Leveraging Heteroskedastic Extreme Value Frameworks to Specify Nested Tree Structures in Weibit Choice Models

Doosun Hong[1], Sunghoon Jang*[2]

[1] Postdoctoral Fellow, Department of Civil and Environmental Engineering,

The Hong Kong Polytechnic University, Hong Kong

[2] Assistant Professor, Department of Civil and Environmental Engineering,

The Hong Kong Polytechnic University, Hong Kong

SHORT SUMMARY

This study explores the application of the Heteroscedastic Extreme Value (HEV) framework to specify nesting structures in Weibit choice models. While the HEV framework has effectively defined nested tree structures in Logit choice models, its potential in Weibit models has not yet been investigated. The nested Weibit (NW) model uniquely addresses heterogeneous covariance between alternatives, unlike the nested Logit (NL) model, which maintains fixed covariance. However, establishing nest specifications within the NW model presents an ongoing empirical challenge. By implementing the HEV Weibit (HEVW) model, we leverage its capacity to estimate individual variances, yielding distinct shape parameters for each alternative in a choice set. This variance analysis can unveil tree structures that may not be immediately evident to analysts relying on intuitive configurations. We demonstrate how to specify these nested structures in Weibit choice models based on HEVW model results, supported by empirical findings from London mode choice behavior.

Keywords: Nested Tree Structures, Heteroskedastic Extreme Value Frameworks, Weibit Choice Models, Mode Choice Behavior

1. INTRODUCTION

The Nested Logit (NL) model has long been a basic in choice modeling, offering a solution to the limitations of the Multinomial Logit (MNL) model. The MNL model, while computationally efficient and widely used, assumes that unobserved error terms are independent and identically distributed (Ben-Akiva and Lerman, 1985). This assumption, however, leads to the well-known "Irrelevance of Irrelevant Alternatives" property, which simplifies substitution patterns and often results in unrealistic choice probabilities. The NL model addresses this issue by introducing a hierarchical structure that allows for correlations between alternatives within the same nest. For example, in transportation mode choice, non-motorized travel modes like walking and cycling can be grouped into a single nest to reflect their shared characteristic. This hierarchical nesting structure partially relaxes the independence assumption, making the NL model more robust for applications where alternatives are correlated. Despite its strengths, the NL model relies heavily on the analyst's intuition to define the nesting structure, which can lead to arbitrary groupings. The choice of nests significantly influences the model's performance, as it determines how alternatives are grouped based on shared unobserved factors. To address this challenge, Hensher (1999) proposed the use of the Heteroscedastic Extreme Value (HEV) framework as a search engine for detecting and refining nesting structures. The HEV Logit (HEVL) model estimates the scale parameters of each alternative independently, capturing the heterogeneity in unobserved variances across alternatives (Bhat, 1995). This approach minimizes the subjectivity involved in nesting decisions by clustering alternatives with similar scale parameters. For instance, alternatives with similar variances, such as different types of public transport modes, can be grouped together systematically. This method has been widely applied in empirical behavior studies (Rasciute and Pentecost, 2010; Genius et al., 2012; Yang et al., 2015; Hensher et al., 2015). However, its application is constrained by the NL model's inherent limitation, which is its assumption of identical error distributions within each nest. This restriction imposes fixed covariances between alternatives within nests, which may oversimplify the substitution patterns and reduce the model's flexibility.

The Multinomial Weibit (MNW) model has been proposed as a robust alternative that relaxes the assumption of identical distribution inherent in the Multinomial Logit (MNL) model. In the MNL framework, error terms are assumed to have fixed variance and follow an independent and identically distributed Gumbel distribution. In contrast, the MNW model introduces flexibility by employing a Weibull distribution (Castillo et al., 2008). Similar to the evolution from MNL to Nested Logit (NL) models, the Nested Weibit (NW) model extends the MNW framework by incorporating a hierarchical structure to account for correlations between alternatives (Gu et al., 2022). The NW model provides several key advantages over other choice models, including flexible covariance structures between alternatives within nests, which allows for a more realistic representation of substitution patterns compared to the fixed covariance assumptions found in NL models. Its ability to model both intra-nest and inter-nest variability makes the NW model particularly suitable for complex behavioral contexts, such as multi-modal transport systems, where alternatives exhibit diverse and overlapping attributes. Despite its enhanced flexibility, the NW model faces the critical challenge of defining an appropriate nesting structure, similar to that found in NL models. Although intuition and domain expertise often guide these decisions, they can introduce subjectivity and bias, potentially leading to suboptimal nests. This situation emphasizes the need for systematic, data-driven approaches to defining nests in NW models, an area that remains an open research question despite advancements in modeling frameworks.

Given these challenges, this study explores the application of the Heteroscedastic Extreme Value (HEV) framework to specify nesting structures in Weibull choice models. Similar to Logit choice models, the HEV Weibull (HEVW) model (Jang and Chen, 2025) can be utilized to define nested tree structures in the NW model. While the HEVL-guided approach has been effective in addressing the nesting limitations of the NL model, its efficacy in the context of the NW model remains unexplored. The NW model's flexibility, characterized by flexible covariances between alternatives within the nest, raises questions about the utility of variance-based clustering methods like HEVL guidance. Consequently, the question arises as to whether, and to what extent, the HEV framework can effectively specify these nested structures in the NW model. To address these research issues, this study investigates the applicability of HEV frameworks to specify nested tree structures in Weibit choice models by presenting empirical evidence based on London mode choice data.

2. METHODOLOGY

Heteroskedastic Extreme Value Weibit (HEVW) model

Very recently, Jang and Chen (2025) proposed the HEVW model. Compared to the HEVL model (Bhat, 1995), the HEVW model allows for a more flexible error structure by capturing intra-alternative heteroscedasticity. The model can be expressed as:

$$U_i = V_i \cdot \varepsilon_i \tag{1}$$

where, V_i is the observed deterministic utility of an alternative i, ε_i is error term which is independently distributed following a Weibull distribution.

Based on the assumption that the error for each alternative has a specific shape parameter (δ_i) in a Weibull distribution, choice probabilities can be formulated as:

$$\begin{aligned}
P_i &= prob\left(U_i \geq U_j\right), \text{ for all } j \neq i, j \in C \\
&= prob\left(V_i\varepsilon_i \geq V_j\varepsilon_j\right) = prob\left(\varepsilon_j \geq \frac{V_i}{V_j}\varepsilon_i\right), \text{ for } V < 0, \varepsilon > 0 \\
&= \int_{\varepsilon_i=0}^{\varepsilon_i=+\infty} \left(1 - \Lambda^W\left(\frac{V_i}{V_1} \cdot \varepsilon_i\right)\right) \cdot \left(1 - \Lambda^W\left(\frac{V_i}{V_2} \cdot \varepsilon_i\right)\right) \cdots \left(1 - \Lambda^W\left(\frac{V_i}{V_J} \cdot \varepsilon_i\right)\right) \cdot \lambda^W(\varepsilon_i) d\varepsilon_i \\
&= \int_{\varepsilon_i=0}^{\varepsilon_i=+\infty} \lambda^W(\varepsilon_i) \cdot \prod_{j \neq i \in I} \left[1 - \Lambda^W\left(\frac{V_i}{V_j} \cdot \varepsilon_i\right)\right] d\varepsilon_i
\end{aligned} \tag{2}$$

where, $\lambda(.)$ and $\Lambda(.)$ are probability density function and cumulative density function of Weibull distribution respectively. The choice probability in equation 2 shows an open form. Jang and Chen (2025) showed that it can be approximated using the Gauss–Laguerre quadrature formula. The, the error variance can be defined as:

$$var(\varepsilon_i) = V^2\left[\Gamma\left(1 + \frac{2}{\delta_i}\right) - \Gamma^2\left(1 + \frac{1}{\delta_i}\right)\right] \tag{3}$$

where, Γ is the Gamma function. Therefore, the error variance is dependent on the not only shape parameter but also observed utility.

The HEVW model is particularly effective in scenarios where alternatives exhibit distinct variances due to differences in observed attributes, such as in datasets characterized by wide attribute ranges or complex multi-modal travel. By combining the strengths of the HEVL model with the added flexibility of the Weibull distribution, the HEVW model provides a robust framework for capturing the complexity of choice behavior in real-world applications.

HEVW model serves as a basis for structuring the nesting framework of NW model

The estimation results derived from the Heteroskedastic Extreme Value (HEV) model provide alternative-specific shape parameters, which directly translate into alternative-specific error variances. In line with Hensher's influential 1999 proposal, this flexible framework establishes a robust foundation for formulating and rigorously testing nesting hypotheses within the context of discrete choice modeling. Researchers employ an iterative process: first identifying a comprehensive set of shape parameters from the HEV model estimations. Subsequently, they conduct intuitive groupings of these parameters based on observed similarities. These groupings are not arbitrary; rather, they reflect a deliberate and informed consideration of the shared characteristics and attributes that demonstrably link certain alternatives. This methodical approach significantly reduces the subjectivity and potential biases associated with manually defining nest structures in traditional nested logit models. The resulting hierarchical structure is not merely intuitive but also statistically well-supported, enhancing the credibility and reliability of the subsequent analysis.

The iterative nature of this approach is particularly valuable in mitigating the risks associated with the often-arbitrary groupings found in conventional nested logit models. The HEV-guided process ensures that any resulting nesting hierarchy is grounded in empirically-observed data patterns, enhancing the model's overall validity and predictive power. This is clearly illustrated in mode choice modeling scenarios. For instance, motorized travel modes such as public transportation and driving are frequently treated as a single homogenous nest in standard analyses. However, applying the HEV-based approach often reveals a more refined structure. The HEV-derived shape parameter clusters may indicate distinct sub-nests within the broader category of motorized transport. This refined structure effectively accounts for intra-nest variability, reflecting more accurately the heterogeneity of preferences among these alternatives and enabling a more precise analysis of substitution patterns between specific modes. For example, different types of public transport (bus versus subway) or various driving options (carpooling versus solo driving) may exhibit distinctly different patterns of variance and scale, leading to the identification of distinct sub-nests, which would have been overlooked by an exclusively intuitive nesting approach. This ultimately leads to more accurate estimations of choice probabilities and a deeper understanding of the underlying decision-making processes.

Nested Weibit (NW) model

The NW model is an extension of the NL model, addressing its limitations by incorporating the flexibility of the Weibull distribution (Kitthamkesorn and Chen, 2017). Both models introduce a hierarchical nesting structure to account for correlations between alternatives within nests, relaxing the Independence of Irrelevant Alternatives (IIA) assumption of the MNL model. The NW model relaxes the identically distributed assumption within nests in that Weibull distribution permits alternatives within the same nest to have different error variances. The choice probability in the NW model is expressed as:

$$P_i = P(u) \cdot P(i|u) \tag{3}$$

Where, $P(u)$ is marginal probability of selecting nest of choosing nest u and $P(i|u)$ is conditional probability of choosing alternative i given nest u. Marginal probability P(u) is given by:

$$P(u) = \frac{(V_u)^{\delta} \cdot \left[\sum_{i \in N_u} \left(V_{i|u}\right)^{-\delta_u}\right]^{\frac{\delta}{\delta_u}}}{\sum_{w \in N} (V_w)^{\delta} \cdot \left[\sum_{j \in N_w} \left(V_{j|w}\right)^{-\delta_w}\right]^{\frac{\delta}{\delta_w}}} \tag{4}$$

Where, V_u and $V_{i|u}$ are deterministic components at the nest and alternative levels, respectively. δ_u and δ are shape parameters for the nest and alternative levels. The conditional probability $P(i|u)$ is expressed as:

$$P(i|u) = \frac{\left(V_{i|u}\right)^{-\delta_u}}{\sum_{j \in N_u} \left(V_{j|u}\right)^{-\delta_u}} \tag{5}$$

A key advantage of the NW model is its ability to represent heteroscedasticity, where alternatives within the same nest exhibit different levels of variance in choice behavior. This flexibility stems from the Weibull distribution, which introduces a perception variance as a function of the deterministic utility (Gu et al., 2022). Specifically, the variance of an alternative's error term in the NW model is proportional to its deterministic utility, providing a more realistic representation of decision-making in scenarios with heterogeneous perceptions.

In summary, its ability to accommodate inter-nest correlations and intra-nest heteroscedasticity makes it a powerful tool for applications in transportation and beyond, where alternatives exhibit overlapping and diverse attributes.

3. EMPIRICAL RESULTS

Data (London Passenger Mode Choice)

The London Passenger Mode Choice (LPMC) dataset serves as the empirical basis for this study, supporting the modeling of passenger behavior using the NW model. This publicly available dataset captures the complexity of urban multi-modal transport networks and provides rich attributes that are essential for understanding diverse choice behaviors.

The data originates from three primary sources: the London Travel Demand Survey (LTDS), routing information from the Google Maps Directions API, and cost modeling based on fare and operational cost structures (Hillel et al., 2018). The LTDS provides socio-demographic information, including household size, income, and vehicle ownership, alongside detailed trip-level data such as trip purpose, departure time, duration, and the observed mode. The dataset comprises 81,086 trips recorded between 2012 and 2015, encompassing four primary modes of transport: walking (17.6%), cycling (3.0%), public transport (35.3%), and driving (44.2%). These modes reflect the diverse mobility options in London and their usage patterns in an urban setting. The dataset's detailed representation of multi-modal transport systems makes it particularly suitable for the NW model, which captures correlations between alternatives within nests, and the HEV Weibit model, which empirically detects heteroscedastic variances that guide the formation of these nests.

Estimation Results

The results highlight the limited impact of using the HEV model as a search engine for constructing the nesting structure of NW models. The analysis focused on comparing the model fit (measured by R-squared) under two scenarios: one with intuitive partitions based on analyst-defined groups and the other guided by HEV model-derived shape parameters. A commonly used distinction between motorized and non-motorized modes was employed for intuitive partitions, based on previous research (Eldeeb et al., 2021; Gumz and Török, 2015).

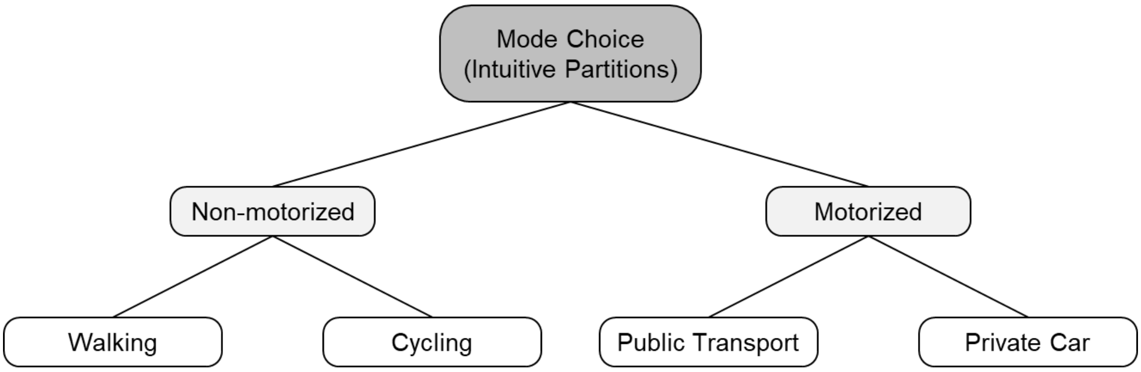

Figure 1: Nest Structure with intuitive partitions (Example)

Table 1: Nested Weibit model results with intuitive partitions

Attributes	**Alternative(s)**	**Parameter (p-value)**
Walking constant	WK	0.000 (fixed)
Cycling constant	CY	-45.922 (0.00)
Public Transport constant	PT	-6.482 (0.00)
Private Car constant	PC	-7.284 (0.00)
Travel time	WK, CY, PT, PC	-1.000 (fixed)
Travel cost	PT, PC	-1.548 (0.00)
Reliability	PC	-21.435 (0.00)
Shape Parameter	WK, CY	0.688 (0.00)
	PT, PC	1.000 (fixed)
Lambda		3.580 (0.00)
Sample size	-	81,086
Rho^2	-	0. 258

* WK: Walking, CY: Cycling, PT: Public Transport, PC: Private Car

Table 2: HEV Weibit model results

Attributes	**Alternative(s)**	**Parameter (p-value)**
Walking constant	WK	0.000 (fixed)
Cycling constant	CY	-43.000 (0.00)
Public Transport constant	PT	-23.860 (0.00)
Private Car constant	PC	-16.621 (0.00)
Travel time	WK, CY, PT, PC	-1.000 (fixed)
Travel cost	PT, PC	-3.073 (0.00)
Reliability	PC	-32.750 (0.00)
Shape Parameter	WK	1.148 (0.00)
	CY	1.000 (fixed)
	PT	3.279 (0.00)
	PC	1.913 (0.00)
Lambda		3.801 (0.00)
Sample size	-	81,086
Rho^2	-	0.341

* WK: Walking, CY: Cycling, PT: Public Transport, PC: Private Car

The HEV-guided nesting structure slightly improved model fit, with R-squared values increasing from 0.258 to 0.262 (**Tables 1** and **3**). The HEV-guided approach established the nesting structure by distinguishing public transport and driving into separate nests, based on the notable differences in their shape parameters identified by the HEV Weibit model. In comparison to the results from the Logit models (where HEV guidance was applied to the Nested Logit model), which showed R-squared values increasing from 0.245 to 0.256, the improvement is smaller in the Weibit models.

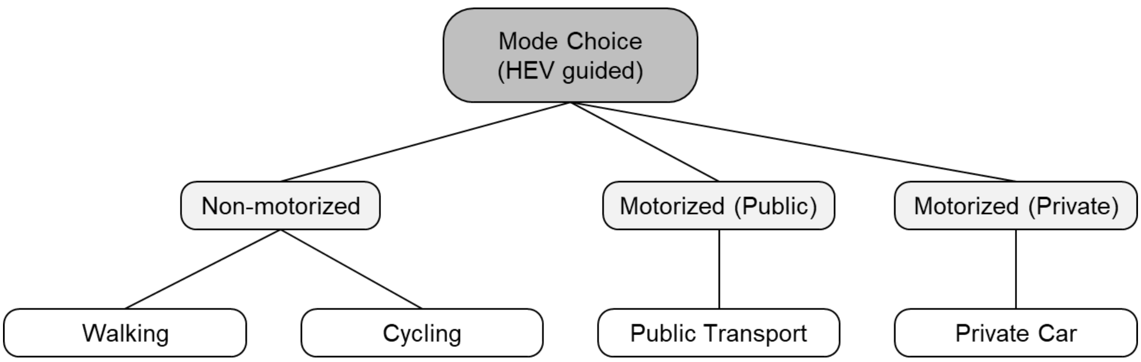

Figure 2: Nest Structure with HEV guide

Table 3: Nested Weibit model results with HEV guide

Attributes	Alternative(s)	Parameter (p-value)
Walking constant	WK	0.000 (fixed)
Cycling constant	CY	-40.794 (0.00)
Public Transport constant	PT	-9.165 (0.00)
Private Car constant	PC	-10.338 (0.00)
Travel time	WK, CY, PT, PC	-1.000 (fixed)
Travel cost	PT, PC	-1.905 (0.00)
Reliability	PC	-27.714 (0.00)
Shape Parameter	WK, CY	1.000 (fixed)
	PT	3.093 (0.00)
	PC	2.041 (0.00)
Lambda		3.234 (0.00)
Sample size	-	81,086
Rho^2	-	0. 262

* WK: Walking, CY: Cycling, PT: Public Transport, PC: Private Car

Unlike the NL model, which assumes identical error distributions within nests and relies heavily on hierarchical structures to improve model performance, the NW model potentially also benefits from its inherent flexibility in accommodating varying covariance structures. This flexibility may enable the model to capture better both intra-nest and inter-nest heterogeneity, which could, in turn, reduce the incremental value of HEV-guided nesting in defining substitution patterns between alternatives. Consequently, while the HEV-guided approach offers a systematic method for identifying potential nests, its impact on improving the overall model fit appears to be modest, possibly due to the intrinsic adaptability of the NW framework.

Nevertheless, a key insight from the HEV-guided approach is the distinction in shape parameters for motorized modes. The HEV Weibit model results suggest notable differences between public transport and driving alternatives, indicating that combining these modes into a single nest may overlook important heterogeneity. HEV-guided clustering could represent a refinement over the intuitive partitioning approach, where motorized modes were grouped together. While these findings imply that separate nests for public transport and driving may better capture observed differences in variability, further empirical validation is needed to confirm the practical significance of this refinement.

CONCLUSIONS

This study investigated the application of the HEV model as a guide for defining nesting structures in the NW model using the LPMC dataset. While the HEV-guided approach slightly improved the model fit (with R-squared increasing from 0.258 to 0.262), the enhancement was modest compared to the results from the Logit models. This outcome highlights the inherent flexibility of the NW model, which already accommodates varying covariance structures and heteroscedasticity, potentially reducing the added value of HEV-guided nesting. Nevertheless, the HEV-guided approach identified significant differences in shape parameters for public transport and driving modes, suggesting that separating these modes into distinct nests better captures the underlying heterogeneity. Despite the modest improvement in overall fit, this refinement highlights the potential of HEV-guided clustering to offer a more data-driven approach for defining nests, although further empirical validation is required to assess its broader applicability.

This paper only shows model fit results from one dataset because of space limits; however, we have much broader findings. The conference presentation will include alternative nesting specifications, analysis of multiple datasets, and a full look at predictive performance.

CKNOWLEDGEMENTS

The work described in this paper was supported by the Hong Kong Polytechnic University (1-BEA2) and Research Grants Council in Hong Kong (F-PPAD). This support is gratefully acknowledged.

REFERENCES

Ben-Akiva, M. E., Lerman, S. R. 1985. *Discrete choice analysis: theory and application to travel demand* (Vol. 9). MIT press.

Bhat, C. R. 1995. A heteroscedastic extreme value model of intercity travel mode choice. *Transportation Research Part B: Methodological*, Vol. 29, No. 6, pp. 471-483.

Castillo, E., Menéndez, J. M., Jiménez, P., Rivas, A. 2008. Closed form expressions for choice probabilities in the Weibull case. *Transportation Research Part B: Methodological*, Vol. 42, No. 4, pp. 373-380.

Eldeeb, G., Mohamed, M., & Páez, A. (2021). Built for active travel? Investigating the contextual effects of the built environment on transportation mode choice. *Journal of Transport Geography*, 96, 103158.

Genius, M., Menegaki, A. N., & Tsagarakis, K. P. (2012). Assessing preferences for wastewater treatment in a rural area using choice experiments. *Water Resources Research*, 48(4).

Gu, Y., Chen, A., & Kitthamkesorn, S. (2022). Weibit choice models: Properties, mode choice application and graphical illustrations. *Journal of Choice Modelling*, Vol. 44, 100373.

Gumz, F., & Török, Á. (2015). Investigation of cordon pricing in Budakeszi. *Periodica Polytechnica Transportation Engineering*, 43(2), 92-97.

Hensher, D. A. (1999). HEV choice models as a search engine for the specification of nested logit tree structures. *Marketing Letters*, 10, 333-343.

Hensher, D. A., Rose, J. M., & Greene, W. H. (2015). *Applied choice analysis*. Cambridge university press.

Jang, S., & Chen, A. (2025). Modeling inter- and intra- alternative heteroscedasticity in trave choice behavior: A heteroscedastic extreme value weibit choice model approach. *Transportation Research Part B: Methodological*, under review.

Kitthamkesorn, S., & Chen, A. (2017). Alternate weibit-based model for assessing green transport systems with combined mode and route travel choices. *Transportation Research Part B: Methodological*, 103, 291-310.

Rasciute, S., & Pentecost, E. J. (2010). A Nested logit approach to modelling the location of foreign direct investment in the Central and Eastern European Countries. *Economic modelling*, 27(1), 32-39.

Yang, C. W., Tsai, M. C., & Chang, C. C. (2015). Investigating the joint choice behavior of intercity transport mode and high-speed rail cabin with a strategy map. *Journal of Advanced Transportation*, 49(3), 297-308.