Communications in Transportation Research 4 (2024) 100121

Contents lists available at ScienceDirect

Communications in Transportation Research

journal homepage: www.journals.elsevier.com/communications-in-transportation-research

Full Length Article

A binary choice model for adoption of an emerging travel mode with unique service features

Yu Gu [a], Anthony Chen [a,b,*], Sunghoon Jang [a], Songyot Kitthamkesorn [c]

[a] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong, 999077, China
[b] The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen, 518000, China
[c] Excellence Center in Infrastructure Technology and Transportation Engineering Department of Civil Engineering, School of Engineering Chiang Mai University, Chiang Mai, 50200, Thailand

ARTICLE INFO

Keywords:
Binary mode choice
Emerging mode
Multiplicative random utility model
Weibull distribution
Future transportation system

ABSTRACT

In the era of emerging technologies, the transportation system is witnessing the introduction of innovative mobility services, such as autonomous vehicles, which possess unique service features that cannot be seen from conventional travel modes. To facilitate the understanding of the behavioral impacts and the adoption of innovative mobilities, a novel binary weibit model with an oddball alternative (BW-O) is developed for the binary choice between conventional and emerging mobilities. The BW-O model explicitly considers the unprecedented (or unique) service features of emerging travel modes while retaining the closed-form choice probability. This study empirically illustrates the application of the BW-O model in the mode choice context. The desirable properties of the BW-O model compared to the existing binary choice models are discussed both theoretically and empirically. In the binary mode choice problem with an emerging travel mode, the unique service features of the emerging mode can lead to the "oddball" effect and "superstar" effect, which play a critical role in the travel behavior and mode adoption. The BW-O model inherently captures both effects by considering a higher perception variance for the emerging mode and asymmetric choice probabilities between different modes. Thus, as revealed by the empirical results, the BW-O model outperforms the basic binary weibit model in terms of both model fit and predictive power. The developed BW-O model is not only applicable to the mode choice problem in transportation systems, but also opens a door for more general class-imbalanced binary choice contexts where an alternative has additional attractiveness and asymmetric choice probability.

1. Introduction

In the era of emerging technology, a variety of innovative transport modes such as the autonomous vehicles (AVs) have been introduced to the transportation system. With the improved transportation service, these emerging modes are expected to be gradually adopted through the competition with conventional modes (Aramrattana and Fu, 2022; Fagnant and Kockelman, 2015; Gu and Chen, 2023; Jansuwan et al., 2021; Olovsson et al., 2022). Compared with conventional modes, the emerging technologies often provide some unique service features, such as the safety concern and autonomous driving of AVs. These features will reduce travel disutility but bring additional subjective uncertainty as travelers have not experienced them in the current services. Aramrattana and Fu (2022) investigated travelers' behavioral adaptations to AVs, which suggests that the introduction of AVs exerts significant but heterogenous impacts on the travel behavior. This may significantly change how travelers perceive the emerging mode and compare it with the conventional ones and hence influence the mode choice behavior (Acharya and Mekker, 2022; Dubey et al., 2022; Song, 2019). Few attempts have been made to specifically consider the effect of unprecedented service features in the forecast of the adoption of emerging modes (Bansal et al., 2021; Haboucha et al., 2017; Jang et al., 2021; Jiang et al., 2019; Ortúzar, 2021). For instance, Dubey et al. (2022) developed an elegant but complicated binary choice modeling framework that is effective to model the additional uncertainty related to AVs but is hungry for the word-of-mouth information, the collection of which requires high cost in terms of money, time, and human resources. To facilitate forecasting the adoption rate of emerging modes, it is imperative to develop advanced and easy-to-implement binary choice models for understanding the choice behavior between conventional and emerging modes.

* Corresponding author. Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong, 999077, China.
E-mail address: anthony.chen@polyu.edu.hk (A. Chen).

https://doi.org/10.1016/j.commtr.2024.100121
Received 4 December 2023; Accepted 18 December 2023
Available online xxxx

Some existing discrete choice models are applicable to model the choice between emerging and conventional modes considering specific features of the binary mode choice. For example, the emerging mode is likely to be preferred and have a rapidly increasing adoption rate during the transition period, which can be driven by the desire for innovation and the growing social influence of novel services (Song, 2019). This phenomenon can be modeled as the "superstar" effect (Chorus, 2018), i.e., the "superstar" alternative attracts much more demand than other alternatives even when its advantage in quality is not as significant. Chorus (2018) modeled the "superstar" effect by transforming systematic utility of the additive random utility model (ARUM) to enlarge the difference in satisfaction. Brathwaite and Walker (2018) developed a series of logit-type binary choice models with asymmetric and closed-form choice probability functions, which are applicable to the class-imbalanced choice between alternatives with distinct demands. However, these models mainly focus on modifying the deterministic part of utility function or the aggregate choice probability expression, while the additional subjective uncertainty arising from the emerging service features is not explicitly modeled.

The additional subjective uncertainty of emerging mode can be modeled as the "oddball" effect, i.e., an oddball alternative has unique attributes that cannot be observed from other regular alternatives in the choice set. Recker (1995) developed a logit-based ARUM to address the "oddball" effect, where a Gumbel distributed random component is assumed for the unique attributes in addition to that assumed for the common attributes shared by all the alternatives. Thus, the oddball alternative has a larger perception variance, which can reflect the additional subjective uncertainty. The logit-based oddball choice model also has the potential to capture the "superstar" effect owing to the asymmetric probability functions and the higher choice probability of the oddball alternative. However, the Gumbel distributed assumption embedded in the logit model and the additive utility function used in the ARUM led to fixed perception variances for both regular and oddball alternatives. This makes the logit-based oddball choice model inadequate to reflect the heterogenous perceptions of service quality provided by different modes or different service features.

This study investigates the adoption of emerging travel mode through a multiplicative random utility model (MRUM) based on the Weibull distribution, which considers the emerging mode as an oddball alternative with unique attributes. The developed weibit oddball choice model retains the closed-form probability expression, which facilitates the model implementation via efficient evaluation and exact estimation solution and guarantees high interpretability of model outcomes. The developed model is applied to investigate the adoption of AVs against the conventional human-driven vehicles (HDVs). The empirical results indicate the benefit of the developed model to address both the "oddball" and "superstar" effects in the binary choice context with an emerging mode like AV. Specifically, the multiplicative disutility function embedded in the model allows disutility-dependent perception variances, which enables considering the heterogeneity in service quality perceptions (Fosgerau and Bierlaire, 2009; Kitthamkesorn and Chen, 2013). On this basis, the developed model can effectively capture the "oddball" effect owing to its flexibility to inherently reflect heterogeneous subjective uncertainties associated with different modes and different service features. Furthermore, the asymmetric choice probability of the developed model facilitates to capture the "superstar" effect, which is applicable to the class-imbalanced choice context between emerging and conventional modes.

The remainder of this paper is organized as follows. Section 2 presents the formulation and properties of the proposed binary mode choice model. Empirical experiments are conducted in Section 3 to verify the applicability of the proposed model for estimation and prediction of the adoption rate of the emerging AVs. Section 4 presents concluding remarks and some directions for future research.

2. Binary choice model between emerging and conventional alternatives

2.1. Binary oddball weibit model formulation

This section presents the binary weibit model with an oddball alternative (BW–O) for the choice between conventional and emerging modes. As an example, we consider the binary choice between a conventional mode (labeled as mode 1 hereafter) and an emerging mode (labeled as mode 2 hereafter). The two modes share a set of common attributes $\overline{\boldsymbol{\tau}}$, while the emerging mode has a set of unique attributes $\tilde{\boldsymbol{\tau}}$, which indicates unprecedented service features and is assumed as an independent random component (Recker, 1995). Consistent with weibit choice models (Fosgerau and Bierlaire, 2009), the proposed BW-O model has multiplicative disutility functions for both alternatives. The perceived disutility of conventional mode V_1 is represented as

$$V_1 = v_1 \cdot \varepsilon_1 \tag{1}$$

where v_1 and ε_1 are the systematic disutility and random error of conventional mode 1, respectively. v_1 is obtained based on the common attributes:

$$v_1 = \sum_{i \in \boldsymbol{I}} \omega^i \cdot \overline{\tau}_1^i \tag{2}$$

where $\boldsymbol{I}$ is the set of common attributes, $\overline{\tau}_1^i$ and ω^i denote the level of attribute i of mode 1 and the coefficient of attribute i, respectively.

To account for the "oddball" effect of the emerging travel mode, the disutility function of mode 2 is constructed following the development of the oddball logit model (Recker, 1995), which includes an additional random error term with respect to the unique attributes. Recently, Gu et al. (2024) extended the oddball logit model to the oddball weibit model using a multiplicative error structure, where the disutility function of the oddball alternative is constructed by multiplying the perceived disutility of common attributes ($\overline{V}_2 = \overline{v}_2 \cdot \overline{\varepsilon}_2$) with that of the unique attributes ($\tilde{V}_2 = \tilde{v}_2 \cdot \tilde{\varepsilon}_2$). The disutility function is then expressed as Eq. (3):

$$V_2 = \overline{V}_2 \cdot \tilde{V}_2 = (\overline{v}_2 \cdot \overline{\varepsilon}_2) \cdot (\tilde{v}_2 \cdot \tilde{\varepsilon}_2) = v_2 \cdot \zeta_2 \tag{3}$$

where $\overline{\varepsilon}_2$ and $\tilde{\varepsilon}_2$ denote the random errors; $\overline{v}_2$ and $\tilde{v}_2$ are the systematic disutility. Let $v_2 = \overline{v}_2 \cdot \tilde{v}_2$ and $\zeta_2 = \overline{\varepsilon}_2 \cdot \tilde{\varepsilon}_2$ denote the total systematic disutility and random error of the emerging mode, respectively. $\overline{v}_2$ and $\tilde{v}_2$ are obtained based on the common attributes and unique attributes, respectively:

$$\begin{aligned} \overline{v}_2 &= \sum_{i \in \boldsymbol{I}} \omega^i \cdot \overline{\tau}_2^i \\ \tilde{v}_2 &= \sum_{j \in \boldsymbol{J}} \omega^j \cdot \tilde{\tau}_2^j \end{aligned} \tag{4}$$

where $\boldsymbol{J}$ is the set of unique attributes. The random error terms used in the BW-O model, ε_1, $\overline{\varepsilon}_2$, and $\tilde{\varepsilon}_2$, are assumed to independently and identically follow the Weibull distribution (λ, α, β). λ and α are the location parameter and the scale parameter, which are set as $\lambda = 0$ and $\alpha = 1$ for simplicity (Kitthamkesorn and Chen, 2013), respectively. This study focuses on the shape parameter β, which implies the level of dispersion.

Following the principle of disutility minimization, the choice probability of the emerging mode P_2 is equivalent to the probability that the emerging travel mode 2 has a lower disutility than the conventional mode 1, which can be expressed as

$$P_2 = P(v_2 \cdot \zeta_2 \leq v_1 \cdot \varepsilon_1) = P\left(\frac{\zeta_2}{\varepsilon_1} \leq \frac{v_1}{v_2}\right) \tag{5}$$

where $P(\cdot)$ denotes the probability function.

Therefore, the choice probability can be obtained based on the cumulative distribution function (CDF) of random variable $\frac{\zeta_2}{\varepsilon_1}$. Based on Proposition 1 presented below, Eq. (5) can be written as

$$P_2 = F_Z\left(\frac{v_1}{v_2}\right) = \left(\frac{v_1}{v_2}\right)^{\beta} \cdot \exp\left[\left(\frac{v_1}{v_2}\right)^{\beta}\right] \cdot E_1\left[\left(\frac{v_1}{v_2}\right)^{\beta}\right] = \varphi \cdot e^{\varphi} \cdot E_1(\varphi) \tag{6}$$

where $F_Z(\cdot)$ denotes the CDF of $\frac{\zeta_2}{\varepsilon_1}$. $\varphi = \frac{(v_2)^{-\beta}}{(v_1)^{-\beta}}$. $E_1(x) = \int_x^{+\infty} \frac{e^{-x}}{x} dx$ is the exponential integral. The choice probability of conventional mode 1 can then be obtained as Eq. (7):

$$P_1 = 1 - P_2 = 1 - \varphi \cdot e^{\varphi} \cdot E_1(\varphi) \tag{7}$$

Proposition 1. The CDF of the quotient between the random variables, $Z = \zeta_2/\varepsilon_1$, can be expressed as

$$F_Z(z) = z^{\beta} \cdot e^{z^{\beta}} \cdot E_1\left(z^{\beta}\right) \tag{8}$$

Proof. From Eqs. (1) and (3), the quotient between the two random error terms can be expressed as $\frac{\zeta_2}{\varepsilon_1} = \frac{\bar{\varepsilon}_2 \cdot \tilde{\varepsilon}_2}{\varepsilon_1} = \left(\frac{\bar{\varepsilon}_2}{\varepsilon_1}\right) \cdot \tilde{\varepsilon}_2$. Given the property of the Weibull distribution (Gu et al., 2022), the first term on the right-hand side, $\frac{\bar{\varepsilon}_2}{\varepsilon_1}$, is the ratio between two independently and identically distributed (IID) Weibull variables, which follows the Log-logistic distribution $(1, \beta)$; while the second term on the right-hand side, $\tilde{\varepsilon}_2$, follows the Weibull distribution $(0,1,\beta)$. Thus, random variable $\frac{\zeta_2}{\varepsilon_1}$ can be expressed as the product of a Log-logistically distributed variable and a Weibull distributed variable.

Now consider two random variables X following the Log-logistic distribution $(1, \beta)$ and Y following the Weibull distribution $(0, 1, \beta)$, their product is $Z = X \cdot Y$. Hence, Y can be expressed based on Z and X as $Y = Z/X$. The probability density function (PDF) of random variable Z can be expressed as Eq. (9):

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) \cdot \left|\frac{1}{x}\right| \cdot f_Y\left(\frac{z}{x}\right) dx = \int_0^{+\infty} \beta x^{\beta-1} \exp\left(-x^{\beta}\right) \cdot \frac{1}{x} \cdot \frac{\beta (z/x)^{\beta-1}}{\left[1 + (z/x)^{\beta}\right]^2} dx = \beta z^{\beta-1} \cdot \int_0^{+\infty} \beta x^{\beta-1} \exp\left(-x^{\beta}\right) \cdot \frac{1}{x^{\beta}} \cdot \frac{1}{\left[1 + (z/x)^{\beta}\right]^2} dx \tag{9}$$

Let $u = x^{\beta}$, Eq. (9) can be expressed as

$$f_Z(z) = \beta z^{\beta-1} \cdot \int_0^{+\infty} u \cdot A(u) du \tag{10}$$

where $A(u) = e^{-u} \cdot \frac{1}{[u+z^{\beta}]^2}$. Let $v = u + z^{\beta}$, the integration of $A(u)$ can be expressed as (Gradshteyn and Ryzhik, 2007):

$$\int_0^{+\infty} A(u) du = \int_{z^{\beta}}^{+\infty} e^{-\left(v - z^{\beta}\right)} \cdot \frac{1}{v^2} dv = e^{z^{\beta}} \left[-\frac{e^{-v}}{v} + E_1(v)\right]\Bigg|_{z^{\beta}}^{+\infty} \tag{11}$$

Taking Eq. (11) into Eq. (10) and using integration by parts, the PDF of random variable Z can be expressed as

$$\begin{aligned} f_Z(z) &= \beta z^{\beta-1} \cdot \left\{\left[\left(v - z^{\beta}\right) \cdot e^{z^{\beta}} \left(-\frac{e^{-v}}{v} + E_1(v)\right)\right]\Bigg|_{z^{\beta}}^{+\infty} - e^{z^{\beta}} \cdot \int_{z^{\beta}}^{+\infty} \left[-\frac{e^{-v}}{v} + E_1(v)\right] dv\right\} = \beta z^{\beta-1} \cdot \left\{0 + e^{z^{\beta}} \cdot E_1\left(z^{\beta}\right) - \left[e^{z} \cdot v E_1(v) - e^{z} \cdot e^{-v}\right]\Big|_{z^{\beta}}^{+\infty}\right\} \\ &= \beta z^{\beta-1} \cdot \left[e^{z^{\beta}} \cdot E_1\left(z^{\beta}\right) + e^{z^{\beta}} \cdot z^{\beta} E_1\left(z^{\beta}\right) - 1\right] \end{aligned} \tag{12}$$

Integrating the PDF shown in Eq. (12) leads to Eq. (8). Thus, the product of the Log-logistically distributed variable $\frac{\zeta_2}{\varepsilon_1}$ and the Weibull distributed variable $\tilde{\varepsilon}_2$, which leads to the variable ζ_2/ε_1 focused by Proposition 1, has the CDF shown in Eq. (8). This completes the proof.

2.2. Model properties

2.2.1. Characterization of different alternatives

This section illustrates the random components considered in the BW-O model through comparison with a logit-based "superstar" effect model based on transforming systematic disutility $v_i' = 25 \cdot v_i$ (Chorus, 2018). As shown in Fig. 1(a), the transformation of systematic utility enlarges the difference in satisfaction but does not influence the random error distribution of each alternative. This implies identical perception variances for both "superstar" and regular alternatives, which may be inadequate to capture the distinct scales of enlarged systematic disutility and the additional uncertainty associated with the unprecedented features of emerging mode.

On the other hand, the developed BW-O model is able to capture both the "oddball" effect and the "superstar" effect by introducing an independent random component for the emerging mode. Fig. 1(b) compares the PDFs of the random components related to emerging and conventional modes with the shape parameter $\beta = 3.7$. Compared with ε_1 following the Weibull distribution, the distribution of ζ_2 is the product of two IID Weibull distribution, whose PDF curve is more right-skewed and has a heavier tail. The changed shape of PDF curve captures the reduced disutility perception and higher uncertainty of the oddball alternative. These characteristics imply the BW-O model is suitable to explicitly account for unique service features of the emerging mode in terms of higher service quality and higher subjective uncertainty.

2.2.2. Perception variances

This section presents the perception variance of the BW-O model, which indicates the subjective uncertainty associated with conventional and emerging modes. The conventional mode shares the same perception variance as in the binary weibit (BW) model, which can be expressed as (Castillo et al., 2008):

$$D(V_1) = E^2(V_1) \cdot \left[\frac{\Gamma(1 + 2/\beta)}{\Gamma^2(1 + 1/\beta)} - 1\right] \tag{13}$$

where $E(V_k)$ denotes the mean disutility of alternative k, $\Gamma(\cdot)$ is the Gamma function.

Based on the assumption that $\bar{\varepsilon}_2$, and $\tilde{\varepsilon}_2$ are independently Weibull distributed with the same shape parameter β, the perception variance of the emerging mode is

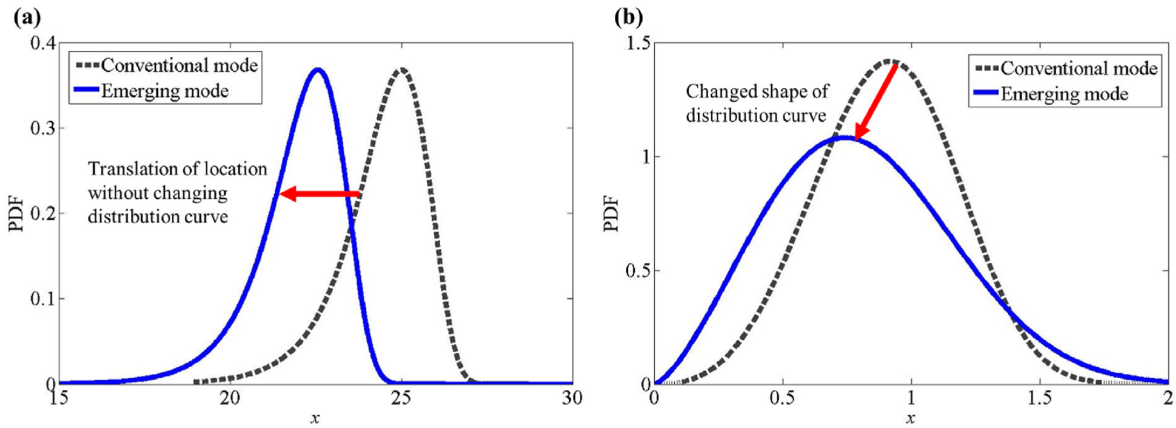

Fig. 1. Comparison of PDFs of random components considered in (a) "superstar" effect model and (b) proposed BW-O model.

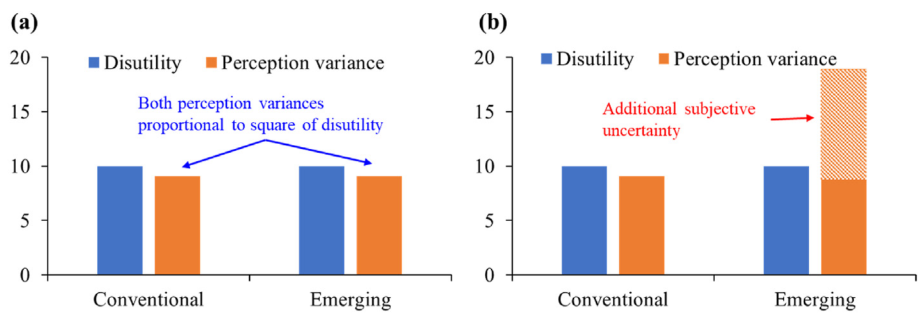

Fig. 2. Comparison of the perception variance between (a) BW model and (b) BW-O model.

$$
\begin{aligned}
D(V_2) &= D(\overline{V}_2 \cdot \tilde{V}_2) = E(\overline{V}_2{}^2) \cdot E(\tilde{V}_2{}^2) - E^2(\overline{V}_2) \cdot E^2(\tilde{V}_2) \\
&= E^2(V_2) \cdot \left[\left(\frac{\Gamma(1+2/\beta)}{\Gamma^2(1+1/\beta)} \right)^2 - 1 \right]
\end{aligned} \tag{14}
$$

Remark. The perception variances of both modes are proportional to square of mean disutility. The proportionalities are dependent on the shape parameter β. Different from the BW model where both alternatives share the same proportionality, the BW-O model allows the oddball alternative to have a larger proportionality than the regular alternative, which indicates the higher subjective uncertainty associated with the emerging mode.

Fig. 2 compares the perception variances in the BW and BW-O models based on an example where conventional and emerging modes have the same disutility of 10 and shape parameter $\beta = 3.7$. The BW model fails to distinguish the different uncertainties associated with different modes (Fig. 2(a)). On the other hand, the BW-O model allows for additional subjective uncertainty for the emerging mode owing to the independent random error assumed for the unprecedented service features (Fig. 2(b)).

2.2.3. Evaluation of the emerging mode choice probability

This section examines the proposed model in terms of its evaluation of the emerging mode choice probability. The BW-O model is compared with the widely used binary logit (BL) and BW models based on a binary choice example, where the disutility of conventional mode is fixed at 5 and the disutility of emerging mode varies from 0 to 10. The shape parameter of BW-O and BW models is $\beta = 3.7$, the scale parameter of BL model is $\theta = 1$.

As shown in Fig. 3(a), different from the symmetric BL model, the BW-O model has an asymmetric choice probability function. The asymmetry property is desirable for the class-imbalanced choice contexts, where the preferred alternative tends to gain a larger increase/decrease

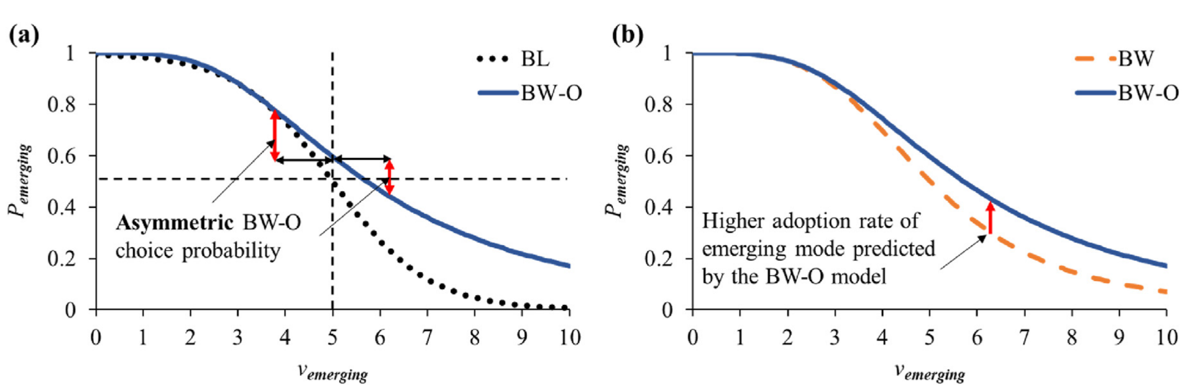

Fig. 3. Comparison of the binary response curves between (a) the BL and BW-O models and (b) the BW and BW-O models.

in choice probability than the under-represented one even under an equal decrease/increase in disutility (Brathwaite and Walker, 2018). This property also implies that the proposed model has the potential to capture the "superstar" effect, as the adoption rate of emerging mode may experience a more rapid increase during the transition to future transportation systems (Song, 2019). Furthermore, the emerging mode tends to have a higher choice probability in the BW-O model than in the BW model (Fig. 3(b)), which is consistent with the "oddball" effect captured by the logit-based oddball choice model (Recker, 1995).

3. Empirical experiments

In this section, we apply synthetic data analysis to examine the performance of the proposed model and generalize the comparative analysis for a population who experience additional subjective uncertainty to the emerging modes.

3.1. Synthetic data generation

Following the classical methodology to generate synthetic data (Williams and Ortúzar, 1982), we generated 200,000 independent observations. The data represents the consumer purchasing choice between HDVs and AVs. Whereas the conventional HDV was described by purchase cost and trip cost, the emerging AV was described by not only purchase cost and trip cost but also penetration rate and parking cost reduction rate reflecting the additional subjective uncertainty (Haboucha et al., 2017; Jiang et al., 2019). Therefore, the AV is deemed as an oddball option in this choice context. The attribute levels were built with random draws from independent truncated normal distribution functions with arbitrary lower and upper bounds based on mean of the levels. To moderate the effect of randomness, we generated 10 normalized random draws, and used the average value, following Jang et al. (2017). Detailed information of attribute level and taste weights are presented in Table 1. Everyone is assumed to behave to minimize disutility following the proposed models.

3.2. Estimation results

The estimation results of both BW and BW-O models are presented in Table 2. Due to the identification issue of shape parameter, the parameter for trip cost was fixed, following Fosgerau and Bierlaire (2009). All parameter estimates are statistically significant at the 95% level. The results show that the BW-O model results in better model fit than BW model in terms of the Bayesian Information Criterion (BIC). Also, the result from the Likelihood Ratio (LR) test indicates that the BW-O model is preferable at the 95% level to the BW model. These results may be rooted in predicted market shares.

Table 3 shows how much the models predicted choice probability for both alternatives. The simulated data indicates that 75.32% of the whole observations choose the conventional vehicle, while 24.68% choose the AV. The BW model over-predicts the market share for conventional vehicles (75.49%) and under-predicts it for AVs (24.51%). On the other hand, the BW-O model shows predictions very similar to the actual observations. In terms of Root Mean Square Error (RMSE), it is only 0.0005. There was only one case where choice was predicted to differ from observation. Compared to the predictions by the BW model, the BW-O model shows higher probability for the emerging alternative (AV), reflecting the 'superstar' effect (Chorus, 2018) consistent with the discussions in Section 2.2.3.

In addition, the BW-O model better recovers the "true" parameters defined to build the simulated data. The results of the *t*-test with a null hypothesis that the parameter estimates in the BW-O model are equal to the "true" parameters indicate that null hypothesis cannot be rejected in all cases at a 95% level of statistical significance (Table 4). On the other hand, the test for the BW model reveals that the null hypothesis is rejected in all cases except for the parameter for the purchase cost.

Table 1
Parameter definition and attribute levels.

Attribute	Taste weight	Alternative		
		Conventional vehicle	Autonomous vehicle	Change[a]
Purchase cost (1,000 USD)	−0.500	30	40	±10
Trip cost (per commuting direction) (USD)	−1.000	1.50	1.25	±1.0
Penetration rate (%)	−0.100	—	10	±20
Discount in insurance (%)	−0.100	—	20	±10
Shape parameter for Weibull distribution	1.500	—	—	—

[a] It indicates that the changes in the explanatory attributes are distributed in accordance with the truncated normal distribution.

Table 2
Estimation results of synthetic data (*t*-values in bracket).

Attribute	Model	
	BW	BW-O
Purchase cost	−0.491 (−9.35)	−0.699 (−6.85)
Trip cost	−1.000 (fixed)	−1.000 (fixed)
Additional cost for autonomous devices	−0.299 (−6.22)	−0.097 (−27.21)
Increase rate in driving insurance	−0.707 (−8.98)	−0.099 (−32.97)
Shape parameter for Weibull distribution	1.335 (54.83)	1.467 (60.68)
Model fit	—	—
Final log-likelihood	−109,923.0	−109,594.8
BIC	219,894.8	219,238.5

Table 3
Simulated and predicted choice probabilities.

	Predicted choice probability		
	Simulated choice probability	BW	BW-O
Conventional vehicle	75.32% (150,643 observations)	75.49% (150,983 observations)	75.32% (150,644 observations)
Autonomous vehicle	24.68% (49,357 observations)	24.51% (49,017 observations)	24.68% (49,356 observations)
RMSE[a]	—	1.0752	0.0032

[a] RMSE was calculated based on the market share.

Table 4
The *t*-test results.

t-value	BW	BW-O
Purchase cost	0.17	−1.94
Additional cost for autonomous devices	−4.14	0.82
Increase rate in driving insurance	−7.71	0.44
Shape parameter for Weibull distribution	−6.78	0.11

Table 5
Predictive ability of the different choice models on a hold-out sample.

	BIC	Correct choice rate
BW	40,665.26	98.36%
BW-O	31,394.42	99.97%

3.3. Validation results

To provide insight into the predictive power of the proposed BW-O model, we conducted the hold-out test. 80% of observations were used to estimate the model, and the hold-out samples, which is 20% of observations were used to validate the model. Table 5 shows the results of the hold-out test. First, consistent with the estimation results of whole observations based on the BIC, the BW-O model outperforms the BW model with respect to the performance on a hold-out sample. Second, the BW-O model has a higher correct choice rate. The BW-O model predicts choice identically to the simulated observations in all but 11 cases of the 40,000 observations (99.97%). In the case of the BW model, it predicts wrong choice, inconsistent with the simulated observations, in 457 cases; therefore, the correct choice rate is 98.36%.

3.4. Result discussion

In both estimation and validation tests, the BW-O model is consistently superior to the BW model in terms of model fit, parameter estimation, model performance, and prediction correctness. This is consistent with the theoretical advantages of the BW-O model illustrated in Section 2, which indicates its ability to account for the unique attributes of the emerging AV mode. The results also imply the importance of addressing both the "oddball" and "superstar" effects in the choice contexts with emerging mobilities that have unprecedented service features. The performance of the developed BW-O model is also verified in the case study based on a real-world mode choice data set as presented in Appendix A, which shows similar results to those obtained from the synthetic dataset.

4. Conclusions

This study develops a multiplicative random utility model focusing on the binary choice between a regular and an oddball alternative. The proposed model serves as a simplified and effective alternative to forecast the adoption rates of emerging modes during the transition from current to future transportation systems. The heterogeneous perceptions of travel disutility can be inherently considered via the multiplicative disutility function based on the Weibull distribution. The additional subjective uncertainty associated with the unprecedented service features of the emerging mode is explicitly modeled, while the closed-form probability expression is retained. Empirical experiments based on a synthetic data set are conducted to show the superiority of the proposed BW-O model compared with the commonly used multiplicative random utility (BW) model. The results indicate that the proposed model provides better model fit and better predictive power when applied to the binary choice context between AVs and HDVs.

Based on the proposed model, there are several potential directions for future studies: (1) extend the proposed model to consider more than one conventional and/or emerging modes; (2) integrate the proposed model in the optimization of infrastructure planning, service design and policy making for promoting emerging technologies; and (3) apply the proposed model to other non-transportation choice contexts with class-imbalanced choice sets, such as tourism destination choice, residential location choice, and shoppers' brand choice (Brathwaite and Walker, 2018; Chorus, 2018).

Replication and data sharing

The data used in this paper can be downloaded from https://transp-or.epfl.ch/pythonbiogeme/examples_swissmetro.html. The software used in this paper can be accessed at https://www.apollochoicemodelling.com/and downloaded for use.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The work is partially funded by the National Natural Science Foundation of China (72071174), the Research Grants Council of the Hong Kong Special Administrative Region (PolyU 15222221 and PolyU 15221922), the Smart Cities Research Institute (CDA9) and the Research Institute of Land and Space (1-CD7N) at the Hong Kong Polytechnic University, and the Chiang Mai University in Thailand. Their support is gratefully acknowledged.

Appendix A. Empirical experiment based on Swissmetro dataset

To verify the applicability of the developed BW-O model in real-world cases, we conduct an experiment based on the Swissmetro data set (Bierlaire et al., 2001). The dataset comprises the stated preference survey data in the mode choice situation with an innovative Swissmetro service. This experiment focuses on the binary mode choice between the conventional train and the innovative Swissmentro for transit passengers that do not have access to car. The number of observations is 1,085. As shown in Table A1, the innovative Swissmetro mode is considered as an oddball alternative with unique attributes (i.e., headway and availability of airline seats).

Table A1
Attributes of travel modes.

Attribute	Travel mode	
	Train	Swissmetro
Common attribute	Train travel time (min)	Swissmetro travel time (min)
	Train travel cost (Swiss franc)	Swissmetro travel cost (Swiss franc)
Unique attribute	—	Headway (min)
	—	Seat configuration

In model estimation, the coefficient of travel cost is normalized to minus unity and the shape parameters of the weibit-based models are estimated (Fosgerau and Bierlaire, 2009). The estimation results from the BW and BW-O models are presented in Table A2.

Table A2
Estimation Results of Swissmetro data (*t*-values in bracket).

Attribute	Model	
	BW	BW-O
Travel time	−6.72 (−3.58)	−2.10 (−2.06)
Travel cost	−1.000 (fixed)	−1.000 (fixed)
Frequency	11.32 (1.85)	0.4 (13.37)
Seat configuration	59.44 (1.74)	0.27 (1.84)
Shape parameter for Weibull distribution	2.65 (10.08)	1.95 (11.15)
Model fit		
Final log-likelihood	−512.75	−505.84
AIC	1033.51	1019.68
BIC	1053.38	1039.56

The results show that the developed BW-O model has a higher log-likelihood and lower values of AIC and BIC, which are similar to those from the synthetic data set described in Section 3. The model comparison results demonstrate that the BW-O model has a clear advantage compared to the basic BW model, which can be attributed to the capability of simultaneously capturing both the "oddball" and "superstar" effects.

References

Acharya, S., Mekker, M., 2022. Importance of the reputation of data manager in the acceptance of connected vehicles. Commun. Transport. Res. 2, 100053.

Aramrattana, M., Fu, J., 2022. Behavioral adaptation of drivers when driving among automated vehicles. J. Intell. Connect. Veh. 5, 309–315.

Bansal, P., Kumar, R.R., Raj, A., Dubey, S., Graham, D.J., 2021. Willingness to pay and attitudinal preferences of Indian consumers for electric vehicles. Energy Econ. 100, 105340.

Bierlaire, M., Axhausen, K., Georg, A., 2001. The acceptance of modal innovation: the case of Swissmetro. In: the 1st Swiss Transport Research Conference, pp. 1–15.

Brathwaite, T., Walker, J.L., 2018. Asymmetric, closed-form, finite-parameter models of multinomial choice. J. Choice Model 29, 78–112.

Castillo, E., Menéndez, J.M., Jiménez, P., Rivas, A., 2008. Closed form expressions for choice probabilities in the Weibull case. Transp. Res. Part B Methodol. 42, 373–380.

Chorus, C.G., 2018. Paving the way towards superstar destinations: models of convex demand for quality. Environ. Plan. B Urban Anal. City Sci. 45, 161–179.

Dubey, S., Sharma, I., Mishra, S., Cats, O., Bansal, P., 2022. A general framework to forecast the adoption of novel products: a case of autonomous vehicles. Transp. Res. Part B Methodol. 165, 63–95.

Fagnant, D.J., Kockelman, K., 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. Transp. Res. Part A Policy Pract. 77, 167–181.

Fosgerau, M., Bierlaire, M., 2009. Discrete choice models with multiplicative error terms. Transp. Res. Part B Methodol. 43, 494–505.

Gradshteyn, I.S., Ryzhik, I.M., 2007. Table of Integrals, Series, and Products. Academic Press, San Diego, USA.

Gu, Y., Chen, A., 2023. Modeling mode choice of customized bus services with loyalty subscription schemes in multi-modal transportation networks. Transport. Res. C Emerg. Technol. 147, 104012.

Gu, Y., Chen, A., Kitthamkesorn, S., 2022. Weibit choice models: properties, mode choice application and graphical illustrations. J. Choice Model 44, 100373.

Gu, Y., Chen, A., Kitthamkesorn, S., Jang, S., 2024. Alternate closed-form weibit-based model for assessing travel choice with an oddball alternative. Transp. Res. Part B Methodol. 179, 102867.

Haboucha, C.J., Ishaq, R., Shiftan, Y., 2017. User preferences regarding autonomous vehicles. Transport. Res. C Emerg. Technol. 78, 37–49.

Jang, S., Caiati, V., Rasouli, S., Timmermans, H., Choi, K., 2021. Does MaaS contribute to sustainable transportation? A mode choice perspective. Int. J. Sustain. Transp. 15, 351–363.

Jang, S., Rasouli, S., Timmermans, H., 2017. Bias in random regret models due to measurement error: formal and empirical comparison with random utility model. Transp. A Transp. Sci. 13, 405–434.

Jansuwan, S., Liu, Z., Song, Z., Chen, A., 2021. An evaluation framework of automated electric transportation system. Transp. Res. Part E Logist. Transp. Rev. 148, 102265.

Jiang, Y., Zhang, J., Wang, Y., Wang, W., 2019. Capturing ownership behavior of autonomous vehicles in Japan based on a stated preference survey and a mixed logit model with repeated choices. Int. J. Sustain. Transp. 13, 788–801.

Kitthamkesorn, S., Chen, A., 2013. A path-size weibit stochastic user equilibrium model. Transp. Res. Part B Methodol. 57, 378–397.

Olovsson, T., Svensson, T., Wu, J., 2022. Future connected vehicles: communications demands, privacy and cyber-security. Commun. Transport. Res. 2, 100056.

Ortúzar, J. de D., 2021. Future transportation: sustainability, complexity and individualization of choices. Commun. Transport. Res. 1, 100010.

Recker, W.W., 1995. Discrete choice with an oddball alternative. Transp. Res. Part B Methodol. 29, 201–211.

Song, F., 2019. Understanding Mode Choice Behaviour when New Modes Come into Play. University of Leeds, Leeds, USA. Ph.D. Dissertation.

Williams, H.C.W.L., Ortúzar, J.D., 1982. Behavioural theories of dispersion and the mis-specification of travel demand models. Transp. Res. Part B Methodol. 16, 167–219.

Yu Gu received his Ph.D. degree from the Department of Civil and Environmental Engineering, the Hong Kong Polytechnic University (PolyU), Hong Kong, China, in 2024. He is currently a postdoctoral fellow with the Department of Civil and Environmental Engineering at PolyU. His research interest includes the transportation network equilibrium model, random utility model, and transportation network reliability, vulnerability, and resilience.

Anthony Chen received his Ph.D. degree from the University of California at Irvine, Irvine, CA, USA, in 1997. From 2015 to 2017, he was the Chang Jiang Chair Professor with Tongji University, Shanghai, China. He is currently a Professor with the Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University (PolyU), Hong Kong, China. Prior to joining PolyU, he was a Professor with the Department of Civil and Environmental Engineering and the Head of the Transportation Division, Utah State University, Logan, UT, USA, for 17 years. He is currently an Associate Editor for *Transportmetrica A: Transport Science*, *Networks and Spatial Economics*, and *Journal of Advanced Transportation*, and an Editorial Board Member of *Transportation Research Part B: Methodological*.

Sunghoon Jang received his Ph.D. degree in transportation planning from the Eindhoven University of Technology under the supervision of Prof. Harry Timmermans in 2018. He is currently an Assistant Professor of transportation engineering at the Hong Kong Polytechnic University. He is an expert in advanced travel behavior modeling, aiming to understand human decision-making process using advanced mathematical models based on mathematical psychology and behavioral economics and to promote citizens' sustainable and active travel behavior to contribute to green transportation.

Songyot Kitthamkesorn received his Ph.D. degree from the Department of Civil and Environmental Engineering, Utah State University, Logan, USA, in 2013. He is currently an Assistant Professor with the Department of Civil Engineering at Chiang Mai University, Thailand. His research interest includes the transportation network equilibrium model, route choice model, and stochastic programming.