Annex B: The Mixed Logit Model (ML)

In accordance with random utility theory (McFadden, 1973, McFadden, 1974), we assume that the utility of individual *n* of choosing alternative *j* in choice situation *t* can be represented as:

$$U_{njt} = \boldsymbol{\beta'}_n . X_{njt} + \varepsilon_{njt} \tag{1}$$

where X_{njt} is a vector of *K* observed attributes related to the alternative *j* of the choice situation *t*; $\boldsymbol{\beta'}_n$ is a vector of preference parameters which explain choices; ε_{njt} is the unobserved error term.

The preference parameters β_k are distributed in the population according to continuous random distributions $f(\beta)$ to be chosen by the analyst (Train, 2009). For the full vector of *K* random coefficients in the model, we may write the full set of random parameters as:

$$\boldsymbol{\beta}_n = \boldsymbol{\beta} + \boldsymbol{\Gamma} . \boldsymbol{v}_n \tag{2}$$

where $\boldsymbol{\Gamma}$ is a lower-triangular matrix that takes care of the possible correlations among coefficients[1], and $\boldsymbol{v}_n$ correspond to the individual specific heterogeneity. If at least one of the elements below the diagonal of $\boldsymbol{\Gamma}$ shows statistical significance, this is supportive of dependence across tastes (Scarpa and Del Giudice, 2004). As individuals were confronted to several choice situations (panel data), we have to consider a sequence of *S* observed choices (s_1, s_2, …, s_S) by the same individual. The probability of observing a sequence *s* conditional on β_n is:

$$L_{ns}(\beta_n) = \prod_{s=s_1}^{s_S} \left[\frac{exp(\beta_n{'}.X_{njs})}{\sum_{q=1}^{J} exp(\beta_n{'}.X_{nqs})} \right] \tag{3}$$

However, the researcher does not know β_n and therefore cannot condition the probability of choosing one alternative on β. The unconditional choice probability is therefore the integral of $L_{ns}(\beta_n)$ over all possible values of β_n is $P_{ns} = \int L_{ns}(\beta). f(\beta). d\beta$. Because the integral in P_{ns} equation does not have a closed form solution, the parameters of the model are estimated by simulated maximum likelihood estimation techniques (Train, 2009).

[1] The matrix **Γ** corresponds to the Cholesky decomposition of the covariance matrix: Γ'.Γ=COV