

Estimating the MDCEV model with Biogeme

Michel Bierlaire Mengyi Wang

January 16, 2024

DRAFT Report TRANSP-OR 24xxxx
Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne
transp-or.epfl.ch

Abstract

The abstract...

1 The MDCEV model

The multiple discrete-continuous extreme value model (MDCEV) is a choice model where the choice of multiple alternatives can occur simultaneously. It has been introduced by Bhat (2005), building on the Karush-Kuhn-Tucker multiple discrete-continuous economic model proposed by Wales and Woodland (1983). In this document, we introduce a generalization of the model, where the derivation is performed for a generic utility function. This is motivated by the need to obtain an implementation that is easily extendible to new models in the future.

Consider an individual, denoted as n , who is presented with a distinct set of items, represented as \mathcal{C}_n , containing J_n items in total. Given a total budget of E_n , this individual decides on purchasing a quantity y_{in} of each item. This decision must verify the following budget constraint:

$$\sum_{i \in \mathcal{C}_n} e_{in} = \sum_{i \in \mathcal{C}_n} p_{in} y_{in} = E_n,$$

where:

- $p_{in} > 0$ is the price per unit of item i for the individual n .
- $e_{in} = p_{in} y_{in}$ represents the total expenditure by individual n on item i .

We assume that the entire budget is spent, and that at least one item is consumed.

Each item i is associated with a utility $U_{in}(y)$, that depends on the consumed quantity y . We identify two properties that the utility functions must have. Both are relatively weak conditions. The first one is related to the optimality conditions, and are verified if the functions are concave, or quasi-concave. The second one is related to the specification of the error terms, needed for the derivation of the econometric model. It is assumed that there exists an order preserving function ϕ such that the random utility can be expressed in additive form, that is such that

$$\phi \left(\frac{\partial U_{in}}{\partial e_i} \right) = V_{in} + \varepsilon_{in}, \quad (1)$$

where V_{in} is the deterministic part, and ε_{in} is a random disturbance. On top of this, the specification of the utility functions should be associated with specific behavioral assumptions. This aspect is not discussed in this document, in order to keep the model as general as possible.

The purchase decisions $y_{1n}, \dots, y_{J_n, n}$ are assumed to be the solution of the following optimization problem:

$$\max_y \sum_{i \in \mathcal{C}_n} U_{in}(y_i) \quad (2)$$

subject to

$$\sum_{i \in \mathcal{C}_n} p_{in} y_i = E_n, \quad (3)$$

$$y_i \geq 0, \quad (4)$$

We refer the reader to Appendix A, where the model is derived from first principles.

2 Model specifications

Biogeme implements the following specifications for the utility function.

2.1 Translated utility function

In the context of a time use model, Bhat (2005) uses the translated utility function introduced by Kim et al. (2002, Eq. (1)):

$$U_{in}(e_i) = \exp(\beta^T x_{in} + \varepsilon_{in})(e_i + \gamma_i)^{\alpha_i}, \quad (5)$$

where $0 < \alpha_i < 1$ and $\gamma_i > 0$ are parameters to be estimated. Note that there is no price involved in this specification, as it models time and not goods. It is equivalent to set $p_i = 1$, for all i in the model described above. Note also that, Kim et al. (2002) and Bhat (2005) impose the following restriction on α_i : $0 < \alpha_i \leq 1$. However, in the context of this implementation, $\alpha_i = 1$ would create a singularity.

We can calculate

$$\frac{\partial U_{in}}{\partial e_i} = \exp(\beta^T x_{in} + \varepsilon_{in}) \alpha_i (e_i + \gamma_i)^{\alpha_i - 1}. \quad (6)$$

In this context, we use the logarithm as the order preserving function to obtain the following specification:

$$\phi \left(\frac{\partial U_{in}}{\partial e_i} \right) = \beta^T x_{in} + \varepsilon_{in} + \ln \alpha_i + (\alpha_i - 1) \ln(e_i + \gamma_i),$$

so that

$$V_{in} = \beta^T x_{in} + \ln \alpha_i + (\alpha_i - 1) \ln(e_i + \gamma_i),$$

and

$$c_{in} = -\frac{\partial V_{in}}{\partial e_i} = \frac{1 - \alpha_i}{e_i + \gamma_i}.$$

In Biogeme, this model is referred to as `translated`.

3 Generalized translated utility function

Bhat (2008) generalizes the above formulation and introduces the following specification, where the utility functions U_{in} are defined as

$$U_{1n} = \exp(\beta^T x_{1n} + \varepsilon_{1n}) \frac{1}{\alpha_1} \left(\frac{e_1}{p_1} \right)^{\alpha_1}, \quad (7)$$

for the “outside good” that is always consumed and, for $i > 1$,

$$U_{in} = \exp(\beta^T x_{in} + \varepsilon_{in}) \frac{\gamma_i}{\alpha_i} \left[\left(\frac{e_i}{p_i \gamma_i} + 1 \right)^{\alpha_i} - 1 \right], \quad (8)$$

where β , $0 < \alpha_i < 1$ and $\gamma_i > 0$ are parameters to be estimated. We can calculate

$$\frac{\partial U_{1n}}{\partial e_1} = \exp(\beta^T x_{1n} + \varepsilon_{1n}) \frac{1}{p_1} \left(\frac{e_1}{p_1} \right)^{\alpha_1 - 1}, \quad (9)$$

and

$$\frac{\partial U_{in}}{\partial e_i} = \exp(\beta^T x_{in} + \varepsilon_{in}) \frac{1}{p_i} \left(\frac{e_i}{p_i \gamma_i} + 1 \right)^{\alpha_i - 1}. \quad (10)$$

We use again the logarithm as the order preserving function to obtain the following specification:

$$\phi \left(\frac{\partial U_{1n}}{\partial e_1} \right) = \beta^T x_{1n} + \varepsilon_{1n} + (\alpha_1 - 1) \ln e_1 - \alpha_1 \ln p_1,$$

and

$$\phi \left(\frac{\partial U_{in}}{\partial e_i} \right) = \beta^T x_{in} + \varepsilon_{in} - \ln p_i + (\alpha_i - 1) \ln \left(\frac{e_i}{p_i \gamma_i} + 1 \right),$$

so that

$$V_{1n} = \beta^T x_{1n} + (\alpha_1 - 1) \ln e_1 - \alpha_1 \ln p_1,$$

and

$$V_{in} = \beta^T x_{in} - \ln p_i + (\alpha_i - 1) \ln \left(\frac{e_i}{p_i \gamma_i} + 1 \right).$$

Finally, we have

$$c_{1n} = -\frac{\partial V_{1n}}{\partial e_1} = \frac{1 - \alpha_1}{e_1},$$

and

$$c_{in} = -\frac{\partial V_{in}}{\partial e_i} = \frac{1 - \alpha_i}{e_i + p_i \gamma_i}.$$

In Biogeme, this model is referred to as `generalized`.

3.1 The γ -profile

If $\alpha_i \rightarrow 0$, (8) collapses to the linear expenditure system (LES) form, defined as follows:

$$U_{1n}(e_1) = \exp(\beta^T x_{1n} + \varepsilon_{1n}) \ln \left(\frac{e_1}{p_1} \right), \quad (11)$$

and

$$U_{in}(e_i) = \exp(\beta^T x_{in} + \varepsilon_{in}) \gamma_i \ln \left(\frac{e_i}{p_i \gamma_i} + 1 \right), \quad (12)$$

where $\beta, \gamma_i > 0$ is a parameter to be estimated. We can calculate

$$\frac{\partial U_{1n}}{\partial e_1} = \exp(\beta^T x_{1n} + \varepsilon_{1n}) \frac{1}{e_1},$$

and

$$\frac{\partial U_{in}}{\partial e_i} = \exp(\beta^T x_{in} + \varepsilon_{in}) \frac{\gamma_i}{e_i + p_i \gamma_i}.$$

Taking again the logarithm, we have

$$\phi \left(\frac{\partial U_{1n}}{\partial e_1} \right) = \beta^T x_{1n} + \varepsilon_{1n} - \ln e_1,$$

and

$$\phi \left(\frac{\partial U_{in}}{\partial e_i} \right) = \beta^T x_{in} + \varepsilon_{in} + \ln \gamma_i - \ln(e_i + p_i \gamma_i),$$

so that

$$V_{1n} = \beta^T x_{1n} - \ln e_1,$$

and

$$V_{in} = \beta^T x_{in} + \ln \gamma_i - \ln(e_i + p_i \gamma_i).$$

Finally, we have

$$c_{1n} = \frac{1}{e_1},$$

and

$$c_{in} = \frac{1}{e_i + p_i \gamma_i}.$$

In Biogeme, this model is referred to as `gamma_profile`.

3.2 The non-monotonic model

Wang et al. (2023) introduce the following specification in order to accommodate non-monotonic preferences, motivated in the context of time consumption. With our notations, the specification is

$$U_{1n}(e_1) = \frac{1}{\alpha_1} e^{\beta^\top x_{1n}} \left(\frac{e_1}{p_1} \right)^{\alpha_1} + (\gamma^\top z_{1n} + \varepsilon_{1n}) e_i.$$

for the outside good, and

$$U_{in}(e_i) = \frac{\gamma_i}{\alpha_i} e^{\beta^\top x_{in}} \left[\left(\frac{e_i}{p_i \gamma_i} + 1 \right)^{\alpha_i} - 1 \right] + (\gamma^\top z_{in} + \varepsilon_{in}) e_i.$$

where β , γ , $0 < \alpha_i < 1$, and $\gamma_i > 0$ are parameters to be estimated. In the original paper, the model has been derived using $p_i = 1$, for all i . We have

$$\frac{\partial U_{1n}}{\partial e_1} = \frac{1}{p_1} e^{\beta^\top x_{1n}} \left(\frac{e_1}{p_1} \right)^{\alpha_1-1} + \gamma^\top z_{1n} + \varepsilon_{1n},$$

and

$$\frac{\partial U_{in}}{\partial e_i} = \frac{1}{p_i} e^{\beta^\top x_{in}} \left(\frac{e_i}{p_i \gamma_i} + 1 \right)^{\alpha_i-1} + \gamma^\top z_{in} + \varepsilon_{in}.$$

Here, we do not need to apply an order preserving function, so that

$$V_{1n} = \frac{1}{p_1} e^{\beta^\top x_{1n}} \left(\frac{e_1}{p_1} \right)^{\alpha_1-1} + \gamma^\top z_{1n}.$$

and

$$V_{in} = \frac{1}{p_i} e^{\beta^\top x_{in}} \left(\frac{e_i}{p_i \gamma_i} + 1 \right)^{\alpha_i-1} + \gamma^\top z_{in}.$$

We have

$$c_{1n} = -\frac{\partial V_{1n}}{\partial e_1} = \frac{1}{p_1^2} e^{\beta^\top x_{1n}} (1 - \alpha_1) \left(\frac{e_1}{p_1} \right)^{\alpha_1-2},$$

and

$$c_{in} = -\frac{\partial V_{in}}{\partial e_i} = \frac{1}{p_i^2} e^{\beta^\top x_{in}} \frac{1 - \alpha_i}{\gamma_i} \left(\frac{e_i}{p_i \gamma_i} + 1 \right)^{\alpha_i-2}.$$

For the outside good, we have

$$\ln c_{1n} = \beta^\top x_{1n} + \ln(1 - \alpha_1) + (\alpha_1 - 2) \ln e_1 - \alpha_1 \ln p_1,$$

and

$$\frac{1}{c_{1n}} = p_1^2 e^{-\beta^\top x_{1n}} \frac{1}{1 - \alpha_1} \left(\frac{e_1}{p_1} \right)^{2-\alpha_1}.$$

For the other goods, we have

$$\ln c_{in} = -2 \ln p_i + \beta^T x_{in} + \ln(1 - \alpha_i) - \ln \gamma_i + (\alpha_i - 2) \ln \left(\frac{e_i}{p_i \gamma_i} + 1 \right).$$

and

$$\frac{1}{c_{in}} = p_i^2 e^{-\beta^T x_{in}} \frac{\gamma_i}{1 - \alpha_i} \left(\frac{e_i}{p_i \gamma_i} + 1 \right)^{2 - \alpha_i}.$$

In Biogeme, this model is referred to as `non_monotonic`.

References

- Bhat, C. R. (2005). A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions, *Transportation Research Part B: Methodological* **39**(8): 679 – 707.
- Bhat, C. R. (2008). The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions, *Transportation Research Part B: Methodological* **42**(3): 274–303. A Tribute to the Career of Frank Koppelman.
URL: <https://www.sciencedirect.com/science/article/pii/S0191261507000677>
- Kim, J., Allenby, G. M. and Rossi, P. E. (2002). Modeling consumer demand for variety, *Marketing Science* **21**(3): 229–250.
- Wales, T. J. and Woodland, A. D. (1983). Estimation of consumer demand systems with binding non-negativity constraints, *Journal of Econometrics* **21**(3): 263–285.
- Wang, M., Ye, X. and Bierlaire, M. (2023). Accommodating non-monotonic preference in multiple discrete-continuous extreme value (MDCEV) choice model: Formulation and application to time use decisions. submitted to TRB.

A Derivation of the model

To simplify the derivation that follows, we write the optimization problem as an equivalent minimization problem, in terms of expenditures instead of quantities:

$$\min_e - \sum_{i \in \mathcal{C}_n} u_{in}(e_i) \quad (13)$$

subject to

$$\sum_{i \in \mathcal{C}_n} e_i = E_n, \quad (14)$$

$$e_i \geq 0. \quad (15)$$

The optimality conditions can be derived from the Lagrangian of the problem. We introduce a Lagrange multiplier λ associated with constraint (14) and a Lagrange multiplier $\mu_i \geq 0$ for each constraint (15). The Lagrangian is defined as

$$\mathcal{L}(y; \lambda, \mu) = - \sum_{i \in \mathcal{C}_n} U_{in}(e_i) + \lambda \left(E_n - \sum_{i \in \mathcal{C}_n} e_i \right) - \sum_{i \in \mathcal{C}_n} \mu_i e_i. \quad (16)$$

The first order optimality conditions state that

$$\frac{\partial \mathcal{L}}{\partial e_i} = - \frac{\partial U_{in}}{\partial e_i} - \lambda - \mu_i = 0, \text{ and } \mu_i e_i = 0, \forall i \in \mathcal{C}_n. \quad (17)$$

Note that we assume that the second order optimization conditions are also verified. This is the case if the utility functions are concave, for instance.

At least one item is consumed. We assume without loss of generality that it is item 1. Therefore $e_{1n} > 0$ and $\mu_1 = 0$. Consequently, (17) can be written

$$\lambda = - \frac{\partial U_{1n}}{\partial e_1}.$$

Consider an alternative $j \neq 1$ such that $e_{jn} > 0$. Using the same argument, we can write

$$\frac{\partial U_{jn}}{\partial e_j} = \frac{\partial U_{1n}}{\partial e_1}.$$

Consider an alternative j such that $e_{jn} = 0$. In this case, $\mu_j \geq 0$ and

$$\frac{\partial U_{in}}{\partial e_i} \leq \frac{\partial U_{1n}}{\partial e_1}.$$

Note that we can transform the utility functions U_{in} with any order preserving function F without changing the solution of the problem. An order preserving function is a strictly increasing function F of one variable such that $F'(u) > 0$. In that case,

$$\phi \left(\frac{\partial U_{in}}{\partial e_i} \right) \leq \phi \left(\frac{\partial U_{1n}}{\partial e_1} \right) \iff \frac{\partial U_{in}}{\partial e_i} \leq \frac{\partial U_{1n}}{\partial e_1}.$$

We now assume that the utility function is a random variable, that can be written in an additive way. More specifically, we assume that there exists an order preserving transform of the utility functions such that

$$\phi \left(\frac{\partial U_{in}}{\partial e_i} \right) = V_{in} + \varepsilon_{in}, \quad (18)$$

where V_{in} is the deterministic part, and ε_{in} is a random disturbance. Therefore, the optimality conditions can be written as

$$V_{in} + \varepsilon_{in} = V_{ln} + \varepsilon_{ln}, \quad \text{if } e_{in} > 0, \quad (19)$$

$$V_{in} + \varepsilon_{in} \leq V_{ln} + \varepsilon_{ln}, \quad \text{if } e_{in} = 0. \quad (20)$$

We assume that the utility functions are defined in such a way that (19) defines a bijective relationship between e_{in} and ε_{in} , for all $i \in \mathcal{C}_n$.

B The distribution of expenditures

We are interested in the distribution of the vector e_n , and we have established that it is a function of the vector of disturbances ε_n : $e_n = H(\varepsilon_n)$. Consequently, if we assume a distribution for ε_n , characterized by a probability density function (pdf) f_ε and a cumulative distribution function (CDF) F_ε , we can characterize the distribution of e_n .

We start by assuming that e_{1n} , the consumed quantity of item 1 is known and non zero. Consequently, the value of ε_{1n} is known as well. In order to derive the pdf evaluated at e , we split the vector e into its positive entries e^+ and its zero entries e^0 , alternative 1 being excluded. In an analogous way, we denote \mathcal{C}^+ and \mathcal{C}^0 the corresponding sets of indices, of size J^+ and J^0 , respectively.

For each $i \in \mathcal{C}^+$, we can use (19) to define $\varepsilon_n = H^{-1}(e_n)$, where $H^{-1} : \mathbb{R}^{J^+-1} \rightarrow \mathbb{R}^{J^+-1}$ is defined as

$$H_i^{-1}(e) = \varepsilon_{i+1,n} = V_{ln}(e_1) - V_{i+1,n}(e_{i+1}) + \varepsilon_{1n}. \quad (21)$$

Therefore, the density function can be decomposed as

$$\begin{aligned} f_e(e^+, e^0 | e_1) &= f_e(e^+ | e^0, e_1) \Pr(e^0 | e_1) \\ &= f_\varepsilon(\varepsilon^+ | \varepsilon^0, \varepsilon_1) \det \left(\frac{\partial H^{-1}}{\partial e} \right) \Pr(e^0 | e_1) \end{aligned}$$

where

$$\begin{aligned} \Pr(e^0 | e_1) &= \Pr(V_{in} + \varepsilon_{in} \leq V_{ln} + \varepsilon_{ln}, \forall i \in \mathcal{C}^0) \\ &= \Pr(\varepsilon_{in} \leq V_{ln} - V_{in} + \varepsilon_{ln}, \forall i \in \mathcal{C}^0) \\ &= F_\varepsilon(\varepsilon_{1n}, 1, \dots, 1, (V_{ln} - V_{in} + \varepsilon_{1n})_{i \in \mathcal{C}^0}). \end{aligned}$$

From (21), we can calculate the entries of the Jacobian $\partial H^{-1}/\partial e$. Indeed,

$$\begin{aligned}\frac{\partial H_i^{-1}(e)}{\partial e_k} &= \frac{\partial V_{1n}}{\partial e_1} \frac{\partial e_1}{\partial e_k} && \text{if } k \neq i+1, \\ &= \frac{\partial V_{1n}}{\partial e_1} \frac{\partial e_1}{\partial e_k} - \frac{\partial V_{i+1,n}}{\partial e_{i+1}} && \text{if } k = i+1.\end{aligned}$$

From (14), we have

$$e_1 = E - \sum_{j \neq 1} e_j,$$

so that

$$\frac{\partial e_1}{\partial e_k} = -1.$$

Consequently,

$$\begin{aligned}\frac{\partial H_i^{-1}(e)}{\partial e_k} &= -\frac{\partial V_{1n}}{\partial e_1} && \text{if } k \neq i+1, \\ &= -\frac{\partial V_{1n}}{\partial e_1} - \frac{\partial V_{i+1,n}}{\partial e_{i+1}} && \text{if } k = i+1.\end{aligned}$$

If we denote

$$c_i = -\frac{\partial V_{1n}}{\partial e_i} \tag{22}$$

the Jacobian has the following structure:

$$\partial H^{-1}/\partial e = \begin{pmatrix} c_1 + c_2 & c_1 & \cdots & c_1 \\ c_1 & c_1 + c_3 & \cdots & c_1 \\ & & \vdots & \\ c_1 & c_1 & \cdots & c_1 + c_{J_n} \end{pmatrix}.$$

Therefore, the determinant is equal to

$$\det(\partial H^{-1}/\partial e) = \left(\prod_{i=1}^{J^+} c_i \right) \left(\sum_{i=1}^{J^+} \frac{1}{c_i} \right).$$

Note that this determinant depends only on the utility function, not on the distribution of the ε_n .

Therefore, the density function of the expenditures is given by

$$f_e(e^+, e^0) = \left(\prod_{i=1}^{J^+} c_i \right) \left(\sum_{i=1}^{J^+} \frac{1}{c_i} \right) \int_{\varepsilon_1 = -\infty}^{+\infty} f_\varepsilon(\varepsilon^+ | \varepsilon^0, \varepsilon_1) F_\varepsilon(\varepsilon_{1n}, \dots) f_\varepsilon(\varepsilon_1) d\varepsilon_1, \tag{23}$$

where $f_\varepsilon(\varepsilon_1)$ is the marginal distribution of ε_1 , and

$$F_\varepsilon(\varepsilon_{1n}, \dots) = F_\varepsilon(\varepsilon_{1n}, 1, \dots, 1, (V_{1n} - V_{in} + \varepsilon_{1n})_{i \in \mathcal{C}^0}).$$

Equation (23) corresponds to Equation 11 in Bhat (2008).

If we assume that the ε_{in} are independent, we obtain the MDCEV model introduced by Bhat (2005). In that case, the density (23) is

$$f_e(e^+, e^0) = \mu^{J^+ - 1} \left(\prod_{i=1}^{J^+} c_i \right) \left(\sum_{i=1}^{J^+} \frac{1}{c_i} \right) \left(\frac{\prod_{i \in \mathcal{C}^+} e^{\mu V_{in}}}{(\sum_{i \in \mathcal{C}_n} e^{\mu V_{in}})^{J^+}} \right) (J^+ - 1)!, \quad (24)$$

where the derivation is available in Bhat (2008). Therefore, the contribution of observation n to the log likelihood is

$$\begin{aligned} \ln f_e(e^+, e^0) = & (J^+ - 1) \ln \mu \\ & + \sum_{i=1}^{J^+} \ln c_i \\ & + \ln \left(\sum_{i=1}^{J^+} \frac{1}{c_i} \right) \\ & + \mu \sum_{i \in \mathcal{C}^+} V_{in} \\ & - J^+ \ln \sum_{i \in \mathcal{C}_n} e^{\mu V_{in}} \\ & + \ln(J^+ - 1)!. \end{aligned}$$

Note that the last term is a constant, and is ignored by Biogeme.