e.*, fundamental characteristics of the respondents' situation). Respondents receive a series of choice sets, each comprising several options. Often, three options are presented: nothing changes (*i.e.*, the status quo) and two alternative options. The use of an opt-out option (status quo) is known to improve realism in choices (Adamowicz and Boxall, 2001; Kontoleon and Yabe, 2003). Respondents then choose their favorite option. Each option has different levels of the attributes. One of these attributes usually represents the monetary contribution of the respondents. Other attributes can include environmental or social implications of the issue under consideration. See Louviere et al. (2000) for a detailed description of the method.

Typically, a DCE survey begins with some information about the issue under investigation (e.g. advantages and disadvantages). Various successive choices (say ten for instance) are then proposed between two scenarios and – in most cases – a status quo option presented in the form of a set of (ten) choice cards. Fictional choices are thus made by respondents among several options. Respondents finish survey by responding to some socio-demographic, economic and/or psychological (attitudinal) questions allowing researchers to analyze the impact of these characteristics on their structure of preferences. This quantitative analysis is conducted using specific econometric estimators relying on Discrete Choice Modelling (DCM) theory.

These estimators are presented in this working paper. Section 2 presents theoretical foundations of the DCE approach, both in preference and WTP spaces, and specify the indirect utility functions

[1]The non-use value of non-market goods is the existence value or bequest value. It is useful when valuing, for instance, issues linked to biodiversity and climate change.

usually estimated in the litterature whether or not alternatives are labeled. Section 3 presents estimators derived from econometric models used to analyze stated choices proposed in DCE surveys to reveal respondents's preferences. Section 4 concludes.

2 Theoretical foundations of the Discrete Choice Experiment approach

The DCE modeling framework relies on Lancaster's characteristics of value theory (Lancaster, 1966) as well as random utility theory (McFadden, 1974). This Section presents both the use of the Random Utility Model (RUM) in DCEs and how indirect utility functions are usually specified.

2.1 Random Utility Model in Discrete Choice Experiment

Since a good may be described by a set of characteristics (Lancaster, 1966), its value becomes the sum of the values of all its characteristics. According to the DCE approach, an alternative $i \in [\![1; I]\!]$ can then be described by a set of K observable characteristics, called attributes, as denoted by $X_i = (x_{i,1}, ..., x_{i,k}, ..., x_{i,K})'$. An individual $n \in [\![1; N]\!]$ is described by A economic and attitudinal characteristics, called socioeconomic variables, denoted $Z_n = (z_{n,1}, ..., z_{n,a}, ..., z_{n,A})'$. The (indirect) utility $V_{n,i}$ is thus given by:

$$V_{n,i} = V(X_i, Z_n) \quad \forall\, n \in [\![1; N]\!]\ ;\ i \in [\![1; I]\!] \tag{1}$$

McFadden (1974) considered that individuals make choices according to a deterministic component based on both their measured characteristics Z and attribute alternatives X, along with some degree of randomness ϵ. The random utility $U_{n,i}$ of an alternative $i \in [\![1; I]\!]$ for respondent $n \in [\![1; N]\!]$ is therefore composed of a deterministic part, i.e. the (indirect) utility $V_{n,i} = V(X_i, Z_n)$, and the stochastic element, $\epsilon_{n,i}$, thereby capturing the unsystematic and unobserved random element of respondent n's choice (Louviere et al., 2000).

$$U_{n,i} = V(X_i, Z_n) + \epsilon_{n,i} \tag{2}$$

Assuming the rationality of individuals, respondents choose the alternative i from a finite set of I alternatives, if its utility, $U_{n,i}$, is greater than the utility derived from any other alternatives j, $U_{n,j}$:

$$U_{n,i} > U_{n,j} \;\Rightarrow\; V_{n,i} + \epsilon_{n,i} > V_{n,j} + \epsilon_{n,j} \quad \forall\, j \neq i\ ;\ (i, j) \in [\![1; I]\!]^2 \tag{3}$$

The probability that a respondent chooses alternative i is the same as the probability that the utility of alternative i is greater than the utility of any other alternative in the choice set (Adamowicz et al., 1998). Following Train (2009), this probability is

$$P_{n,i} = P\left\{U_{n,i} > U_{n,j} \;\; \forall\, j \neq i\ ;\ (i, j) \in [\![1; I]\!]^2\right\} \tag{4}$$

$$\Leftrightarrow P_{n,i} = P\left\{V_{n,i} + \epsilon_{n,i} > V_{n,j} + \epsilon_{n,j} \quad \forall\, j \neq i\ ;\ (i, j) \in [\![1; I]\!]^2\right\} \tag{5}$$

$$\Leftrightarrow P_{n,i} = P\left\{\epsilon_{n,j} - \epsilon_{n,i} < V_{n,i} - V_{n,j} \quad \forall\, j \neq i\ ;\ (i, j) \in [\![1; I]\!]^2\right\} \tag{6}$$

2.2 Specification of the indirect utility functions

According to Equation (2), the random utility $U_{n,i}$ is composed of a deterministic component, $V_{n,i} = V(X_i, Z_n)$, and a stochastic element, $\epsilon_{n,i}$. Before estimating an econometric model, the deterministic part of the utility function $V_{n,i} = V(X_i, Z_n)$ must be specified. The linear specification is often chosen in the literature for the sake of simplicity. We have thus introduced the column vector of parameters $\beta_n = (\beta_{1,n,i}, ...\beta_{K,n,i})$, which are the coefficients quantifying the (linear) influence of the K attributes on utility that individual n associates with the I available alternatives.

An Alternative Specific Constant (ASC_i) term is usually introduced to capture the effect of unobserved influences (i.e. omitted variables) on the utility function for the *i-th* alternative. It is a dummy variable assuming a value 1 for the *i-th* alternative, and 0 otherwise. In its most general form, the model is specified such that the utility of an alternative is expressed as follows:

$$U_{n,i} = (\eta_i + \phi_{n,i} Z_n) ASC_i + \beta_{n,i} X_i + (X_i)' \Gamma_i Z_n + \epsilon_{n,i} \tag{7}$$

where $\phi_{n,i} = (\phi_{n,i,1}, ..., \phi_{n,i,A})$ are the coefficients representing the direct influence of the A socioeconomic and psychological variables on the utility of the N individuals for alternative I. The matrix Γ_i of size $(K \times A)$, is composed of coefficients $\gamma_{i,k,a}$; it serves to capture the effect of individuals' characteristics $z_{n,a}$ on attribute $x_{i,k}$.

2.2.1 Non labeled choice experiment

When alternatives are unlabeled, the choice options given are thus symmetrical, differing only in the attributes assigned X_i. Even if a reference choice bias were to exists, it could be represented by a simple Alternative Specific Constant. Therefore, the previous coefficients $\beta_{n,i}$, $\phi_{n,i}$ and Γ_i do not depend on alternative i; they can be simplified into β_n, ϕ_n and Γ. Let's observe that for each individual, $\phi_n Z_n$ remains constant across all alternatives i and therefor does not influence the model; hence, this item can be removed, yielding:

$$U_{n,i} = \beta_n X_i + (X_i)' \Gamma Z_n + \epsilon_{n,i} \tag{8}$$

This underlines that for a non-labeled choice model, one of the main solution in order to include the influence of a socio-economic variable $z_{n,a}$ is to introduce it through an interaction term with the alternative attributes $x_{k,i}$.

2.3 WTP-space approach

In their seminal paper, Train and Weeks (2005) found that the WTP distributions they derived from preference models had an "unreasonably" large variance in comparison with WTP distributions derived from utility models in the WTP-space. This result has been confirmed in subsequent studies (Mabit et al., 2006; Scarpa et al., 2008; Rose and Masiero, 2010). According to the WTP-space approach, marginal WTP estimates are directly obtained, unlike with the preference space approach, where a ratio is to be computed between the non-cost attribute and the cost attribute.

Welfare measures can be determined in the form of marginal WTP/WTA, *i.e.* for a unit change of a given attribute, by computing the MRS between the considered attribute and the cost attribute

(Louviere et al., 2000). Since utilities are modeled as linear functions of the attributes, the *MRS* between two attributes is the ratio between the corresponding coefficients.[2] Thus for continuous attributes[3], the WTP for an attribute $k \in [\![1; K-1]\!]$ equals the ratio of the attribute's coefficient to the cost coefficient, $\beta_{cost,n}$:[4]

$$WTP_{k,n} = \frac{dx_{cost,i}}{dx_{k,i}} = \frac{dU_n/dx_{k,i}}{dU_n/dx_{cost,i}} = \frac{\partial V_n/\partial x_{k,i}}{\partial V_n/\partial x_{cost,i}} = \frac{\beta_{k,n}}{\beta_{cost,n}} \tag{9}$$

where $x_{cost,i}$ is the cost attribute, *i.e.* the K-th attributes, $x_{K,i}$, in X_i.

Equation (8) corresponds to the preference-space model. Combined with Equation (9), utility (without considering any socio-demographic characteristic effects) can be rewritten in the WTP-space as follows:

$$U_{n,i} = \beta_{cost,n} x_{cost,i} + \sum_{k=1}^{K-1} WTP_{k,n} x_{k,i} + \epsilon_{n,i} \tag{10}$$

When adding the effects of individuals' characteristics, it is necessary for all $\gamma_{k,a}$ (which measure the effects of individuals' characteristics on attribute preferences) to be divided by the cost attribute's coefficient $\beta_{cost,n}$ in order to generate their estimation in the WTP-space.[5] The newly formed $\gamma_{k,a}$ in the WTP-space therefore measure the effect of the individuals' characteristic $z_{n,a}$ on the WTP for attribute k. It is possible to consider these effects as being directly included in individuals' WTP so as to facilitate model interpretation.

The WTP is now defined as follows:

$$WTP_{k,n} = WTP_{base,k,n} + \sum_{a=1}^{A} \gamma_{k,a} z_{n,a} \quad \text{with} \quad WTP_{base,k,n} = \frac{\beta_{k,n}}{\beta_{cost,n}} \tag{11}$$

Equation (10) becomes:

$$U_{n,i} = \beta_{cost,n} x_{cost,i} + \sum_{k=1}^{K-1} \left(WTP_{base,k,n} + \sum_{a=1}^{A} \gamma_{k,a} z_{n,a} \right) x_{k,i} + \epsilon_{n,i} \tag{12}$$

3 Econometric models

Different econometric models can be used to analyze discrete choice data. Indeed, the heterogeneity of preferences can be modeled to different degrees, depending on (i) assumptions made on the distribution of the column vector of parameters $\beta_n = (\beta_{1,n,i}, ...\beta_{K,n,i})$ and (ii) the way socioeconomic variables are introduced. This Section presents Conditional Logit (CL), Random Parameter Logit (RPL), Latent Class (LC) and Integrated Choice and Latent Variable (ICLV) models.

[2] It should be noted that the derivative of the unobserved part of the utility function is supposed to be zero with respect to both attributes.

[3] See Appendix B for attributes modeled as dummy variables and specific cases of both RPL and LC models.

[4] The cost attribute is considered to be the last one, for purpose of consistency with the previous notations.

[5] We have opted to maintain the same variable for the effect of the individuals' characteristics in both the preference and WTP-space in order to simplify notations.

3.1 Conditional Logit models

The CL model is a cornerstone for analyzing discrete choice data and has been widely used in DCE. This model has however several well-known limits. An important drawback is that it assumes homogeneous preferences across respondents, meaning that the probability that an agent n chooses alternative i in a choice set S, is considered fixed across all individuals ($\beta_n = \beta$, $\Gamma_n = \Gamma \ \forall \ n$), while we can expect the preferences to vary. Two other important drawbacks are uncorrelated unobserved components and the hypothesis of the independence of irrelevant alternatives (IIA).

IIA implies that the relative probabilities of two options being chosen are unaffected by the introduction or removal of other alternatives. The (multinomial) logit probability that a respondent n chooses a particular alternatives i is [6]:

$$P_{n,i} = \frac{e^{V_{n,i}}}{\sum_j e^{V_{n,j}}} \tag{13}$$

The hypothesis of irrelevant alternatives implies that the relative probabilities of two options, here i and h, being chosen are unaffected by the introduction or removal of other alternatives. This is an underlying assumption of the CL as shown by Equation (14):

$$\frac{P_{n,i}}{P_{n,h}} = \frac{\frac{e^{V_{n,i}}}{\sum_j e^{V_{n,j}}}}{\frac{e^{V_{n,h}}}{\sum_j e^{V_{n,j}}}} = \frac{e^{V_{n,i}}}{e^{V_{n,h}}} \tag{14}$$

If the IIA property is violated then the CL model does not fit the data. Results will be biased, leading to unrealistic predictions, and hence a discrete choice model that does not require the IIA property should be used. Moreover, it does not uses the information that the same individual make several choices, thus cannot be as accurate as a model taking into account this supplementary information.

3.2 Random Parameter Logit models

Compared to the CL model, the RPL model (Mcfadden and Train, 2000; Train, 2009) relaxes the IIA hypothesis and is able to take the heterogeneity of preferences into account. Indeed, the preferences parameters β are allowed to vary randomly across respondents, in allowing for the fact that different decision-makers may have different preferences: $\beta_n \neq \beta_m \ \ \forall \ n \neq m \ ; \ (n,m) \in [\![1;N]\!]^2$. As such, conditional on the individual-specific parameters and error components, we can define the (multinomial) logit probability[7] that respondent n chooses a specific alternative i for a given β:

$$P_{n,i}|\beta = L_{n,i}(\beta) = \frac{e^{V_{n,i}(\beta)}}{\sum_j e^{V_{n,j}(\beta)}} \tag{15}$$

[6] See Appendix A for more details.

[7] As the error term is assumed to be independent and identically distributed (iid) Type I Extreme Value. Note that Appendix A details calculation to obtain this probability.

Without taking into account the cross-effects of socio-demographic characteristics, the individual specific utility is simply modeled by:

$$U_{n,i} = \beta_n X_i + \epsilon_{n,i} \tag{16}$$

where $\epsilon_{n,i} \sim$ iid extreme value type I, and $\beta_n \sim g(\beta|\Omega)$.

Since β_n is not known, the unconditional choice probability of person n choosing alternative i is the integral of $P_{n,i}|\beta$ over the distribution of β:

$$P_{n,i} = \int L_{n,i}(\beta) f(\beta|\Omega) d\beta \tag{17}$$

$f(\beta|\Omega)$ is the density describing the distribution of preferences across individuals. Ω is a vector of the true parameters of the taste variation, e.g. fixed parameters of the distribution representing the mean and standard deviation of β_n within the population. The true distribution remains unknown, so, in theory, any distribution could be applied (Hensher and Greene, 2003). A common assumption made in the literature is to assume that random parameters are normally distributed, but the one associated with the cost attribute.[8] The latter is usually kept fixed or supposed to follow a log-normal distribution in order to avoid a "wrong" sign (*i.e.* negative) for a share of respondents.

The choice probability in equation (17) cannot be calculated exactly because the integral does not have a closed form solution in general. This integral is approximated through simulations. For a given value of the parameters Ω, a value of β is drawn from its distribution. Using this draw, the logit formula in (15) is calculated. This process is repeated for many draws, and the mean of the resulting $L_{n,i}(\beta)$ is taken as the approximate choice probability yielding equation (18):

$$SP_{n,i} = \frac{1}{R} \sum_{r=1}^{R} L_{n,i}(\beta_r) \tag{18}$$

where R is the number of draws of β, and SP is the simulated probability that an individual n chooses alternative i.

With a RPL model, one can consider that parameters are random but not independent between them. Train (2009) describes how the simulation and the estimation of the parameters can be done when preference parameters are assumed to follow a random normal joint distribution $\beta \sim \mathcal{N}(\mu, \Sigma)$ where μ is the mean vector of parameters and Σ the covariance matrix for each individual.

3.3 Latent Class models

Another way to relax the hypothesis of the IIA and to take account for the heterogeneity in respondents' preferences is to estimate a LC model. In the latter, each respondent is sorted into a number of classes C in which preferences are assumed to be homogeneous with respect to attributes. Preferences are allowed to be heterogeneous between each latent class segment $c \in C$.

[8]Random parameters are generally supposed to be normally distributed in the RPL model because it is the most easily applied distribution allowing for both negative and positive preferences.

Compared to Equation (15), the logit probability that respondent n prefers a specific alternative i over alternatives j is no more defined for a given β but becomes conditional on class c. Indeed, the β's are now assumed to follow a discrete distribution and belong to one class c of C classes. Thus, the conditional probability that respondents who are members of class c choose alternative i is:

$$P_{n,i}|\beta_c = \frac{e^{V_{n,i}(\beta_c)}}{\sum_j e^{V_{n,j}(\beta_c)}} \;\forall\, c \in [\![1, ..., C]\!] \tag{19}$$

where β_c' is the vector of preferences parameters specific to each class c, representing the average importance of each attribute for respondents belonging to c.

The unconditional probability of individual n selecting choice option i can be expressed as:

$$P_{n,i} = \sum_{c=1}^{C} (\Pi_{n,c} P_{n,i}|\beta_c) = \sum_{c=1}^{C} \left(\Pi_{n,c} \frac{e^{\beta_c X_i}}{\sum_j e^{\beta_c X_j}} \right) \tag{20}$$

where $\Pi_{n,c}$ is the probability of membership of respondent n in class c:

$$\Pi_{n,c} = \frac{e^{\phi_c Z_n}}{\sum_h e^{\phi_h Z_n}} \tag{21}$$

where Z_n is the vector of psychometric constructs and socioeconomic characteristics, and ϕ is the vector of parameters associated to Z_n (Boxall and Adamowicz, 2002).

According to Equation (21), the probability of belonging to a class c with specific preferences is probabilistic, and depends on the social, economic and attitudinal characteristics of the respondents. Combining Equation (20) and Equation (21), it comes that the LC model assumes that respondent characteristics affect his choice indirectly through their impact on segment membership. Note that ϕ_c includes $C-1$ class membership parameters with ϕ_C being normalized to zero for identification. All other coefficients ϕ_c are thus interpreted relative to this normalized class.

3.4 Integrated Choice and Latent Variable models

3.4.1 General Mathematical specification

Hybrid models are composed with a Discrete Choice Model (DCM) and a Structural Equation Model (SEM). The DCM is very similar to the previous models, the difference being in the introduction of the Latent Variables into the deterministic utility function V:

$$U_{n,i} = V(X_i, Z_n, LV_n) + \epsilon_{n,i} \tag{22}$$

where $LV_n = (LV_{n,1}, ..., LV_{n,Q})$ is a $(Q \times 1)$ vector containing the Q latent variables for individual n and $\epsilon_n = (\epsilon_{n,1}, ..., \epsilon_{n,I})$ is the $(I \times 1)$ vector of iid error terms, for the I alternatives, following a extreme value distribution type I with 0 mean and a covariance matrix $\Sigma_{\epsilon,n}$.

The structural equation of the SEM gives information on the distribution of the latent variables given the observed socio-economic characteristics:

$$LV_{n,q} = S(Z_n) + v_{n,q} \tag{23}$$

where $v_n = (v_{n,1}, ..., v_{n,Q})$ is the $(Q \times 1)$ vector of iid error terms for the Q latent variables, *i.e.* $v_n \sim \mathcal{N}(0, \Sigma_{v,n})$, with $\Sigma_{v,n}$ the covariance matrix.

The measurement equation of the discrete choice part of the model is the same as the other DCM:

$$Y_{n,i} = \begin{cases} 1 & \text{if} \quad U_{n,i} > U_{n,j} \quad \forall\, j \neq i \;;\; (i,j) \in [\![1; I]\!]^2 \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

The measurement equation of the SEM uses the values of the attitudinal indicators as dependent variables, and explain their values with the help of the latent variables:

$$I^*_{n,l} = M(LV_n) + w_{n,l} \tag{25}$$

where I^*_n is a $(L \times 1)$ vector containing the L attitudinal indicators used to measure the latent variables. $w_n = (w_{n,1}, ..., w_{n,L})$ is the $(L \times 1)$ vector of iid error terms for the L attitudinal indicators, *i.e.* $w_n \sim \mathcal{N}(0, \Sigma_{w,n})$, with $\Sigma_{w,n}$ the covariance matrix.

The attitudinal indicators are almost always measured with ordinal variables based on Likert-scale items. The most simple option is to consider these ordinal variables as continuously normally distributed (Bouscasse, 2018). However, a more accurate way to take into account the ordinal nature of the indicators is to add a threshold model equation to the measurement model of the SEM, which is usually done in the SEM literature (Kamargianni et al., 2015; Giansoldati et al., 2020). Therefore $I^*_{n,l}$ is a latent variable coming from a threshold model equation:

$$I_{n,l} = \begin{cases} 1 \;\text{ if } -\infty < I^*_{n,l} \leq \tau_{1,l} \\ 2 \;\text{ if } \tau_{1,l} < I^*_{n,l} \leq \tau_{2,l} \\ ... \\ D \;\text{ if } \tau_{D-1,l} < I^*_{n,l} \leq \infty \end{cases} \tag{26}$$

$I^*_{n,l}$ is the exact but unobserved dependant variable and D is the total number of categories (D could vary between all attitudinal indicators L, but here was supposed constant to simplify notations). $\tau_{d,l}$ are the threshold for $I^*_{n,l}$ that give the probability of observing the answer $I_{n,l}$. The probability that the answer $I_{n,l} = d$ is equal to

$$P(I_{n,l} = d) = P(\tau_{d-1,l} < I^*_{n,l} \leq \tau_{d,l}) = F_v(\tau_{d,l}) - F_v(\tau_{d-1,l}) \tag{27}$$

3.4.2 Specification for a non labeled choice experiment

To simplify notations, we consider here that V depends only on the attribute alternatives X_i, so that the individual characteristics Z_n are only included in the Structural equation of the SEM. A linear specification is adopted and the RUM equation becomes:

$$U_{n,i} = \beta_{n,i} X_i + \xi_{n,i} LV_n + X_i' \Lambda_{n,i} LV_n + \epsilon_{n,i} \tag{28}$$

The matrix $\Lambda_{n,i}$ of size $(K \times Q)$ is composed of coefficients $\lambda_{n,i,k,q}$, capturing the cross-effect of latent variable $LV_{n,q}$ on attribute $x_{i,k}$.

In the case of non labeled DCE, the previous coefficients do not depend on the alternative i. Therefore they can be simplified as β_n, ξ_n and Λ_n. $\xi_n LV_n$ can be removed because constant between the alternatives i. This leads to the following equation:

$$U_{n,i} = \beta_n X_i + (X_i)' \Lambda_n LV_n + \epsilon_{n,i} \tag{29}$$

Like in ICLV empirical models found in the literature, structural (23) and measurement (25) equations are usually chosen as linear specifications to limit the number of parameters and enable convergence of the algorithm.

$$LV_{n,q} = \gamma_{n,q} Z_n + \upsilon_{n,q} \tag{30}$$

$$I^*_{n,l} = \nu_{n,l} LV_n + w_{n,l} \tag{31}$$

To enable better understanding of psychological mechanisms, attitudinal indicators are assumed to be explained by only one latent variable: the vector $\nu_{n,l}$ is composed of $Q-1$ zeros and only one non-zero coefficient for the corresponding latent variable. This assumption is made in every empirical ICLV studies we consulted.

Maximum likelihood techniques are used to estimate the unknown parameters. Assuming the error components $(\epsilon_n, \upsilon_n, w_n)$ are independent, the joint probability of the observable variables I_n and Y_n, conditional on the exogenous variable X and Z_n is:

$$\begin{aligned} f_{Y,I}(Y_n, I_n | X, Z_n; \beta_n, \Lambda_n, \gamma_n, \nu_n, \Sigma_{\epsilon,n}, \Sigma_{\upsilon,n}, \Sigma_{w,n}) = \\ \int_{LV} P(Y_n | X, Z_n, LV_n; \beta_n, \Lambda_n, \Sigma_{\epsilon,n}) \\ f_{LV}(LV_n | X, Z_n; \gamma_n, \Sigma_{\upsilon,n}) f_I(I_n | LV_n; \nu_n, \Sigma_{w,n}) dLV_n \end{aligned} \tag{32}$$

with f_{LV_n} and f_{I^*} the density functions depending on the error terms γ_n and ν_n. The first term of the integral corresponds to the choice model, the second term corresponds to the measurement equation, and the third term to the structural equation.

4 Conclusion

Models presented in this working paper belong to the Generalized Extreme Value (GEV) models. This large class of models is adapted to the case where the variable to be explained is a qualitative variable whose cannot be ranked in relation to each other. They aim at explaining people's choice where the choice set is exhaustive, made up of finite number of mutually exclusive alternatives described by attributes. Although seemingly restrictive, these conditions are found in many cases in everyday life such as mode of transport and/or car ownership, home location, food choices, medical or voting decisions, etc.

Despite an RP resurgence, most discrete choice applications are currently based on SP/SC data where respondents have to choose between different hypothetical scenarios. In a DCE survey respondents are faced, by construction, with multiple choice situations which allow researcher to obtain more information per respondent. Analyst can control settings, in the manner of experimental sciences, by (i) choosing not only attributes but also their levels and (ii) designing SP surveys that encourages trade-off behaviour. The latter is an area of research in its own right. It aims to obtain as much information as possible with as few choice situations as possible. Indeed, respondents cannot be offered all the possible combinations between the different levels of the different attributes as they would get tired and would eventually not respond with sufficient concentration.

Compared to RP data, where real world choices are analysed, SP methods produce data that can look at hypothetical choice scenarios by proposing to choose alternatives that do not exist yet.

It is a strength as it is one of the few methods that allow for an *ex ante* evaluation of public policies that do not yet exist. It may also be a curse as respondents make fictional choices rather than real ones. This drawback is referred to in the literature as hypothetical bias. This is a critical issue: if surveys are not taken seriously by respondents, it eventually yields to non reliable data. Hypothetical bias arises when choice tasks are difficult to complete or choices are not incentive-compatible. One solution to the latter is to emphasize consequentiality of the choices (referred to as *"introducing consequentiality"*) – and to control for – in the survey, which is a necessary condition for incentive compatibility (see Johnston et al., 2017).

Appendices

Appendix A The multinomial logit model

Assuming $\epsilon_{n,i}$ being iid and following a type I extreme-value distribution, i.e. a standard Gumbel distribution, we thus specify a conditional logit model (or multinomial logit model).

The cumulative distribution function F and the density function f of each $\epsilon_{n,i}$ are given by:

$$F(\epsilon_{n,i}) = e^{-e^{-\epsilon_{n,i}}} \tag{33}$$

$$f(\epsilon_{n,i}) = e^{-\epsilon_{n,i}} e^{-e^{-\epsilon_{n,i}}} \tag{34}$$

Since the unobserved components are independent, we can multiply Eq. (6) to obtain the probability of individual n choosing alternative i, conditional on $\epsilon_{n,i}$:

$$P_{n,i}|\epsilon_{n,i} = \prod_{j \neq i} P\left\{\epsilon_{n,j} < V_{n,i} - V_{n,j} + \epsilon_{n,i}\right\} \tag{35}$$

$$= \prod_{j \neq i} e^{-e^{-(V_{n,i} - V_{n,j} + \epsilon_{n,i})}} \tag{36}$$

The non conditional probability for an agent n to choose the alternative i is therefore the integration of $P_{n,i}|\epsilon_{n,i}$ over the distribution of $\epsilon_{n,i}$:

$$P_{n,i} = \int \left(\prod_{j \neq i} e^{-e^{-(V_{n,i} - V_{n,j} + \epsilon_{n,i})}} \right) e^{-\epsilon_{n,i}} e^{-e^{-\epsilon_{n,i}}} d\epsilon_{n,i} \tag{37}$$

By replacing $\epsilon_{n,j}$ with s, equation (37) becomes:

$$P_{n,i} = \int_{s=-\infty}^{+\infty} \left(\prod_{j \neq i} e^{-e^{-(V_{n,i} - V_{n,j} + s)}} \right) e^{-s} e^{-e^{-s}} ds \tag{38}$$

As $V_{n,i} - V_{n,i} = 0$, we have:

$$P_{n,i} = \int_{s=-\infty}^{+\infty} \left(\prod_{j \neq i} e^{-e^{-(V_{n,i} - V_{n,j} + s)}} \right) e^{-s} e^{-e^{-(V_{n,i} - V_{n,i} + s)}} ds \tag{39}$$

and the last term can be introduced into the product,

$$P_{n,i} = \int_{s=-\infty}^{+\infty} \left(\prod_{j} e^{-e^{-(V_{n,i} - V_{n,j} + s)}} \right) e^{-s} ds \tag{40}$$

By removing the first exponential from the product, we obtain:

$$P_{n,i} = \int_{s=-\infty}^{+\infty} exp\left(-\sum_{j} e^{-(V_{n,i} - V_{n,j} + s)} \right) e^{-s} ds \tag{41}$$

$$P_{n,i} = \int_{s=-\infty}^{+\infty} exp\Big(- e^{-s} \sum_j e^{-(V_{n,i}-V_{n,j})}\Big) e^{-s} ds \tag{42}$$

We now define $t = e^{-s}$. The expression $-e^{-s}ds$ therefore gives dt and note that t approaches zero (resp. positive infinity) if s tends to infinity (resp. negative infinity) as:

$$P_{n,i} = \int_{t=+\infty}^{0} exp\Big(- t \sum_j e^{-(V_{n,i}-V_{n,j})}\Big)(-dt) \tag{43}$$

that is to say:

$$P_{n,i} = \int_{t=0}^{+\infty} exp\Big(- t \sum_j e^{-(V_{n,i}-V_{n,j})}\Big) dt \tag{44}$$

This expression is now easy to integrate and allows us to obtain expression in equation (15).

$$P_{n,i} = \frac{exp(- t \sum_j e^{-(V_{n,i}-V_{n,j})})}{- \sum_j e^{-(V_{n,i}-V_{n,j})}}\Bigg]_0^{+\infty} \tag{45}$$

$$P_{n,i} = 0 - \frac{1}{- \sum_j e^{-(V_{n,i}-V_{n,j})}} \tag{46}$$

$$P_{n,i} = \frac{1}{\sum_j e^{-(V_{n,i}-V_{n,j})}} = \frac{e^{V_{n,i}}}{\sum_j e^{V_{n,j}}} \tag{47}$$

where $P_{n,i}$, the (multionomial) logit probability, only depends on observable components.

The CL model is estimated using maximum likelihood procedures. The probability that a respondent n chooses a particular alternative is $\prod_i (P_{n,i})^{y_{n,i}}$ with $y_{n,i} = 1$ if the alternative i is chosen and zero otherwise. Assuming the independence in choices of each respondent, the likelihood and log-likelihood functions are given by:

$$L(\beta) = \prod_{n=1}^{N} \prod_i (P_{n,i})^{y_{n,i}} \tag{48}$$

$$LL(\beta) = \sum_{n=1}^{N} \sum_i y_{n,i} ln(P_{n,i}) \tag{49}$$

with:

$$P_{n,i} = \frac{e^{V_{n,i}}}{\sum_j e^{V_{n,j}}} \tag{50}$$

Appendix B Computing WTP/WTA: some complements

For attributes modeled as dummy variables, the $WTP^l_{k,n}$ associated with attribute k and category l is

$$WTP^l_{k,n} = \frac{\beta^l_{k,n}}{\beta_{cost,n}} \tag{51}$$

The variation of utility associated with a change of the attribute k from the status quo to category l is a measure of the willingness to pay to do so.

Since the RPL model assumes attributes' coefficients are randomly distributed, we then have the following result in this case:

$$E[WTP_k, n] = \frac{E[\beta_{k,n}]}{E[\beta_{cost,n}]} \tag{52}$$

$$E[WTP^l_{k,n}] = \frac{E[\beta^l_{k,n}]}{E[\beta_{cost,n}]} \tag{53}$$

Hole (2007) explains *"since the logit model is typically estimated using maximum likelihood, which implies that the coefficients in the model are asymptotically normally distributed, it is reasonable to assume that WTP is given by the ratio of two normally distributed variables when the model is estimated using a large sample. The distribution of the ratio of two normally distributed variables has been derived by Fieller (1932) and Hinkley (1969), who show that the distribution is approximately normal when the coefficient of variation of the denominator variate (in this case* $\beta_{cost,n}$ *), is negligible."*

For a LC model, the WTP for the individual n for a variation of the attribute k can be computed per class as:

$$WTP^c_k = \Pi^*_{n,c} \cdot \frac{\beta^c_k}{\beta^c_{cost}} \tag{54}$$

where c are the latent classes, β^c_k the parameter associated to attribute k for each latent class c, β^c_{cost} the parameter associated to the cost attributes for each latent class c, and $\Pi^*_{n,c}$ the posterior estimate of the individual-specific class probability of membership of respondents n in class c.

For each models, the estimated standard deviations and confidence intervals around the mean of the WTP estimates may be obtained using different methods, depending whether or not the previous property is verified. The four main methods are *i)* the delta, *ii)* the Fieller (Fieller, 1954), *iii)* the Krinsky and Robb (Krinsky and Robb, 1986, 1990) and *iv)* the bootstrap methods. In calculating a WTP, it is important that both parameters used in the calculation be statistically significant, otherwise no meaningful WTP measures can be established.

References

Adamowicz, V. and Boxall, P. (2001). Future directions of stated choice methods for environment valuation. *Choice experiments: A new approach to environmental valuation, London*, pages 1–6.

Adamowicz, W., Boxall, P. C., Williams, M., and Louviere, J. (1998). Stated preference approaches for measuring passive use values: Choice experiments and contingent valuation. *American Journal of Agricultural Economics*, 80(1).

Bouscasse, H. (2018). Integrated choice and latent variable models: A literature review on mode choice. Working Papers GAEL n° 07/2018.

Boxall, P. C. and Adamowicz, W. L. (2002). Understanding Heterogeneous Preferences in Random Utility Models: A Latent Class Approach. *Environmental and Resource Economics*, 23(4):421–446.

Fieller, E. C. (1932). The distribution of the index in a normal bivariate population. *Biometrika*, 24.

Fieller, E. C. (1954). Some problems in interval estimation. *Journal of the Royal Statistical Society*, Series B (16).

Giansoldati, M., Rotaris, L., Scorrano, M., and Danielis, R. (2020). Does electric car knowledge influence car choice? Evidence from a hybrid choice model. *Research in Transportation Economics*, 80:100826.

Hensher, D. A. and Greene, W. H. (2003). The Mixed Logit model: The state of practice. *Transportation*, 30:44.

Hinkley, D. V. (1969). On the ratio of two correlated normal variables. *Biometrika*, 56.

Hole, A. R. (2007). A comparison of approaches to estimating confidence intervals for willingness to pay measures. *Health Economics*, 16.

Hoyos, D. (2010). The state of the art of environmental valuation with discrete choice experiments. *Ecological Economics*, 69(8):1595–1603.

Johnston, R., Boyle, K., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T., Hanemann, M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., and Vossler, C. (2017). Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, 4(2):319–405.

Kamargianni, M., Dubey, S., Polydoropoulou, A., and Bhat, C. (2015). Investigating the subjective and objective factors influencing teenagers' school travel mode choice - an integrated choice and latent variable model. *Transportation Research Part A: Policy and Practice*, 78:473–488.

Kontoleon, A. and Yabe, M. (2003). Assessing the impacts of alternative 'opt-out'formats in choice experiment studies: consumer preferences for genetically modified content and production information in food. *Journal of Agricultural policy and Resources*, 5(1):1–43.

Krinsky, I. and Robb, A. L. (1986). On approximating the statistical properties of elasticities. *The Review of Economics and Statistics*, 68.

Krinsky, I. and Robb, A. L. (1990). On approximating the statistical properties of elasticities: a correction. *Review of Economics and Statistics*, 72.

Lancaster, K. J. (1966). A new approach to consumer theory. *The Journal of Political Economy*, 74(2):132–157.

Louviere, J. J., Hensher, D. A., and Swait, J. D. (2000). *Stated Choice Methods, Analysis and Applications*. Cambridge University Press.

Mabit, S., Caussade, S., and Hess, S. (2006). Representation of taste heterogeneity in willingness to pay indicators using parameterization in willingness to pay space. *Proceedings ETC*.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In *Frontiers of econometrics*. Academic Press, New York.

Mcfadden, D. and Train, K. (2000). Mixed MNL models for discrete response. *J. Appl. Econ.*, page 24.

Rose, J. M. and Masiero, L. (2010). A comparison of the impacts of aspects of prospect theory on wtp/wtaestimated in preference and wtp/wta space. *European Journal of Transport and Infrastructure Research*, 10(4).

Scarpa, R., Thiene, M., and Train, K. (2008). Utility in willingness to pay space: a tool to address confounding random scale effects in destination choice to the alps. *American Journal of Agricultural Economics*, 90(4):994–1010.

Train, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press, 2nd ed. edition.

Train, K. and Weeks, M. (2005). Discrete choice models in preference space and willingness-to-pay space. In *Applications of simulation methods in environmental and resource economics*, pages 1–16. Springer.

Retrouvez toute la collection
https://www.ifpenergiesnouvelles.fr/article/les-cahiers-leconomie

IFP SCHOOL
228 - 232 avenue Napoléon Bonaparte
92852 Rueil-Malmaison
www.ifpschool.com

IFP Energies nouvelles
1-4 avenue de Bois-Préau
92852 Rueil-Malmaison
www.ifpenergiesnouvelles.fr