Multiple Discrete-Continuous Choice Models with Flexible and Partially Monotonic Utility Functions

Huichang Lee[1], Jason Hawkins[2], Prateek Bansal*[3], Dong-Kyu Kim[4], Eui-Jin Kim[5]

[1] PhD Student, Department of Civil and Environmental Engineering, Seoul National University, Republic of Korea

[2] Assistant Professor, Department of Civil Engineering, University of Calgary, Canada

[3] Assistant Professor, Department of Civil and Environmental Engineering, National University of Singapore, Singapore

[4] Professor, Department of Civil and Environmental Engineering, Seoul National University, Republic of Korea

[5] Assistant Professor, Department of Transportation Systems Engineering, Ajou University, Republic of Korea

SHORT SUMMARY

The multiple discrete-continuous extreme value (MDCEV) model estimates individuals' preferences for multiple alternatives and their usage, while accounting for satiation–the diminishing marginal utility from consuming additional units of each alternative–in a closed form. However, existing models lack flexibility due to specific assumptions about the utility function (e.g., monotonicity and parabolicity), leading to poor finite sample properties and prediction errors when the true data-generating process deviates from these assumptions. This study relaxes these assumptions by specifying the satiation parameters using lattice networks (LN), piecewise linear functions that flexibly model nonlinear attribute effects and employ multilinear interpolation to capture complex attribute interactions. The proposed MDCEV-LN demonstrated high predictive accuracy in budget allocation in a Monte Carlo study. At the same time, it maintains interpretability with added flexibility to accommodate various functional forms, including traditional log-linear satiation trend. Thus, MDCEV-LN offers an accurate, flexible and interpretable framework for discrete-continuous choice analysis.

Keywords: Multiple discrete-continuous choice model; Deep neural networks; Lattice networks; Partial monotonicity; Interpretability.

1. INTRODUCTION

Activity-based models (ABMs) describe travel demand as the outcome of activity time-use decisions. ABMs estimate how individuals seek to fulfill their preferences to engage in various recreational and social activities and distribute time across these activities within a given time budget (Arentze and Timmermans, 2004; Bhat, 2005).

Multiple discrete-continuous choice models are suitable to elicit such activity time-use decisions as they can jointly model discrete alternative choices and continuous budget allocation (Hanemann, 1984). Early discrete-continuous choice models (Hanemann, 1984; Dubin and McFadden, 1984) were limited to extreme corner solutions, where only a single alternative could be chosen (Hanemann et al., 2024). Kim et al. (2002) extended this framework to general corner solutions, allowing the simultaneous choice of multiple alternatives through a translated nonlinear additive utility function. However, their approach lacked a closed form and was computationally intractable. Bhat (2005, 2008) addressed this limitation by proposing the multiple discrete-continuous extreme value (MDCEV) model, introducing a multiplicative log-extreme value error term into the utility function. The MDCEV model, as an extension of the multinomial logit (MNL) model, provides a closed-form choice probability functions for modeling multiple discrete and continuous preferences.

However, existing MDCEV studies typically assume, based on domain knowledge, how satiation over time affects the utility function. These assumptions may not align with intuitive or observed behavior (e.g., monotonically increasing utilities fail to capture attributes with potential non-monotonic patterns) (Wang and Ye, 2024). Neural networks can significantly enhance the predictability of utility-based models by flexibly capturing the nonlinear relationships (Sifringer et al., 2020; Han et al. 2022). However, the excessive flexibility of neural networks may lead to violation of domain knowledge assumptions or result in misinterpretations (Kim and Bansal, 2024). There is a need to specify utility functions through more interpretable yet flexible functions than traditional neural networks that can accommodate various functional forms.

To this end, the lattice network (LN) (You et al., 2017) can flexibly represent utility functions as piecewise linear functions. A recent study by Kim and Bansal (2024) demonstrated the application of LN to specify utility functions in a discrete choice model that maintains partial monotonicity for a subset of attributes (e.g., utility decreases as travel cost increases) while offering flexibility comparable to neural networks.

This study proposes a novel method to flexibly and interpretably estimate the parameters of the MDCEV model using LNs. Compared to traditional MDCEV models, the proposed approach captures nonlinear and high-dimensional relationships while providing a more flexible representation of how individuals allocate time across various activities.

2. METHODOLOGY

In this section, we review the structure of the MDCEV model and explain the principles of lattice networks (LN). We then describe the specific architecture of lattice networks used to estimate the parameters of the MDCEV model.

MDCEV model

Bhat (2008) proposed a utility function that accommodates multiple discreteness in a closed form. Without loss of generality, assume that the first good (e.g., staying home activity) is the essential Hicksian composite outside good. The utility function in the traditional MDCEV model is expressed as follows:

$$U(\boldsymbol{t}) = \frac{1}{\alpha_1}\psi_1 t_1^{\alpha_1} + \sum_{k=2}^{K} \frac{\gamma_k}{\alpha_k}\psi_k \left\{\left(\frac{t_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right\} \tag{1}$$

where $U(\boldsymbol{t})$ is quasi-concave, increasing, and continuously differentiable with respect to the consumption quantity $(K \times 1)$-vector $\boldsymbol{t}$ ($t_k \geq 0$ for all k). The parameters ψ_k, α_k, and γ_k are associated with good k (e.g., activity-specific parameters). The function $U(\boldsymbol{t})$ is valid if $\psi_k > 0$, $\alpha_k \leq 1$, and $\gamma_k > 0$ for all k. Here, ψ_k represents the baseline marginal utility, or the marginal utility at zero consumption. γ_k and α_k serve as satiation parameters, influencing the consumption level of good k. Specifically, γ_k controls satiation by shifting consumption quantity threshold where utility starts diminishing, while α_k modulates satiation intensity by controlling how quickly utility diminishes as consumption increases. There is no γ_1 term for the first good (i.e., the composite outside good like staying home) because it is assumed that individuals always allocate some portion of their budget to it, regardless of other choices. In practice, disentangling the effects of γ_k and α_k is often challenging. Bhat (2008) suggests estimating models with both the α-profile and γ-profile, then selecting the specification with better statistical fit. Generally, the γ-profile has demonstrated better performance than the α-profile (Bhat et al., 2016) and is therefore more frequently adopted (Calastri et al., 2017; Jian et al., 2017). Moreover, the γ-profile provides a more intuitive explanation of consumption patterns (Pinjari and Bhat, 2021). When estimating the γ-profile, it is assumed that $\alpha_k \to 0\ \forall k$, resulting in the utility function taking the following form:

$$U(\boldsymbol{t}) = \psi_1 \ln t_1 + \sum_{k=2}^{K} \gamma_k \psi_k \ln\left\{\left(\frac{t_k}{\gamma_k}+1\right)\right\} \tag{2}$$

Bhat (2018) pointed out that the $\ln t_1$ term in the traditional MDCEV formulation necessitates the prediction of the continuous value of the outside good for calculating the discrete choice probability because small differences in $\ln t_1$ significantly affect the utility due to the logarithmic nature. This dependence leads to a tight linkage between the discrete and continuous components, as the $\ln t_1$ term relies on the specific combination of alternatives consumed (Bhat, 2018). To better separate the discrete component from the continuous component, Bhat (2018) proposed a linear-in-consumption utility for the outside good. Under this assumption, the L_γ-profile (linear utility for the outside good combined with a γ-profile for the inside goods) utility function is expressed as follows:

$$U(\boldsymbol{t}) = \psi_1 t_1 + \sum_{k=2}^{K} \gamma_k \psi_k \ln\left\{\left(\frac{t_k}{\gamma_k}+1\right)\right\} \tag{3}$$

This enables the prediction of discrete choices even in the absence of observations on continuous consumption, as the exact consumption quantity t_1 is less critical, given that it is proportionally increase the utility.

Lattice networks

The traditional MDCEV model assumes a monotonically increasing utility with continuous consumption, limiting its ability to capture non-monotonic utility patterns (Wang and Ye, 2024).

Consequently, utility misspecification in utility-based models can induce bias in inference of attribute effects and reduce the predictability (Sifringer et al., 2020). Deep neural networks (DNN) flexibly model the complex relationships in large-scale data, without relying on theoretical assumptions (van Cranenburgh et al., 2022). However, their complex structure result in low interpretability (Lipton, 2018) and sometimes produce counter-intuitive inference regarding attribute effects (Wang et al., 2021).

Lattice networks (LN) offer both flexibility and interpretability for discrete choice modeling by representing utility functions in a piecewise linear form (Kim and Bansal, 2024). LN consist of three layers: input calibrators, lattice functions, and output calibrators. The input calibrators transform real-valued inputs into values within a specific interval using a piecewise linear function, preparing them for the lattice layer. **Figure 1** illustrates an example of the piecewise linear transform in the input calibrators.

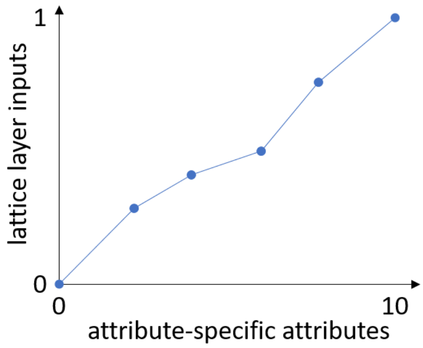

Figure 1: Piecewise Linear Transform in the Input Calibrators

Each lattice function operates as a linearly interpolated multidimensional look-up table. Let the dimension of the lattice function be denoted as S, and assume that the output from the input calibrators is normalized to the range $[0,1]$. Each S-dimensional look-up table takes the unit hypercube $[0,1]^S$ as input, and has 2^S parameters $\theta \in \mathbb{R}^{2^S}$ corresponding to the outputs at the vertices of the unit hypercube. Multilinear interpolation is used to estimate values between the vertices. Let the linear interpolation weights be denoted as $w(x)\colon [0,1]^S \to [0,1]^{2^S}$. These weights can be expressed as $w(x)^T\theta$, where each component of $w(x)$ is given by:

$$w(x)[j] = \prod_{d=1}^{S} x[d]^{v_j[d]}(1 - x[d])^{1-v_j[d]} \tag{4}$$

where $v_j[\cdot]$ is the coordinate vector of the jth vertex of the unit hypercube, and $j = 1, \dots, 2^S$. For example, as shown in **Figure 2**, if $x = (0.3, 0.7)$, $w(x) = \big((1 - 0.3)(1 - 0.7), (1 - 0.3)0.7, 0.3(1 - 0.7), 0.3 \cdot 0.7\big) = (0.21,0.49,0.09,0.21)$.

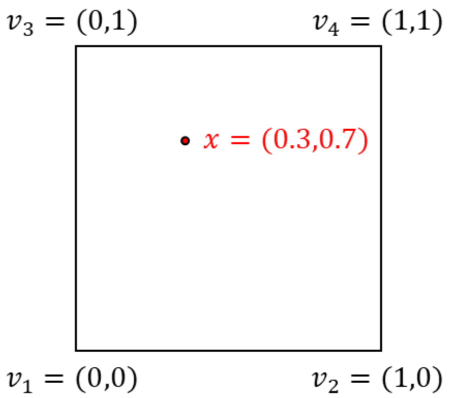

Figure 2: Multilinear Interpolation in the Lattice Function

The output calibration layer converts the lattice function's output as utility. The output calibration layer has exactly the same function structure as the input calibration layer.

Model structure

This study presents two data-driven models for activity time-use analysis. The first model, MDCEV-DNN, is a standard DNN that predicts discrete-continuous choice based on individual-specific attributes. The second model, MDCEV-LN, integrates DNN and LN, where DNN estimates ψ_k and LN estimates γ_k in **Equation 3**, respectively. **Figure 3** shows the structure of the MDCEV-DNN model. This model uses DNN to transform individual-specific attributes x into parameter ψ_k for each activity. Applying the softmax function to ψ_k outputs activity choice probabilities $p(1), \dots, p(K)$ that sum to 1. Multiplying these probabilities by the total budget T (24 hours for activity-time use analysis) provides prediction of time allocated to each activity. While this model can predict activity time-use using only individual-specific attributes, it has the disadvantage of being unable to separate the discrete and continuous component.

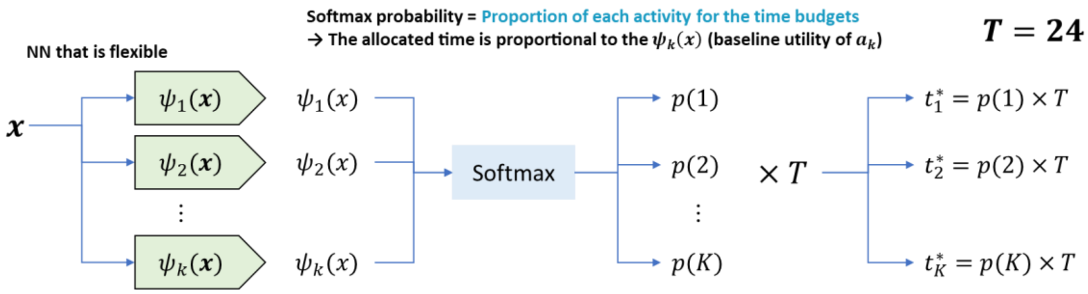

Figure 3: MDCEV-DNN Model Structure

The MDCEV-DNN model is trained to minimize the reconstruction error between the activity time allocation $\boldsymbol{t}$ and reconstructed $\boldsymbol{t}^*$. The reconstruction error is expressed as shown in **Equation 5**.

$$MSE = \frac{1}{K}\sum_{k=1}^{K}\left|t_k^* - t_k\right|^2 \tag{5}$$

The MDCEV-LN model is designed to separately capture the two parameters of the L_γ-profile in **Equation 3**, ψ_k and γ_k. **Figure 4** illustrates the structure of the MDCEV-LN model. The DNN transforms individual-specific attributes $\boldsymbol{x}$ into parameter ψ_k, which is associated with discrete choice. In this context, the DNN learns complex relationships between individual-specific attributes and ψ_k in a data-driven manner. The LN represents the satiation parameter γ_k as a piecewise linear function over time $\boldsymbol{t}$. As in **Equation 3**, the utility for each alternative is calculated by multiplying ψ_k and t_1 for the outside good, and multiplying ψ_k and γ_k for inside goods. Finally, applying softmax function to the utility and multiplying by the total budget T outputs the predicted time allocation for each activity. The MDCEV-LN model is trained to minimize the same reconstruction error as the MDCEV-DNN.

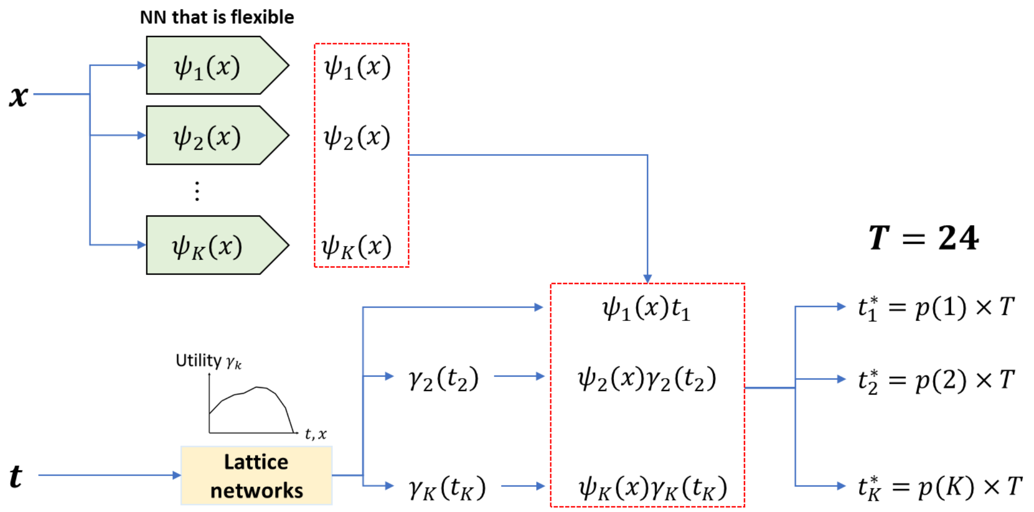

Figure 4: MDCEV-LN Model Structure

3. RESULTS AND DISCUSSION

We conducted a simulation study using synthetic data to evaluate the performances of the proposed models. Following Saxena et al. (2022a), we generated L_γ-profile simulation data for three scenarios. For detailed information about the synthetic data generation, refer to Saxena et al. (2022a). The three scenarios are distinguished by the proportion of inside goods as follows:

(a) Scenario 1: Total budget of 50,000 units, with very small inside goods consumption averaging less than 1% of the total budget
(b) Scenario 2: Total budget of 1,000 units, with moderate inside goods consumption averaging 16% of the total budget
(c) Scenario 3: Total budget of 1,000 units, with significant inside goods consumption averaging 43% of the total budget

We trained the proposed models using 80% of the synthetic data and evaluated their performance on the remaining 20% test set. The model performance was evaluated using the Brier score and Root Mean Squared Error (RMSE). The Brier score (Brier, 1950) measures the accuracy of probability prediction by quantifying distance between predicted probabilities and actual outcomes. For N individuals where K alternatives can be selected, the Brier score is defined as shown in **Equation 6**.

$$Brier\ Score = \frac{1}{N}\sum_{n=1}^{N}\sum_{k=1}^{K}\left(p_{n,k}^{*} - p_{n,k}\right)^2 \quad (6)$$

$p^*_{n,k}$ and $p_{n,k}$ represent the predicted and observed choice probabilities for the kth alternative in the nth individual, respectively. The Brier score decreases as prediction accuracy increases. **Table 1** shows the performance of MDCEV, MDCEV-DNN and MDCEV-LN across the three scenarios.

Table 1: Performance Evaluation in the Simulation Study

	Metrics	Model	Alt 1	Alt 2	Alt 3	Alt 4
Scenario 1	Brier score	MDCEV	0.0251	0.0033	0.0072	0.0143
		MDCEV-DNN	0.0203	0.0031	0.0067	0.0124
		MDCEV-LN	**0.0006**	**0.0004**	**0.0004**	**0.0005**
	RMSE	MDCEV	7,928	2,860	4,246	5,976
		MDCEV-DNN	7,120	2,765	4,095	5,575
		MDCEV-LN	**1,255**	**965**	**1,005**	**1,115**
Scenario 2	Brier score	MDCEV	0.0299	0.0036	0.0088	0.0186
		MDCEV-DNN	0.0246	0.0034	0.0077	0.0162
		MDCEV-LN	**0.0009**	**0.0004**	**0.0004**	**0.0004**
	RMSE	MDCEV	173.0	60.3	94.1	136.5
		MDCEV-DNN	156.8	58.3	87.6	127.4
		MDCEV-LN	**30.8**	**20.1**	**19.1**	**20.1**
Scenario 3	Brier score	MDCEV	0.1331	0.0786	0.0088	0.0625
		MDCEV-DNN	0.0928	0.0603	0.0083	0.0507
		MDCEV-LN	**0.0030**	**0.0015**	**0.0011**	**0.0016**
	RMSE	MDCEV	364.8	280.3	93.8	250.1
		MDCEV-DNN	304.7	245.6	90.9	225.1
		MDCEV-LN	**54.5**	**38.6**	**32.6**	**39.5**

Alt 1 represents the outside good, while Alt 2 through Alt 4 correspond to inside goods. The original MDCEV model showed low predictability despite the simulation data mimicking its data generation process. This aligns with the inconsistency of the original MDCEV reported in previous research (Saxena et al., 2022b). MDCEV-DNN showed higher predictability compared to MDCEV; however, the satiation effect cannot be isolated. The MDCEV-LN outperformed MDCEV-DNN in both Brier score and RMSE, demonstrating superior predictability. This suggests that the parameters captured by the MDCEV-LN structure accurately reconstruct the decision-making process of the L_γ-profile simulation data. **Figure 5** compares the satiation effects estimated by the LN with the true log-linear satiation effects of the simulation data across all scenarios. Note that utility is computed as the product of $\psi_k(x)$ and $\gamma_k(x)$; hence, their scales may differ. The vertical dotted lines in the figure represent the 95th percentile for each alternative. Values significantly exceeding the 95th percentile (on the right) are deemed invalid. As shown in the figure, the estimated γ-functions across all scenarios and alternatives maintain a log-linear function trend, being monotonically increasing but with gradually decreasing slopes.

The simulation study results across three scenarios with varying assumptions about the proportion of inside goods suggest that MDCEV-LN has potential to understand satiation effects without relying on handcrafted assumptions about the utility function, extending its applicability to various discrete-continuous choice modeling contexts such as activity time-use, energy consumption and vehicle purchases.

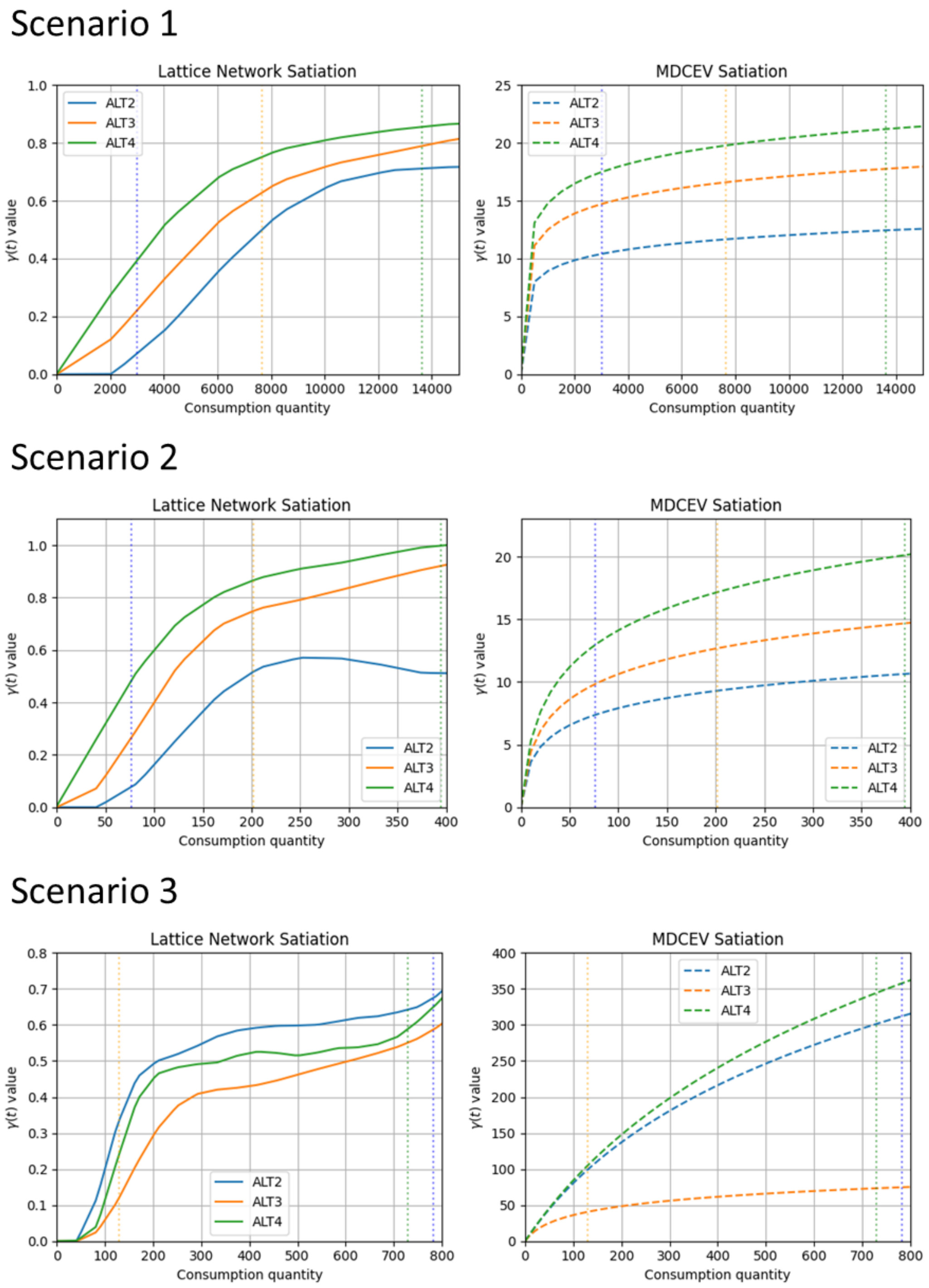

Figure 5: Satiation Effects Estimated by the MDCEV-LN Model

4. CONCLUSIONS

Traditional MDCEV models that rely on hand-crafted utility specifications suffer from issues of inference bias and low predictability. This study proposed MDCEV-LN that estimates MDCEV model's satiation parameters using LN. LN derive flexible yet interpretable utility functions without utility specification by employing piecewise linear functions and multilinear interpolation.

We generated synthetic data following the L_γ-profile and conducted a simulation study to recover the parameters using MDCEV-LN. In the simulation study, MDCEV-LN showed higher predictability compared to the benchmarked original MDCEV and MDCEV-DNN. Also, it efficiently separates baseline marginal utility from satiation effect and captures the log-linear trend of the satiation parameter. MDCEV-LN's high performance across multiple scenarios with varying proportions of inside goods suggests its potential for application in various discrete-continuous choice contexts, including activity time-use, energy consumption, and vehicle purchases.

This study only conducted a simulation study based on the L_γ-profile; however, the flexible yet interpretable characteristics of LN have the potential to outperform existing models on more complex utility specification with non-linear and interaction effects. We are conducting more extensive simulation studies using a non-linear specification in the data generating process to evaluate the potential of MDCEV-LN in capturing the combination of monotonic and non-monotonic relationships between activity duration and satiation effects, which are challenging to model with existing approaches.

ACKNOWLEDGEMENTS

This work is financially supported by Korea Ministry of Land, Infrastructure and Transport (MOLIT) as Innovative Talent Education Program for Smart City, by Basic Science Research Programs through the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) (No.2022R1A2C2012835). Eui-Jin Kim was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No.RS-2023-00246523). The authors used OpenAI's ChatGPT to correct the typos and the grammar of this manuscript. The authors verified the accuracy, validity, and appropriateness of any content generated by the language model.

REFERENCES

Arentze, T. A., Timmermans, H. J. 2004. A learning-based transportation oriented simulation system. *Transportation Research Part B: Methodological*, Vol. 38, No. 7, pp. 613-633.

Bhat, C. R. 2005. A multiple discrete–continuous extreme value model: formulation and application to discretionary time-use decisions. *Transportation Research Part B: Methodological*, Vol. 39, No. 8, pp. 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*, Vol. 42, No. 3, pp. 274-303.

Bhat, C. R., Astroza, S., Bhat, A. C., Nagel, K. 2016. Incorporating a multiple discrete-continuous outcome in the generalized heterogeneous data model: Application to residential self-selection effects analysis in an activity time-use behavior model. *Transportation Research Part B: Methodological*, Vol. 91, pp. 52-76.

Bhat, C. R. 2018. A new flexible multiple discrete–continuous extreme value (MDCEV) choice model. *Transportation Research Part B: Methodological*, Vol. 110, pp. 261-279.

Brier, G. W. 1950. Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, Vol. 78, No. 1, pp. 1-3.

Calastri, C., Hess, S., Daly, A., Carrasco, J. A. 2017. Does the social context help with understanding and predicting the choice of activity type and duration? An application of the Multiple Discrete-Continuous Nested Extreme Value model to activity diary data. *Transportation Research Part A: Policy and Practice*, Vol. 104, pp. 1-20.

Dubin, J. A., McFadden, D. L. 1984. An econometric analysis of residential electric appliance holdings and consumption. *Econometrica: Journal of the Econometric Society*, pp. 345-362.

Han, Y., Pereira, F. C., Ben-Akiva, M., Zegras, C. 2022. A neural-embedded discrete choice model: Learning taste representation with strengthened interpretability. *Transportation Research Part B: Methodological*, Vol. 163, pp. 166-186.

Hanemann, M., Labandeira, X., Labeaga, J. M., Vásquez-Lavín, F. 2024. Discrete-continuous models of residential energy demand: A comprehensive review. *Resource and Energy Economics*, p. 101426.

Hanemann, W. M. 1984. Discrete/continuous models of consumer demand. *Econometrica: Journal of the Econometric Society*, pp. 541-561.

Jian, S., Rashidi, T. H., Dixit, V. 2017. An analysis of carsharing vehicle choice and utilization patterns using multiple discrete-continuous extreme value (MDCEV) models. *Transportation Research Part A: Policy and Practice*, Vol. 103, pp. 362-376.

Kim, E. J., Bansal, P. 2024. A new flexible and partially monotonic discrete choice model. *Transportation Research Part B: Methodological*, Vol. 183, p. 102947.

Kim, J., Allenby, G. M., Rossi, P. E. 2002. Modeling consumer demand for variety. *Marketing Science*, Vol. 21, No. 3, pp. 229-250.

Lipton, Z. C. 2018. The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, Vol. 16, No. 3, pp. 31-57.

Pinjari, A. R., Bhat, C. 2021. Computationally efficient forecasting procedures for Kuhn-Tucker consumer demand model systems: Application to residential energy consumption analysis. *Journal of Choice Modelling*, Vol. 39, p. 100283.

Saxena, S., Pinjari, A. R., Bhat, C. R. 2022. Multiple discrete-continuous choice models with additively separable utility functions and linear utility on outside good: Model properties and characterization of demand functions. *Transportation Research Part B: Methodological*, Vol. 155, pp. 526-557.

Saxena, S., Pinjari, A. R., & Paleti, R. 2022. A multiple discrete-continuous extreme value model with ordered preferences (MDCEV-OP): Modelling framework for episode-level activity participation and time-use analysis. *Transportation Research Part B: Methodological*, Vol. 166, pp. 259-283.

Sifringer, B., Lurkin, V., Alahi, A. 2020. Enhancing discrete choice models with representation learning. *Transportation Research Part B: Methodological*, Vol. 140, pp. 236-261.

Van Cranenburgh, S., Wang, S., Vij, A., Pereira, F., Walker, J. 2022. Choice modelling in the age of machine learning-discussion paper. *Journal of choice modelling*, Vol. 42, p. 100340.

Wang, S., Mo, B., Zhao, J. 2021. Theory-based residual neural networks: A synergy of discrete choice models and deep neural networks. *Transportation Research Part B: Methodological*, Vol. 146, pp. 333-358.

Wang, M., Ye, X. 2024. Development of Multiple Discrete-Continuous Extreme Value Model with Non-Monotonic Utilities: Formulation and Application to Outdoor Non-Mandatory Time-Use Decisions. *Transportation Research Record*, Vol. 2678, No. 9, pp. 655-669.

You, S., Ding, D., Canini, K., Pfeifer, J., Gupta, M. 2017. Deep lattice networks and partial monotonic functions. *Advances in Neural Information Processing Systems*, No. 30.