On allowing endogenous minimum consumption bounds in the Multiple Discrete Continuous Choice Model: An application to expenditure patterns

Andrea Pellegrini[1], John Rose[1]

[1] Institute of Transport and Logistics Studies,

Business School, University of Sydney

SHORT SUMMARY

This study develops a novel econometric model that allows for the endogenous identification of minimum bounds on consumption. This is done by combining a censored Tobit model with a Multiple Discrete Continuous Extreme Value model. Whilst the former is employed to identify the minimum consumption of a good/service based upon the socio-demographic characteristics of the consumer, the latter is used to assess multiple discrete continuous consumption patterns. The proposed modelling framework is applied to investigate individuals' expenditure behaviour with the attention being placed on the following expenditure categories: Transport, Shopping, Child Care, Entertainment, Household Bills, and Rent/Mortgage.

Keywords: Multiple Discrete Continuous Decisions, Minimum Consumption Patterns, Expenditure Behavior, Censored Tobit Model

1. INTRODUCTION

Numerous consumption choices result in economic-agents selecting multiple goods or services at the same time. Along with the selection of the type of good to purchase, consumers also usually decide how much of the selected product to consume, with the latter representing a continuous quantity dimension within the underlying decision-making process. These choice situations have thus far been widely analysed in the economics literature via the use of multiple discrete continuous (MDC) demand models. Unlike traditional single discrete choice models, MDC methods allow for assessing interior (goods are assumed to be imperfect substitutes) and corner solutions (goods are treated as perfect substitutes), whilst also accounting for potential satiation effects. The latter effects typically arise with increasing consumption of a good or a service. Since its advent, the Multiple Discrete Continuous Extreme Value (MDCEV) model (Bhat, 2005; 2008) has become the benchmark framework within the transportation literature for analysing individuals' decisions being discrete and continuous in nature. Further, the MDCEV model has been extended in a variety of directions. These are, for example, a) heterogeneity in parameters (Bhat et al., 2016; Shobani et al., 2013), b) flexible utility profiles (Palma and Hess, 2022; Pellegrini et al., 2021; Bhat, 2018, Bhat et al., 2015), and c) multiple constraints (Mondal and Bhat, 2021; Castro et al., 2012).

A downside of many MDC models is that they fail to account for the impact of minimum consumption patterns for goods/services on the analysis of consumer behaviour. For example, it is reasonable to believe that leisure activities (going to a museum) cannot be performed without devoting a certain minimum amount of time for it, or travelling to holiday destinations cannot be done without individuals spending a certain minimum amount of monetary expenditure. The first attempt to incorporate lower bounds on consumption quantities traces back to Van Nostrand et al. (2012) who recast the utility function formulated in Bhat (2008) to capture the existence of a minimum amount of time allocated to location destination vacation choices (see, Astroza et al., 2017, for an application involving time-use data). Specifically, the authors observed the minimum time that each individual allotted for every destination under investigation and imposed this value as the lower bound when estimating the log-likelihood function. The empirical evidence from the study suggested the implementation of lower bounds on consumption patterns provided more realistic predictions of the time use allocation decisions, whilst also resulting in a better goodness of fit relative to the traditional MDCEV model. Recently, Saxena et al. (2021) proposed an extension of the MDCEV model in which the analyst is able to integrate both lower and upper bounds on consumption choices into a trackable and flexible framework. Despite such improvements, lower bounds on consumption are imposed disregarding the socio-demographic and economics characteristics of the consumers. It is quite unrealistic to assume that households with children versus childless households allot the same minimum time value for entertainment activities. Likewise, house owners versus tenants are unlikely to engage with similar minimum household related expenses. As such, it is likely necessary to exploit individual consumer differences to better understand consumer preference behaviour.

Given the above considerations, the aim of this paper is to develop a model of MDC demand wherein lower bounds on consumptions are specified based upon the demographic information of different decision-makers. To do this, we propose using a censored Tobit regression model to identify minimum consumption patterns for the discrete alternatives, which subsequently serve as lower bounds within a MDCEV model. The developed framework is adopted for investigating consumer expenditure behaviour of 858 Australian residents of New South Wales (NWS), Australia. The expenditure categories involved in this study comprise of nine different categories, consisting of transport, shopping, childcare, entertainment, household bills, and rent/mortgage payments, savings, miscellaneous costs, and other expenditure items.

METHODOLOGY

The framework that we formulate for this study consists of the joint estimation of two empirical methods, namely the censored Tobit model and the MCEV model. Specifically, a censored Tobit model is first employed to the identify the minimum amount of money that each respondent i spent on the expenditure category j, such that

$y_{ij}^* = \beta_0 + \beta K + \epsilon_{ij}, \epsilon_{ij} \sim N(0, \sigma_j^2)$ (1)

where K is a $N \times T$ matrix that describes the socio-demographic and economic characteristics of the consumers, y_{ij}^* is a latent variable, β is a vector $T \times 1$ of parameters to be estimated, and ϵ_{ij} are independently and identically normally distributed error terms. The amount of money spent by the respondent *i* on the alternative *j* is specified such that

$$y_{ij} = \begin{cases} y_{ij}^* \text{ if } y_{Lj} < y_{ij}^* < y_{Uj} \\ y_{Lj} \text{ if } y_{ij}^* < y_{Lj} \\ y_{Uj} \text{ if } y_{ij}^* > y_{Uj} \end{cases} \quad (2)$$

In the above equation, the Tobit model is assumed to be bounded between (y_{Lj}, y_{Uj}), with y_{Lj} being the below censored point whereas y_{Uj} representing the above censored point. By computing the conditional expectation, $y_{ij}^0 = E\left[y_{ij} \middle| y_{ij}^* \text{ if } y_{Lj} < y_{ij}^* < y_{Uj}, X\right]$, we are able to endogenously determine the minimum expenditure that respondents made for each alternative *j*. The estimated minimum expenditure is then integrated into the utility function of the MDCEV model,

$U_i(\boldsymbol{Y}) = y_{i1}\psi_1 + \sum_{j=2}^{J} \delta_{ij}(y_{ij})$. (3)

In Equation (3), y_{ij} are inside goods (i.e., the expenditure categories under scrutiny) whilst y_{i1} is the linear outside good whose role is to acknowledge the possibility of consumers spending money on expenditure categories other than those analysed in the analysis. The appeal of a linear formulation for the outside good resides in the fact that the corresponding first derivate equals one and hence drops out in the calculation of the KKT conditions for optimality. Further, $\delta_{ij}(y_{ij})$ can be written such as

$$\begin{aligned} \delta_{ij}(y_{ij}) &= \psi_{ij} y_{ij} \ \text{ if } y_{ij} \leq y_{ij}^0 \\ \delta_{ij}(y_{ij}) &= \psi_{ij} y_{ij}^0 + \gamma_{ij}\psi_{ij} \ln\left(\frac{y_{ij} - y_{ij}^0}{\gamma_{ij}} + 1\right) \text{ if } y_{ij} > y_{ij}^0 \end{aligned} \quad (4)$$

where γ_{ij} govern satiation patterns, ψ_{ij} is the baseline marginal utility at the point of zero expenditure, and y_{ij}^0 is the minimum required expenditure of an alternative *j* (if it is selected). Both γ_{ij} and ψ_{ij} can be further parametrized to incorporate demographic and economic variables attached to the consumer *i* as below

$$\begin{aligned} \gamma_{ij} &= \exp(\delta r_{ij}) \\ \psi_{ij} &= \exp(\alpha r_{ij} + \tau_j) \end{aligned} \quad (5)$$

where τ_j are independently and identically error terms with a type 1 extreme value distribution

The underlying assumption is that the economic-agent *i* is assumed to maximize the utility function expressed in Equation (3), $U_i(\boldsymbol{Y})$, subject to a monetary constraint $y_{i1} + \sum_{j=2}^{J} y_{ij} = \mathrm{M}_i$, where M_i represents the available monetary budget. The optimal expenditure allocations $y_{ij}^* (j = ,1 \ldots, J)$ can be obtained by forming the Lagrangian function and applying the KKT conditions for optimality, such that

$$\mathcal{L} = U(Y) - \lambda\left(y_{i1} + \sum_{j=2}^{J} y_{ij} - M\right), (6)$$

where λ is the Lagrangian multiplier associated with the specified monetary budget constraint. The KKT conditions for the optimal expenditure allocations for the individual *i* are given by:

$$u_i'(y_{ij}^*) - \lambda = 0, \text{ if } y_{ij}^* > 0, j = 1, \dots, J$$
$$u_i'(y_{ij}^*) - \lambda < 0, \text{ if } y_{ij}^* = 0, j = 1, \dots, J \; (7)$$

where $\lambda = \psi_1$, with $\psi_1 = \exp(\tau_1)$

The KKT conditions formalized in Equation (7) can be re-written as:

$$V_{ij} + \tau_j = \psi_1 \text{ if } y_{ij}^* > 0, j = 1, \dots, J$$
$$V_{ij} + \tau_j < \psi_1 \text{ if } y_{ij}^* = 0, j = 1, \dots, J \; (8)$$

where $V_{ij} = \psi_{ij}$ if $y_{ij} \leq y_{ij}^0$ whilst $V_{ij} = \psi_{ij}\left(\frac{y_{ij} - y_{ij}^0}{\gamma_{ij}} + 1\right)^{-1}$ if $y_{ij} > y_{ij}^0$

The maximum likelihood estimation method is next adopted for estimating the parameters of the two methodological approaches involved in the optimization process.

2. DATA

Respondents 18 years and older drawn from New South Wales Australia were recruited using the online panel QOR Surveys (www. https://www.qorsurveys.com.au/) between 27th March and 20th April 2022 and asked to complete an online survey associated with environmental issues related to the Murray Darling Basin area. As part of the survey, respondents were asked to provide a summary of their average monthly household expenditure across nine different expenditure categories. The nine expenditure categories consisted of expenditure on transport, shopping, child care, entertainment, household bills, and rent/mortgage payments, savings, miscellaneous costs, and other expenditure items. After providing information regarding actual expenditure, respondents were next asked what would be the absolute minimum amount of expenditure that could be allocated to each category over a given month. As such, the survey captured both actual expenditure as well as the minimum perceived expenditure amount across each expense category.

A total of 2,056 respondents completed the survey. After extensive data cleaning involving removing data from respondents whose total survey response time was considered to be too fast to meaningfully complete the survey, who undertook straight lining behaviour in answering attitudinal questions, and who provided inconsistent responses to survey questions such as suggesting minimum expenditure amounts were necessary that were greater than actual amounts reported, the final sample consists of data collected from 858 respondents. The socio-demographic characteristics of this final sample match the recent 2021 Census values, with the sole exception of being slightly more skewed towards males (54.6% relative to 49.3%).

3. RESULTS

In addition to the proposed model, we also estimated a MDCEV model in which minimum bounds on consumption are imposed without accounting for the differences among consumers (bounds are exogenous to the model). The goodness of fit measures reveal that the framework developed herein outperforms the model with exogenous bounds on expenditure patterns. Focusing on the empirical findings, we found that large households (with at least one adult) are more likely to spend on utility bills than single-person households. The presence of kids results in households spending more on childcare followed by entertainment and shopping. Full-time workers are inclined to allocate a larger portion of the monetary budget to transportation and utility bills categories. Finally, we conducted a simulation study to compare the forecasting performance of the two models estimated in this study. The evidence suggests that the proposed modelling framework provides more accurate predictions relative to the model where bounds on consumptions are exogenously specified, suggesting that respondents do indeed have different minimum consumption patterns.

4. CONCLUSIONS

In this study, we restrict our attention to the formulation of a novel econometric model that permits the identification of minimum bounds on consumption based upon the sociodemographic and economics characteristics of the decision-makers. Unlike traditional applications where minimum bounds are exogenous to the model, we exploit the features of a censored Tobit model to endogenously determine the minimum amount of money that can be potentially spent by respondents. The estimated minimum consumption value is then used in the MCEV model which is employed to assess individuals' expenditure behaviour.

REFERENCES

1. A., Mondal and C.R. Bhat. A New Closed Form Multiple Discrete-Continuous Extreme Value (MDCEV) Choice Model with Multiple Linear Constraints. Transportation Research Part B, (2021) Vol. 147, pp. 42-66.
2. A., Pellegrini, Pinjari, A.R., and R. Maggi. A multiple discrete continuous model of time use that accommodates non-additively separable utility functions along with time and monetary budget constraints. Transportation Research Part A: Policy and Practice 144, (2021) 37-53.
3. A., Sobhani, N., Eluru, and A., Faghih-Imani. A latent segmentation based multiple discrete continuous extreme value model. Transportation Research Part B: Methodological. 58, (2013) 154-169.
4. C. Van Nostrand, V. Sivaraman, and A.R. PinjariAnalysis of long-distance vacation travel demand in the United States: A multiple discrete-continuous choice framework. Transportation, 40 (1) (2013), pp. 151-171.
5. C.R. Bhat, S. Astroza, and A.C. Bhat. On Allowing a General Form for Unobserved Heterogeneity in the Multiple Discrete-Continuous Probit Model: Formulation and Application to Tourism Travel. Transportation Research Part B: Methodological, 86 (2016), pp. 223-249.
6. C.R. Bhat. A Multiple Discrete-Continuous Extreme Value Model: Formulation and Application to Discretionary Time-Use Decisions. Transportation Research Part B: Methodological., 39 (8) (2005), pp. 679-707.
7. C.R. Bhat. A new flexible multiple discrete–continuous extreme value (MDCEV) choice model. Transportation Research Part B: Methodological., 110 (2018), pp. 261-279.
8. D. Palma, and S. Hess/ Some Adaptations of Multiple Discrete-Continuous Extreme Value (MDCEV) Models for a Computationally Tractable Treatment of

Complementarity and Substitution Effects, and Reduced Influence of Budget Effects. Transportation Research Part B: Methodological, (2022).
9. M. Castro, R. Paleti, and C.R. Bhat. A latent variable representation of count data models to accommodate spatial and temporal dependence: Application to predicting crash frequency at intersections. Transportation Research Part B: Methodological., 46 (1) (2012), pp. 253-272.
10. S Saxena, AR Pinjari, A Roy, and R Paleti. Multiple discrete-continuous choice models with bounds on consumptions. Transportation Research Part A: Policy and Practice 149, 237-265.
11. S., Astroza, A.R. Pinjari, C.R. Bhat, and S.R. Jara-Diaz. A Microeconomic Theory-based Latent Class Multiple Discrete-Continuous Choice Model of Time Use and Goods Consumption. Transportation Research Record, (2017) Vol. 2664, pp. 31-41.