

Title: Comparing the mixed logit estimates and true parameters under informative and uninformative heterogeneity: a simulated discrete choice experiment

Authors: Maksat Jumamyradov; Benjamin M. Craig; William Greene; Murat Munkin.

Correspondence to: Maksat Jumamyradov; maksat@usf.edu; 4202 E Fowler Ave, Tampa, Florida, 33620, USA

Conflicts of interest/competing interest: Maksat Jumamyradov received funding from the EuroQol Research Foundation. Benjamin M. Craig is a member the EuroQol Research Foundation.

Funding: The EuroQol Research Foundation funded the doctoral research of Maksat Jumamyradov under the project titled: Sequential relief of child health problems (304-PHD).

Acknowledgement: The authors would like to thank the EuroQol Research Foundation for their support of Maksat Jumamyradov's dissertation.

Abstract

In discrete choice experiments (DCEs), differences between respondents' preferences may be associated with observable or unobservable factors. Unobservable heterogeneity, related to latent factors associated with the choices of individuals, may be modelled using correlated (i.e. informative heterogeneity) or uncorrelated (i.e. uninformative heterogeneity) individual-specific parameters of a logit model. In this study, we simulated unobservable heterogeneity among DCE respondents and compared the results of the maximum simulated likelihood (MSL) estimation of the mixed logit model when correctly specified and mis-specified. These results show that the MSL estimates are biased and can differ greatly from the true parameters, even when correctly specified. Before estimating a mixed logit model, we highly recommend that choice modellers conduct simulation analyses to assess the potential extent of biases before relying on the MSL estimates, particularly their variances and correlations, and then ultimately determine which model specification produces the least bias.

Keywords: Mixed Logit, Unobservable Heterogeneity, Maximum Simulated Likelihood.

1. Introduction

In a **discrete choice experiment (DCE)**, respondents are assigned sets of alternatives and asked to make choices under a hypothetical situation. The alternatives in each set are typically described using a descriptive system of multi-level attributes. Depending on the hypothetical situation, the choice sets may include a referent alternative (i.e., opt-out), for example, “watchful waiting” (e.g., Campbell and Erdem, 2019). In health preference research (HPR), the primary purpose of a DCE is to test and estimate the effects of attribute levels and different alternatives on preferential choice behaviors (Gonzalez, 2019). A common mantra in HPR is that choice defines value (Craig et al., 2017).

In a DCE, the effects of different attribute levels and alternatives on choice behaviors may vary between respondents due to differences in their preferences (i.e., **preference heterogeneity**). Such differences may be systematically associated with observable or unobservable factors. **Observable heterogeneity** refers to observed (by the researcher) factors associated with individuals' choices such as age or gender (e.g., Craig et al., 2022). On the other hand, **unobservable heterogeneity** refers to latent factors associated with the choices of individuals (e.g., risk perception).

Unobserved heterogeneity may be described using individual-specific parameters, which may or may not be correlated. **Uninformative heterogeneity** implies that individual-specific parameters vary but are uncorrelated within the sample. **Informative heterogeneity** means that individual parameters are correlated through latent factors; therefore, fixing one parameter for a respondent affects the distribution of another. Jumamyradov et al. (2023) simulated DCE data under uninformative and informative heterogeneity and compared the conditional logit (CL) estimates to their true values. They found that the extent and form of unobservable heterogeneity may result in biased CL estimates, and these biases substantively affect willingness-to-pay estimates. It is not typically feasible to estimate individual-specific parameters directly; however, some researchers allow for random parameters, namely mixed logit specifications. As a natural extension of this prior simulation study, this paper investigates the effects of alternative mixed logit specifications on the biases caused by unobservable heterogeneity.

Informative heterogeneity may be separated into two subtypes based on which parameters are correlated. **Substitution patterns** (e.g., Brownstone and Train, 1999) refer to the presence of a latent factor that causes an association with alternative-specific constants (ASCs). For example, the classic red-bus-blue-bus problem refers to a positive correlation between the two bus alternatives, where a red bus is a substitute for a blue one (e.g., Quandt, 1970). An ASC is not the same as a nominal attribute. In this classic problem, every set includes three fixed alternatives, namely a car, a red bus, and a blue bus. If the color and mode were nominal attributes and varies independently, some sets might include three buses. In health economics, a more common example is a choice set with multiple fixed treatment options (e.g., a brand-name medication and a generic medication) as well as “watchful waiting” (i.e., no medication).

Taste patterns (e.g., Revelt and Train, 1998) refer to the presence of a latent factor that causes an association between attribute coefficients more generally (i.e., attribute importance). In health economics, each alternative may vary by out-of-pocket price, wait times, and other attributes. The coefficients of these attributes may be correlated. For example, individuals who are willing to pay more for their health care may also prefer more luxurious services (e.g., a brand-name medication).

Imagine a specific alternative (e.g., a brand-name medication). Under uninformative interpersonal heterogeneity, the effect of a specific brand varies between persons, but this effect is not associated with another alternative's effect (i.e., substitution pattern) or the effects of any attributes (i.e., taste pattern).

Alternatively, latent factors may induce informative heterogeneity, such as a substitution pattern between brands or a taste pattern between a brand and out-of-pocket cost.

For simplicity, we generated preferential choice behaviors for a standard DCE with a full factorial design (see Appendix A) and individual-specific unobserved coefficients drawn from a lognormal distribution. Each task represents a choice between three alternatives to illustrate uninformative and informative heterogeneity, including substitution and taste patterns. Each alternative is described using a nominal attribute and a numeric one, namely: medication and out-of-pocket price (see Appendix A for out-of-pocket price). For example, the correlation between the effects of brand and generic is a substitution pattern, and the correlation between the effects of medication and price is a taste pattern. Willingness-to-pay for a brand-name medication is the ratio of the brand-name medication and price effects. This DCE was constructed to mimic the classic red-bus-blue-bus example.

Generally, a mixed logit is the same as a conditional logit, except it allows for random parameters (i.e., more flexible) and is estimated using maximum simulated likelihood (MSL). The MSL estimator has become a widely used simulation-based method among applied economists to numerically evaluate models the likelihood function of which cannot be expressed in a closed-form and, therefore, cannot be evaluated analytically using the maximum likelihood (ML) estimator. However, the MSL estimator is biased for a fixed number of simulations, which is a results of the log transformation of the simulated likelihood (see Gourieroux and Monfort, 1990, Lee 1995, and Hajivassiliou et al., 1994, Train, 2009). Furthermore, recently Jumamyradov and Munkin (2022) showed that the MSL estimator produce significant biases when applied to the bivariate normal and bivariate Poisson-lognormal models. We expect that similar biases might appear in the MSL estimation of the mixed logit model within the context of a DCE.

Apart from simulation bias and other nuances (e.g., number of Halton draws), determining which parameters should be allowed to be random in the mixed logit model depends on the researcher's assumptions. Knowing the specification of the unobservable heterogeneity in advance can be challenging. Therefore, this simulation study examines the performance of the MSL estimation of the mixed logit model targeting three questions:

1. What happens when the mixed logit model is correctly specified (**simulation bias**)?
2. What happens when the mixed logit model erroneously restricts a correlation (**under-specification**)?
3. What happens when the mixed logit model erroneously allows a correlation (**over-specification**)?

More specifically, we compare the mixed logit estimates and true parameters under uninformative and informative heterogeneity to assess differences due to simulation and misspecification. In this paper, we emphasized the estimation of the correlations, not the variances. Note that increasing the variances only increases the scale of the biases and has no effect on the correlations.

Depending on the experimental design, analysts may take two general approaches when accounting for interpersonal heterogeneity in panel data analysis. First, in rare cases, the analyst can estimate individual-specific parameters per respondent if the DCE has a large number of tasks per respondent relative to the number of parameters (i.e., fixed effects). When this is not the case, the analyst may estimate a random parameter model by making distributional assumptions about the prevalence of parameters within the population (e.g., mixed logit). To estimate a random-parameter model, the analyst will need to specify the correlation structure between the random parameters, often with little prior information.

Since its advent in 1980s, mixed logit estimation has been increasingly utilized in the analysis of DCE responses (e.g., air travel: Hossain et al., 2018; healthcare: Soekhai et al., 2019). For instance, Clark et al. (2014) report that while only 3% of health-related DCE papers published during 1990-2000 utilized the mixed logit model in their estimations, that number increased to 21% during 2009-2012. Our null hypothesis is that the mixed logit estimation accurately describes the unobservable heterogeneity in DCE responses. The rest of the paper is organized as follows. Section 2 presents the mixed logit model, the DCE simulations, MSL estimation, and comparison procedure details. Section 3 presents the results and concludes.

2. Methods

a. The Mixed Logit Model

The theoretical framework of choice analysis is based on the random utility maximization (RUM) theory (e.g., Marschak, 1960). According to RUM theory, utility function $U_{itj} = V_{itj} + \varepsilon_{itj}$ of individual $i = 1, \dots, N$ for alternative $j = 1, \dots, J$ at choice task $t = 1, \dots, T$ can be decomposed into deterministic part of utility V_{itj} (representative utility) which is observed by the researcher, and random part of utility ε_{itj} which is not observed by the researcher.

Following McFadden (1974), individual i will choose alternative j if and only if the probability that utility associated with alternative j is higher than utilities associated with all other alternatives

$$\begin{aligned} P_{itj} &= P(U_{itj} > U_{itk}, \forall k \neq j) \\ P_{itj} &= P(V_{itj} + \varepsilon_{itj} > V_{itk} + \varepsilon_{itk}, \forall k \neq j) \\ P_{itj} &= P(\varepsilon_{itk} - \varepsilon_{itj} < V_{itj} - V_{itk}, \forall k \neq j) \end{aligned} \quad (1)$$

To derive the choice probabilities, we need to make distributional assumptions about the random part of utility. The conditional logit (CL) model is derived under the assumption that ε_{itj} is independently and identically distributed (IID) with extreme value type 1 (EV1) distribution (e.g., Revelt and Train, 1998; Brownstone and Train, 1999; McFadden and Train, 2000). As a result, the difference of two IID EV1 random error terms ($\varepsilon_{itk} - \varepsilon_{itj}$) has a logistic distribution. This implies that the choice probabilities of the conditional logit model can be expressed in terms of logistic distribution with cumulative distribution function

$$P_{itj} = \frac{1}{1 + \sum_{k=1}^J \exp[(V_{itk} - V_{itj})]}, \forall k \neq j \quad (2)$$

Under the conditional logit, representative utility V_{itj} is a function of alternative attributes, typically linear in parameters. However, the mixed logit extends the conditional logit through individual-specific parameters (e.g., Hensher et al., 2005). The mixed logit model can account for various kinds of unobservable heterogeneity, but its flexibility comes with a cost. The choice probabilities of the mixed logit model do not have a closed-form expression. Therefore, the estimation of the mixed logit model relies on simulation-based methods such as the maximum simulated likelihood estimator (MSL) (e.g., Greene, 2012). In this paper, **simulation bias** refers to the difference between estimated and true parameter values, when the estimated mixed logit model is correctly specified.

Furthermore, we examine biases due to **misspecification** of the mixed logit models. In particular, a model may be under-specified when the researcher erroneously restricts a correlation (i.e. **under-specification**

bias) between random parameters. Similarly, a researcher erroneously assumes that random parameters are correlated with each other (i.e. **over-specification bias**). We conducted a series of simulations to assess differences between mixed logit estimates and true parameters due to simulation and misspecification biases.

b. Simulation

Using a standard DCE, we generate choices between three alternatives, where the third alternative serves as a referent (i.e., opt-out; Campbell and Erdem, 2019). For identification of the parameters, we normalize the model by assuming that the coefficients of the opt-out alternative are equal to zero. Therefore, the deterministic part of the opt-out alternative will be represented as $V_{it3} = 0$. For example, imagine the decision between a brand-name medication, generic medication or doing nothing. Hence, we assume that the representative utilities of three alternatives will take the following form

$$\begin{aligned} V_{it1} &= \alpha_{i1} - \beta_i x_{t1} \\ V_{it2} &= \alpha_{i2} - \beta_i x_{t2} \\ V_{it3} &= 0 \end{aligned} \quad (3)$$

and therefore, utility differences between each alternative and opt-out will be

$$\begin{aligned} Z_{it1} &= \alpha_{i1} - \beta_i x_{t1} + \eta_{it1} & \eta_{it1} &\sim \text{Logistic}(0,1) \\ Z_{it2} &= \alpha_{i2} - \beta_i x_{t2} + \eta_{it2} & \eta_{it2} &\sim \text{Logistic}(0,1) \end{aligned} \quad (4)$$

where x_{t1} and x_{t2} are out-of-pocket prices (see the Appendix A for the values of out-of-pocket price), α_{i1} and α_{i2} are alternative-specific constants (ASCs), β_i is a random price coefficient, $Z_{it1} = U_{it1} - U_{it3}$ and $Z_{it2} = U_{it2} - U_{it3}$ are two utility differences, and finally $\eta_{it1} = \varepsilon_{it1} - \varepsilon_{it3}$ and $\eta_{it2} = \varepsilon_{it2} - \varepsilon_{it3}$ are differences in IID EV1 error terms. Under this specification, the willingness-to-pay for each brand is the ratio of the ASC and price coefficient, α_{i1}/β_i and α_{i2}/β_i . Notice that we have a negative sign in front of the price coefficient β_i to represent the fact that people dislike paying more.

In our model specification, we use two ASCs (α_{i1} and α_{i2}) to allow for their correlation. Alternatively, a nominal attribute for opt-in (α_i) may be included instead of two ASCs. However, that specification would deviate from the red-bus-blue-bus example. Similar in health economics is the case of choosing between three ways to see heartburn relief where the alternatives are a brand-name medication, a generic medication or doing nothing. It seems appropriate to allow the ASCs for the brand-name and generic medications to be correlated, and not combined as a treatment attribute.

Since the random out-of-pocket cost β_i enters the denominator, careful consideration of its distribution is important. Daly et al. (2012) show that whether the distribution of WTP has finite moments depends on the distribution of β_i . In this paper, we assume that all three random variables have a lognormal distribution such as

$$\begin{aligned} \beta_i &= \exp(\beta + v_{\beta_i}) \\ \alpha_{1i} &= \exp(\alpha_1 + v_{\alpha_{1i}}) \\ \alpha_{2i} &= \exp(\alpha_2 + v_{\alpha_{2i}}) \end{aligned}$$

where $(v_{\alpha_{1i}}, v_{\alpha_{2i}}, v_{\beta_i}) \sim N((0,0,0), \Sigma)$ has a multivariate normal distribution with the following covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_\alpha & \sigma_1\sigma_\beta\rho_\beta \\ \sigma_1\sigma_2\rho_\alpha & \sigma_2^2 & 0 \\ \sigma_1\sigma_\beta\rho_\beta & 0 & \sigma_\beta^2 \end{bmatrix} \quad (5)$$

with $\sigma_1^2 = \text{Var}(v_{\alpha_{1i}})$, $\sigma_2^2 = \text{Var}(v_{\alpha_{2i}})$, $\sigma_\beta^2 = \text{Var}(v_{\beta_i})$, $\rho_\alpha = \text{Corr}(v_{\alpha_{1i}}, v_{\alpha_{2i}})$, and $\rho_\beta = \text{Corr}(v_{\alpha_{1i}}, v_{\beta_i})$. In this covariance matrix, informative heterogeneity can further be separated based on whether the correlation is between the alternative-specific constants, ρ_α (i.e. **substitution patterns**) or alternative attributes, ρ_β (i.e. **taste patterns**). Notice that we do not generate data with $\text{Corr}(v_{\alpha_{2i}}, v_{\beta_i})$, because we found no substantive differences between the data generated with $\text{Corr}(v_{\alpha_{1i}}, v_{\beta_i})$ or $\text{Corr}(v_{\alpha_{2i}}, v_{\beta_i})$.

The mean and variance of lognormal random variables β_i and α_{1i} are

$$\begin{aligned} E(\beta_i) &= \exp(\beta + \sigma_\beta^2/2) \\ E(\alpha_{1i}) &= \exp(\alpha_1 + \sigma_1^2/2) \\ \text{Var}(\beta_i) &= \exp(2\beta + \sigma_\beta^2)[\exp(\sigma_\beta^2) - 1] \\ \text{Var}(\alpha_{1i}) &= \exp(2\alpha_1 + \sigma_1^2)[\exp(\sigma_1^2) - 1] \end{aligned}$$

Notice that the WTP, which is a ratio of two lognormal random variables, also has a lognormal distribution

$$WTP_{\alpha_1} = \frac{\alpha_{1i}}{\beta_i} = \frac{\exp(\alpha_1 + v_{\alpha_{1i}})}{\exp(\beta + v_{\beta_i})} = \exp(\alpha_1 - \beta + v_{\alpha_{1i}} - v_{\beta_i})$$

with mean and variance as following

$$\begin{aligned} E(WTP_{\alpha_1}) &= \exp\left(\alpha_1 - \beta + \frac{\sigma_1^2 - 2\sigma_1\sigma_\beta\rho_\beta + \sigma_\beta^2}{2}\right) \\ \text{Var}(WTP_{\alpha_1}) &= \exp(2(\alpha_1 - \beta) + \sigma_1^2 - 2\sigma_1\sigma_\beta\rho_\beta + \sigma_\beta^2)[\exp(\sigma_1^2 - 2\sigma_1\sigma_\beta\rho_\beta + \sigma_\beta^2) - 1] \end{aligned}$$

In the data generation process, the true means of the representative utility parameters (equation 3) were fixed at $\alpha_1 = \alpha_2 = \beta = 1$ (e.g., brand-name medication, generic medication, and out-of-pocket price, respectively). Furthermore, we generate various covariance matrices (equation 5) based on the true values we choose for the variances and correlations, and we refer to each distinct covariance matrix as a “parameter set” (see Table 1). For each parameter set, we generated 100 datasets ($R = 100$), each including 25 choices ($T = 25$) for 200 respondents ($N = 200$) given a standard DCE with a full factorial experimental design. We acknowledge that this design may include dominated alternatives. However, since this is a simulation study, the existence of dominated alternatives should not be a concern.

The simulated covariance matrices fall into three distinct groups: the first group (i.e., Group 1) represents uninformative heterogeneity (i.e., uncorrelated random parameters; $\rho_\alpha = \rho_\beta = 0$) with different variances, and the last two groups (i.e., Group 2 and Group 3) represent informative heterogeneity (i.e., correlated random parameters; $\rho_\alpha \neq 0$ or $\rho_\beta \neq 0$) where the variances are fixed at 0.15. For instance, in Group 1 we assume that σ_1^2 , σ_2^2 and σ_β^2 range from 0.05 to 0.25 with increments of 0.1 (i.e. $\sigma_1^2 = \sigma_2^2 = \sigma_\beta^2 = \{0.05, 0.15, 0.25\}$ or $\sigma_1^2 = \sigma_2^2 = \sigma_\beta^2 = \{0.05: 0.1: 0.25\}$) and we assume that $\rho_\alpha = \rho_\beta = 0$. Since

there are 3 different values for each σ_1^2 , σ_2^2 and σ_β^2 , we can create $3^3=27$ combinations of parameter sets for the true covariance matrix.

In the Group 2 we generate the data with fixed variances $\sigma_1^2 = \sigma_2^2 = \sigma_\beta^2 = 0.15$ and different values for ρ_β ranging from -0.95 to 0.95 with increments of 0.05 (i.e. $\rho_\beta = \{-0.95:0.05:0.95\}$). In the Group 3 we generate the data with fixed variances $\sigma_1^2 = \sigma_2^2 = \sigma_\beta^2 = 0.15$ and different values for ρ_α ranging from 0.05 to 0.95 with increments of 0.05 (i.e. $\rho_\alpha = \{0.05:0.05:0.95\}$). Therefore, in Group 2 and Group 3, there are 38 and 19 nonzero correlation values, respectively. For brevity, we define $\sigma^2 = \{\sigma_1^2, \sigma_2^2, \sigma_\beta^2\}$ and $\rho = \{\rho_\alpha, \rho_\beta\}$.

Table 1: Parameter Sets for the Simulation of Informative and Uninformative Heterogeneity

	Group	σ_1^2	σ_2^2	σ_β^2	ρ_α	ρ_β	Parameter sets
Uninformative Heterogeneity	1 No Patterns	0.05:0.1:0.25	0.05:0.1:0.25	0.05:0.1:0.25	0	0	27
Informative Heterogeneity	2 Taste Patterns	0.15	0.15	0.15	0	$-0.95:0.05:0.95$	38
	3 Substitution Patterns	0.15	0.15	0.15	$0.05:0.05:0.95$	0	19

c. Estimation

To analyze simulation and misspecification biases in each simulated dataset with various parameter sets, we estimated different specifications of the mixed logit model using 500 Halton draws. For simulation bias, we estimate the mixed logit model assuming that the researcher knows the accurate specification of the data generation process. For instance, when the data is generated using the Group 1, we estimate the mixed logit model assuming $\rho_\alpha = \rho_\beta = 0$. When the data are generated using the Group 2, we estimate the mixed logit model assuming $\rho_\alpha = 0$ and $\rho_\beta \neq 0$. Lastly, when the data are generated using the Group 3, we estimate the mixed logit model assuming $\rho_\beta = 0$ and $\rho_\alpha \neq 0$.

To analyze misspecification bias, we estimate an inaccurate specifications of the mixed logit model by either restricting (i.e., under-specification) or unrestricting (i.e. over-specification) correlations in the estimation process. For instance, to analyze the under-specification bias, we generate the data using the Group 2 or Group 3, however, we estimate the mixed logit model where we restrict $\rho_\alpha = \rho_\beta = 0$. To analyze the over-specification bias, we generate the data using the Group 1, however, we estimate the model where we allow $\rho_\beta \neq 0$ or $\rho_\alpha \neq 0$. The goal in these analysis is to examine whether MSL estimation of the mixed logit model produce accurate estimates, and more importantly, to see if the results produce accurate WTP.

Overall, in this paper we estimate three different mixed logit models, which differ based on correlation restrictions. The first model, Model 1, assumes that $\rho_\alpha = \rho_\beta = 0$. The second model, Model 2, assumes that $\rho_\beta \neq 0$ but $\rho_\alpha = 0$. In the third model, Model 3, we assume that $\rho_\alpha \neq 0$ but $\rho_\beta = 0$.

In Table 2 we summarize different biases we analyze in this paper (i.e., simulation, over-specification and under-specification biases). Notice that if we generate the data using the Group 1 and estimate the mixed logit model using the Model 1, occurring bias in the estimated parameters will represent the simulation bias. Similarly, if we generate the data using the Group 1 and estimate the Model 2, occurring bias in the estimated parameters will represent the over-specification bias.

Table 2: Different models and groups of data generation processes.

	Group 1	Group 2	Group 3

	No Patterns: $\rho_\beta = \rho_\alpha = 0$	Taste: $\rho_\beta \neq 0, \rho_\alpha = 0$	Substitution: $\rho_\beta = 0, \rho_\alpha \neq 0$
Model 1 No Patterns: $\rho_\beta = \rho_\alpha = 0$	Simulation bias	Under-specification bias, restricted ρ_β	Under-specification bias, restricted ρ_α
Model 2 Taste: $\rho_\beta \neq 0, \rho_\alpha = 0$	Over-specification bias, unrestricted ρ_β	Simulation bias	
Model 3 Substitution: $\rho_\beta = 0, \rho_\alpha \neq 0$	Over-specification bias, unrestricted ρ_α		Simulation bias

d. Comparison between True and Estimated Values

First, we describe simulation bias by comparing the true and estimated values obtained from accurate specification of uninformative (i.e. Group 1) and informative heterogeneity (i.e., Group 2 and Group 3). As a measure of biases, we use the Wald distance of the estimated and true values based on the estimated standard errors (see Greene, 2012). Second, we describe under-specification bias by comparing the true and estimated values obtained from Group 2 (i.e., restricted $\rho_\beta = 0$) and Group 3 (i.e., restricted $\rho_\alpha = 0$). Lastly, we describe over-specification bias by comparing the true and estimated values obtained from the Group 1 (i.e., unrestricted $\rho_\beta \neq 0$ or $\rho_\alpha \neq 0$).

3. Results

This section summarizes the results that we believe ideally address the questions put forward in the introduction regarding simulation bias, under-specification (i.e., restricted correlation) and over-specification (i.e. unrestricted correlation) of MSL estimation of the mixed logit model. **Standard error of** each parameter was calculated by dividing its standard deviation over 100 simulations by the square root of 100.

3.1. Simulation Bias

As mentioned earlier, simulation bias represents the bias occurring when the mixed logit model is correctly specified. To analyze the simulation bias, we present Table 3 where we show MSL results of the mixed logit model with the Group 1 data generation process. Since there are 27 parameter sets with different true variances, in Table 3 we present the results of only nine combinations of variances. The rest of the results can be found in the Supplementary Materials. The true values of variances in Table 3 are chosen such that σ_1^2 and σ_2^2 each has three values $\{0.05, 0.15, 0.25\}$, and σ_β^2 does not vary ($\sigma_\beta^2 = 0.15$). In Table 3, we show the estimated parameters as well as the true and estimated WTP with their respective standard errors in brackets.

There are three observations that we would like to point out about the simulation bias on Group 1. First, notice that the MSL estimator produces biased results for α_1 , α_2 and β in all nine specifications presented in the table. Specifically, when the true coefficients are $\alpha_1 = 1$, $\alpha_2 = 1$, $\beta = 1$, with variances $\sigma_1^2 = 0.05$, $\sigma_2^2 = 0.05$ and $\sigma_\beta^2 = 0.15$ (i.e., first column), the estimated coefficients are equal to 1.117 (0.003), 1.118 (0.003) and 1.121 (0.003), respectively, all three of which are separated from their true values by approximately 40 standard errors.

Second observation is that the estimated variances are biased as well. For instance, when the true variances are $\sigma_1^2 = 0.05$, $\sigma_2^2 = 0.05$ and $\sigma_\beta^2 = 0.15$ (i.e., first column), MSL estimates are 0.222 (0.003), 0.219 (0.003) and 0.378 (0.003) which are separated from their true values by more than 57, 56 and 43 standard errors, respectively. Therefore, the null hypothesis that the estimated coefficients and variances are equal to their respective true values may be rejected for all these parameters.

Third observation is that even though the parameter estimates are biased, the estimated WTP results are close to their true values. For instance, when the true variances are $\sigma_1^2 = 0.05$, $\sigma_2^2 = 0.05$ and $\sigma_\beta^2 = 0.15$

(i.e., first column), the estimated $E(WTP_{\alpha_1})$ and $Var(WTP_{\alpha_1})$ are -1.097 (0.004) and 0.259 (0.004), which are separated from their respective true values -1.105 and 0.270 by less than three standard errors (i.e. typical margin or error). Furthermore, all 36 estimated WTP parameters (i.e., 9 columns, 2 mean values and 2 variances each) are within the typical margin of error. Therefore, we conclude that although there is a simulation bias in the coefficients and variances of the Group 1, it does not substantively affect the estimated WTP. Similar patterns can be seen in results presented in the Supplementary Materials.

Table 3: Model 1 Estimation for Group 1 (i.e., Simulation Bias)

	$\sigma_1^2 = 0.05$ $\sigma_2^2 = 0.05$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.15$ $\sigma_2^2 = 0.05$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.25$ $\sigma_2^2 = 0.05$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.05$ $\sigma_2^2 = 0.15$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.15$ $\sigma_2^2 = 0.15$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.25$ $\sigma_2^2 = 0.15$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.05$ $\sigma_2^2 = 0.25$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.15$ $\sigma_2^2 = 0.25$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.25$ $\sigma_2^2 = 0.25$ $\sigma_\beta^2 = 0.15$
$\alpha_1 = 1$	1.117 (0.003)	1.110 (0.004)	1.117 (0.005)	1.114 (0.003)	1.105 (0.004)	1.109 (0.005)	1.119 (0.003)	1.110 (0.004)	1.111 (0.005)
$\alpha_2 = 1$	1.118 (0.003)	1.107 (0.003)	1.116 (0.003)	1.116 (0.004)	1.107 (0.004)	1.110 (0.004)	1.123 (0.005)	1.115 (0.005)	1.108 (0.004)
$\beta = 1$	1.121 (0.003)	1.113 (0.003)	1.116 (0.004)	1.120 (0.004)	1.114 (0.004)	1.111 (0.003)	1.125 (0.003)	1.115 (0.004)	1.121 (0.004)
σ_1^2	0.222 (0.003)	0.383 (0.003)	0.493 (0.004)	0.225 (0.003)	0.388 (0.003)	0.494 (0.004)	0.219 (0.007)	0.383 (0.004)	0.500 (0.005)
σ_2^2	0.219 (0.003)	0.232 (0.003)	0.218 (0.005)	0.384 (0.003)	0.386 (0.004)	0.387 (0.003)	0.493 (0.005)	0.490 (0.004)	0.500 (0.004)
σ_β^2	0.378 (0.003)	0.377 (0.003)	0.373 (0.008)	0.375 (0.003)	0.380 (0.003)	0.378 (0.004)	0.364 (0.011)	0.378 (0.004)	0.374 (0.004)
True $E(WTP_{\alpha_1})$	-1.105	-1.162	-1.221	-1.105	-1.162	-1.221	-1.105	-1.162	-1.221
True $E(WTP_{\alpha_2})$	-1.105	-1.105	-1.105	-1.162	-1.162	-1.162	-1.221	-1.221	-1.221
Est. $E(WTP_{\alpha_1})$	-1.097 (0.004)	-1.154 (0.005)	-1.219 (0.007)	-1.095 (0.005)	-1.150 (0.006)	-1.215 (0.007)	-1.097 (0.004)	-1.153 (0.005)	-1.208 (0.008)
Est. $E(WTP_{\alpha_2})$	-1.098 (0.005)	-1.098 (0.005)	-1.104 (0.004)	-1.152 (0.006)	-1.153 (0.006)	-1.160 (0.006)	-1.215 (0.007)	-1.215 (0.007)	-1.204 (0.007)
True $Var(WTP_{\alpha_1})$	0.270	0.472	0.734	0.270	0.472	0.734	0.270	0.472	0.734
True $Var(WTP_{\alpha_2})$	0.270	0.270	0.270	0.472	0.472	0.472	0.734	0.734	0.734
Est. $Var(WTP_{\alpha_1})$	0.259 (0.004)	0.453 (0.008)	0.719 (0.017)	0.258 (0.005)	0.463 (0.009)	0.715 (0.017)	0.265 (0.006)	0.457 (0.010)	0.719 (0.020)
Est. $Var(WTP_{\alpha_2})$	0.258 (0.005)	0.266 (0.005)	0.266 (0.005)	0.450 (0.008)	0.462 (0.009)	0.466 (0.010)	0.712 (0.018)	0.704 (0.015)	0.707 (0.015)

In Table 4, we present MSL parameter estimates for Group 2 (taste patterns). Notice that we do not estimate σ_1^2 because we restrict it at its true value 0.15 for identification. Since in Group 2 there are 38 parameter sets with different true correlations (see Table 1), in Table 4 we present only the results of eight parameter sets. The rest of the results can be found in the Supplementary Materials.

Similar to the results we found in Table 3, the MSL estimator produces biased results in the Group 2 as well. Specifically, in all eight specifications presented in Table 4, the estimated mean coefficients (i.e., α_1 , α_2 , β) are separated from their true values by approximately 40 standard errors. Furthermore, in all eight columns of Table 4, the estimated variances (i.e., σ_2^2 , σ_β^2) are separated from their true values by approximately 70 standard errors. Therefore, the null hypothesis that the estimated coefficients and variances are equal to their respective true values may be rejected.

On the contrary, we can see in Table 4 that MSL results of the estimated ρ_β are relatively accurate than the estimated coefficients and variances. For instance, the true and estimated ρ_β are separated from each other by more than three standard errors only in two out of eight columns. Specifically, when the true correlation is $\rho_\beta = -0.75$ and $\rho_\beta = -0.35$, their estimated values are -0.798 (0.013) and -0.398 (0.013), respectively, both of which are separated from their true values by 3.7 standard errors.

However, similar to the results we observed in Table 3, even though there are biases in the estimated coefficients, variances and correlation, the estimated mean and variance of WTP are mostly close to their true values. There are only a few instances where the WTP parameters are separated from their true

values by more than 3 standard errors (i.e., biased). For instance, the estimated $E(WTP_{\alpha_1})$ is biased when the true correlation $\rho_\beta = 0.55$ and $\rho_\beta = 0.75$, the estimated $E(WTP_{\alpha_2})$ is biased when the true correlation $\rho_\beta = -0.75$, and finally the estimated $Var(WTP_{\alpha_2})$ is biased when the true correlation $\rho_\beta = 0.35$ and $\rho_\beta = 0.55$. Therefore, the simulation bias in the Group 2 (taste patterns) does not seem to affect the estimated WTP.

Table 4: Model 2 Estimation for Group 2 (i.e., Simulation Bias)

	$\rho_\beta = -0.95$	$\rho_\beta = -0.75$	$\rho_\beta = -0.55$	$\rho_\beta = -0.35$	$\rho_\beta = 0.35$	$\rho_\beta = 0.55$	$\rho_\beta = 0.75$	$\rho_\beta = 0.95$
$\alpha_1 = 1$	1.113 (0.004)	1.107 (0.004)	1.108 (0.004)	1.106 (0.004)	1.106 (0.004)	1.110 (0.004)	1.106 (0.004)	1.121 (0.004)
$\alpha_2 = 1$	1.115 (0.004)	1.102 (0.004)	1.106 (0.004)	1.108 (0.004)	1.109 (0.003)	1.112 (0.003)	1.114 (0.004)	1.123 (0.004)
$\beta = 1$	1.115 (0.004)	1.117 (0.004)	1.116 (0.004)	1.114 (0.003)	1.117 (0.003)	1.118 (0.003)	1.121 (0.004)	1.131 (0.004)
$\sigma_\gamma^2 = 0.15$	0.379 (0.003)	0.385 (0.003)	0.385 (0.004)	0.378 (0.004)	0.391 (0.004)	0.382 (0.003)	0.381 (0.004)	0.379 (0.004)
$\sigma_\beta^2 = 0.15$	0.385 (0.003)	0.380 (0.004)	0.383 (0.004)	0.381 (0.003)	0.371 (0.003)	0.368 (0.003)	0.373 (0.003)	0.380 (0.003)
ρ_β	-0.943 (0.009)	-0.798 (0.013)	-0.588 (0.015)	-0.398 (0.013)	0.294 (0.010)	0.527 (0.008)	0.734 (0.006)	0.948 (0.002)
True $E(WTP_{\alpha_1})$	-1.340	-1.300	-1.262	-1.225	-1.102	-1.070	-1.038	-1.008
True $E(WTP_{\alpha_2})$	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162
True $Var(WTP_{\alpha_1})$	1.427	1.167	0.942	0.749	0.262	0.165	0.084	0.015
True $Var(WTP_{\alpha_2})$	0.472	0.472	0.472	0.472	0.472	0.472	0.472	0.472
Est. $E(WTP_{\alpha_1})$	-1.337 (0.009)	-1.293 (0.008)	-1.259 (0.009)	-1.221 (0.007)	-1.095 (0.004)	-1.063 (0.004)	-1.025 (0.003)	-0.998 (0.002)
Est. $E(WTP_{\alpha_2})$	-1.160 (0.005)	-1.143 (0.006)	-1.150 (0.006)	-1.151 (0.006)	-1.149 (0.005)	-1.147 (0.005)	-1.147 (0.006)	-1.148 (0.006)
Est. $Var(WTP_{\alpha_1})$	1.425 (0.031)	1.189 (0.028)	0.978 (0.025)	0.769 (0.016)	0.272 (0.005)	0.165 (0.003)	0.085 (0.002)	0.016 (0.001)
Est. $Var(WTP_{\alpha_2})$	0.463 (0.008)	0.451 (0.008)	0.464 (0.010)	0.450 (0.009)	0.454 (0.008)	0.436 (0.009)	0.442 (0.009)	0.447 (0.008)

In Table 5, we present the results for the Group 3 (substitution patterns) where we find similar patterns in terms of the estimated parameters and WTP. Specifically, in all nine columns presented in Table 5, the estimates of coefficients and variances are separated from their true values by more than three standard errors, and therefore, the MSL estimator produces biased results for the coefficients α_1 , α_2 and β , and variances σ_1^2 and σ_2^2 .

However, similar to the results in Table 4, the MSL estimator produces relatively more accurate results for the correlation parameter ρ_α . For instance, in Table 5, MSL produces biased result for ρ_α only when its true value is $\rho_\alpha = 0.5$, which is separated from its estimated value 0.532 (0.009) by 3.5 standard errors. However, regardless of these simulation biases occurring in the estimation of parameters, they do not substantively affect the estimated WTP.

Table 5: Model 3 Estimation for Group 3 (i.e., Simulation Bias)

	$\rho_\alpha = 0.1$	$\rho_\alpha = 0.2$	$\rho_\alpha = 0.3$	$\rho_\alpha = 0.4$	$\rho_\alpha = 0.5$	$\rho_\alpha = 0.6$	$\rho_\alpha = 0.7$	$\rho_\alpha = 0.8$	$\rho_\alpha = 0.9$
$\alpha_1 = 1$	1.112 (0.004)	1.112 (0.004)	1.118 (0.004)	1.115 (0.004)	1.113 (0.004)	1.111 (0.004)	1.110 (0.003)	1.117 (0.004)	1.111 (0.004)
$\alpha_2 = 1$	1.107 (0.004)	1.116 (0.005)	1.120 (0.004)	1.114 (0.004)	1.115 (0.004)	1.110 (0.004)	1.111 (0.003)	1.116 (0.004)	1.110 (0.004)
$\beta = 1$	1.118 (0.004)	1.127 (0.004)	1.120 (0.004)	1.121 (0.003)	1.117 (0.004)	1.120 (0.004)	1.109 (0.004)	1.117 (0.004)	1.120 (0.004)
$\sigma_\gamma^2 = 0.15$	0.387 (0.004)	0.392 (0.004)	0.390 (0.004)	0.390 (0.004)	0.382 (0.003)	0.393 (0.004)	0.390 (0.004)	0.388 (0.003)	0.390 (0.003)
$\sigma_\beta^2 = 0.15$	0.379 (0.004)	0.366 (0.003)	0.369 (0.003)	0.373 (0.003)	0.382 (0.003)	0.374 (0.003)	0.381 (0.003)	0.370 (0.003)	0.372 (0.003)
ρ_α	0.126 (0.014)	0.237 (0.013)	0.321 (0.013)	0.426 (0.011)	0.532 (0.009)	0.616 (0.008)	0.711 (0.007)	0.806 (0.005)	0.904 (0.002)
True $E(WTP_{\alpha_1})$	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162
True $E(WTP_{\alpha_2})$	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162

True $Var(WTP_{\alpha_1})$	0.472	0.472	0.472	0.472	0.472	0.472	0.472	0.472	0.472
True $Var(WTP_{\alpha_2})$	0.472	0.472	0.472	0.472	0.472	0.472	0.472	0.472	0.472
Est. $E(WTP_{\alpha_1})$	-1.153	-1.138	-1.154	-1.151	-1.156	-1.148	-1.163	-1.156	-1.147
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	(0.005)
Est. $E(WTP_{\alpha_2})$	-1.148	-1.145	-1.159	-1.152	-1.157	-1.150	-1.166	-1.155	-1.147
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)
Est. $Var(WTP_{\alpha_1})$	0.460	0.429	0.447	0.449	0.466	0.450	0.468	0.450	0.444
	(0.008)	(0.006)	(0.007)	(0.007)	(0.008)	(0.008)	(0.007)	(0.008)	(0.006)
Est. $Var(WTP_{\alpha_2})$	0.461	0.445	0.459	0.457	0.463	0.461	0.479	0.454	0.450
	(0.010)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.010)	(0.008)

3.2. Under Specification Bias

Next, we analyze under-specification bias. Recall that in Table 4 we present the Model 2 estimation (i.e., correct specification) results of the Group 2 and analyze simulation bias. In Table 6, we present the Model 1 estimation (i.e., restricted $\rho_\beta = 0$) results for the Group 2. By comparing Tables 4 and 6, we infer the effects of restricting ρ_β on the mixed logit estimates apart from simulation bias.

There are three observations that we would like to point out. The first observation relates to the estimated coefficients α_1 , α_2 , and β . The second observation is related to the estimated variances σ_2^2 and σ_β^2 . Finally, the third observation relates to the estimated WTP parameters.

First, restricting a positive ρ_β has a different effect on the estimated coefficients than restricting a negative ρ_β . Specifically, restricting a negative ρ_β slightly increases the bias in the estimated coefficients α_1 , α_2 , and β . However, restricting a positive ρ_β substantively reduces the bias in the coefficients, countering the simulation bias. For instance, in Table 6, when the true $\rho_\beta = -0.95$, the estimated α_1 , α_2 , and β are 1.122 (0.004), 1.118 (0.004) and 1.130 (0.004), which are separated from their true values relatively more than their counterparts in Table 4, implying that restricting ρ_β increases the under-specification bias. However, when the true value is $\rho_\beta = 0.95$, the estimated α_1 , α_2 , and β are 1.104 (0.004), 1.085 (0.004) and 1.097 (0.004), which are substantively closer to their true values than the estimated coefficients in Table 4.

Second, when we compare the estimated variances, there are a few non-negligible differences between the estimates we find in Table 4 and Table 6. For instance, among all nine columns presented in Table 4, the largest and smallest estimated σ_2^2 are 0.391 (0.004) and 0.378 (0.004), which occur when the true correlations are $\rho_\beta = 0.35$ and $\rho_\beta = -0.35$, respectively. However, in Table 6, the estimated σ_2^2 monotonically increases from 0.351 (0.003) to 0.480 (0.004) when the true correlations are $\rho_\beta = -0.95$ and $\rho_\beta = 0.95$, respectively. Given that the true value is $\sigma_2^2 = 0.15$, the results in Table 4 and Table 6 suggests that restricting ρ_β increases the bias in the estimated σ_2^2 .

Similarly, the largest and smallest estimated σ_β^2 in Table 4 are 0.368 (0.003) and 0.385 (0.003), respectively. However, the estimated σ_β^2 in Table 6 monotonically decreases from 0.434 (0.003) to 0.150 (0.006) when the true correlations are $\rho_\beta = -0.95$ and $\rho_\beta = 0.95$, respectively. Therefore, we can conclude that restricting $\rho_\beta = 0$ adds more variation to the estimated variances σ_2^2 and σ_β^2 .

Finally, recall that the estimated mean and variance of WTP in Table 4 are mostly close to their true values. Specifically, in Table 4, only 4 out of 36 estimated WTP parameters are separated from their respective true values by more than 3 standard errors (i.e., 9 columns with 4 estimated WTP parameters each). However, when we compare the WTP results from Table 4 to the results we observe in Table 6, we note that restricting ρ_β increased the bias in the mean and variance of WTP. Specifically, 25 out of 36 estimated WTP parameters are separated from their respective true values by more than 3 standard errors.

Therefore, we can conclude that although mis-specifying the mixed logit model by restricting ρ_β may sometimes positively affect the estimated coefficients and variances, however, it negatively affects the estimated WTP parameters.

Table 6: Under Specification for Group 2 (Restricted ρ_β)

	$\rho_\beta = -0.95$	$\rho_\beta = -0.75$	$\rho_\beta = -0.55$	$\rho_\beta = -0.35$	$\rho_\beta = 0.35$	$\rho_\beta = 0.55$	$\rho_\beta = 0.75$	$\rho_\beta = 0.95$
$\alpha_1 = 1$	1.122 (0.004)	1.120 (0.004)	1.116 (0.004)	1.113 (0.004)	1.105 (0.004)	1.106 (0.004)	1.099 (0.004)	1.104 (0.004)
$\alpha_2 = 1$	1.118 (0.004)	1.110 (0.004)	1.112 (0.004)	1.113 (0.004)	1.104 (0.004)	1.102 (0.004)	1.096 (0.004)	1.085 (0.004)
$\beta = 1$	1.130 (0.004)	1.132 (0.004)	1.128 (0.004)	1.122 (0.003)	1.112 (0.003)	1.107 (0.003)	1.103 (0.003)	1.097 (0.004)
$\sigma_1^2 = 0.15$	0.470 (0.004)	0.455 (0.004)	0.438 (0.004)	0.421 (0.004)	0.334 (0.003)	0.295 (0.003)	0.234 (0.003)	0.115 (0.005)
$\sigma_2^2 = 0.15$	0.351 (0.003)	0.355 (0.003)	0.361 (0.004)	0.361 (0.004)	0.410 (0.004)	0.420 (0.003)	0.439 (0.004)	0.480 (0.004)
$\sigma_\beta^2 = 0.15$	0.434 (0.003)	0.429 (0.004)	0.422 (0.004)	0.409 (0.003)	0.337 (0.003)	0.296 (0.003)	0.250 (0.003)	0.150 (0.006)
True $E(WTP_{\alpha_1})$	-1.340	-1.300	-1.262	-1.225	-1.102	-1.070	-1.038	-1.008
True $E(WTP_{\alpha_2})$	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162
Est. $E(WTP_{\alpha_1})$	-1.221 (0.008)	-1.205 (0.007)	-1.192 (0.007)	-1.180 (0.006)	-1.114 (0.004)	-1.092 (0.004)	-1.058 (0.003)	-1.029 (0.002)
Est. $E(WTP_{\alpha_2})$	-1.157 (0.006)	-1.145 (0.005)	-1.150 (0.006)	-1.153 (0.006)	-1.144 (0.005)	-1.138 (0.005)	-1.130 (0.005)	-1.125 (0.005)
True $Var(WTP_{\alpha_1})$	1.427	1.167	0.942	0.749	0.262	0.165	0.084	0.015
True $Var(WTP_{\alpha_2})$	0.472	0.472	0.472	0.472	0.472	0.472	0.472	0.472
Est. $Var(WTP_{\alpha_1})$	0.767 (0.016)	0.708 (0.014)	0.651 (0.015)	0.583 (0.011)	0.319 (0.006)	0.232 (0.005)	0.142 (0.003)	0.045 (0.002)
Est. $Var(WTP_{\alpha_2})$	0.498 (0.010)	0.483 (0.009)	0.489 (0.011)	0.470 (0.010)	0.433 (0.008)	0.399 (0.008)	0.379 (0.007)	0.374 (0.008)

Like with Tables 4 and 6, Tables 5 and 7 present the Model 3 estimation (i.e., correct specification) results and the Model 1 estimation (i.e., restricted $\rho_\alpha = 0$) results of the Group 3, respectively. By comparing Tables 5 and 7, we infer the effects of restricting ρ_α on the mixed logit estimates apart from simulation bias. Similarly, there are three observations that we would like to point out.

First, we observe that restricting ρ_α does not have a substantive effect on the estimated coefficients α_1 , α_2 , and β . The estimated coefficients in Table 5 and Table 7 are mostly similar to each other. Second observation is that in Table 7, when we increase the true correlation ρ_α from 0.1 to 0.9, the estimated values of σ_1^2 and σ_2^2 both monotonically decrease from 0.369 (0.004) and 0.374 (0.004) to 0.172 (0.004) and 0.179 (0.003), respectively. However, in Table 5, the largest and smallest estimated σ_2^2 are 0.392 (0.004) and 0.383 (0.003), respectively, which is fairly constant compared to the estimated σ_2^2 in Table 7. Given that the true value in Table 7 is $\sigma_2^2 = 0.15$, these results suggest that restricting $\rho_\alpha = 0$ decreases the bias in the estimated σ_2^2 , especially when the true ρ_α is high.

However, when we compare the estimated σ_β^2 in Table 5 and Table 7, the results suggest that restricting ρ_α in the estimation of the mixed logit model negatively affects the estimated σ_β^2 . Specifically, the estimated σ_β^2 in Table 7 ranges from 0.377 (0.003) to 0.450 (0.003), however, in Table 5 it only ranges from 0.366 (0.003) to 0.382 (0.003). Given that the true value is $\sigma_\beta^2 = 0.15$, these results suggest that restricting $\rho_\alpha = 0$ increases the bias in the estimated σ_β^2 .

Finally, when we compare the estimated WTP parameters between Table 5 and Table 7, we see that restricting ρ_α substantially increases the bias in the mean and variance of WTP. For instance, in Table 5, there are only 7 out of 36 WTP parameters that are separated from their respective true values by more than 3 standard errors, with an average distance of 2.1 standard errors. However, in Table 7, there are 32 estimated WTP parameters that are biased, with an average distance 8.7 standard errors between the

true and estimated values. Therefore, we can conclude that omitting ρ_α in MSL estimation of the mixed logit model, negatively affects the mean and variance of WTP.

Table 7: Under Specification for Group 3 (Restricted ρ_α)

	$\rho_\alpha = 0.1$	$\rho_\alpha = 0.2$	$\rho_\alpha = 0.3$	$\rho_\alpha = 0.4$	$\rho_\alpha = 0.5$	$\rho_\alpha = 0.6$	$\rho_\alpha = 0.7$	$\rho_\alpha = 0.8$	$\rho_\alpha = 0.9$
$\alpha_1 = 1$	1.110 (0.004)	1.108 (0.004)	1.111 (0.003)	1.109 (0.004)	1.105 (0.004)	1.103 (0.004)	1.099 (0.003)	1.107 (0.004)	1.101 (0.004)
$\alpha_2 = 1$	1.106 (0.004)	1.114 (0.005)	1.116 (0.004)	1.109 (0.004)	1.109 (0.004)	1.102 (0.004)	1.101 (0.003)	1.106 (0.004)	1.100 (0.004)
$\beta = 1$	1.118 (0.004)	1.124 (0.004)	1.118 (0.004)	1.118 (0.003)	1.113 (0.004)	1.116 (0.004)	1.104 (0.004)	1.111 (0.004)	1.109 (0.004)
$\sigma_1^2 = 0.15$	0.369 (0.004)	0.364 (0.003)	0.352 (0.004)	0.325 (0.003)	0.315 (0.003)	0.291 (0.003)	0.268 (0.004)	0.226 (0.003)	0.172 (0.004)
$\sigma_2^2 = 0.15$	0.374 (0.004)	0.365 (0.003)	0.351 (0.003)	0.335 (0.003)	0.307 (0.003)	0.295 (0.003)	0.268 (0.003)	0.227 (0.003)	0.179 (0.003)
$\sigma_\beta^2 = 0.15$	0.387 (0.004)	0.377 (0.003)	0.386 (0.003)	0.396 (0.003)	0.412 (0.003)	0.413 (0.003)	0.429 (0.003)	0.431 (0.004)	0.450 (0.003)
True $E(WTP_{\alpha_1})$	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162
True $E(WTP_{\alpha_2})$	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162	-1.162
True $Var(WTP_{\alpha_1})$	0.472	0.472	0.472	0.472	0.472	0.472	0.472	0.472	0.472
True $Var(WTP_{\alpha_2})$	0.472	0.472	0.472	0.472	0.472	0.472	0.472	0.472	0.472
Est. $E(WTP_{\alpha_1})$	-1.147 (0.006)	-1.131 (0.005)	-1.141 (0.006)	-1.133 (0.005)	-1.138 (0.005)	-1.124 (0.005)	-1.133 (0.005)	-1.124 (0.006)	-1.117 (0.005)
Est. $E(WTP_{\alpha_2})$	-1.145 (0.005)	-1.138 (0.005)	-1.145 (0.006)	-1.136 (0.005)	-1.139 (0.005)	-1.124 (0.005)	-1.136 (0.005)	-1.124 (0.005)	-1.117 (0.005)
Est. $Var(WTP_{\alpha_1})$	0.444 (0.010)	0.411 (0.007)	0.418 (0.009)	0.392 (0.007)	0.407 (0.008)	0.374 (0.007)	0.380 (0.006)	0.345 (0.008)	0.331 (0.006)
Est. $Var(WTP_{\alpha_2})$	0.449 (0.009)	0.417 (0.007)	0.416 (0.007)	0.405 (0.008)	0.398 (0.007)	0.377 (0.007)	0.381 (0.006)	0.345 (0.007)	0.334 (0.006)

3.3. Over-specification Bias

Next, we analyze over-specification bias. Recall that in Table 3 we present the Model 1 estimation (i.e., correct specification) results of the Group 1 and analyze simulation bias. In Table 8, we present the Model 2 estimation (i.e., unrestricted $\rho_\beta \neq 0$) results for the Group 1. By comparing Tables 3 and 8, we infer the effects of unrestricting ρ_β on the mixed logit estimates apart from simulation bias.

There are three observations that we would like to point out. The first observation relates to the estimated coefficients α_1 , α_2 , and β . The second observation is related to the estimated variances σ_2^2 and σ_β^2 , as well as to the estimated correlation ρ_β . Finally, we analyze the effects on the estimated WTP parameters.

First, notice that the estimated coefficients α_1 , α_2 and β are biased in all nine columns presented in Table 8. However, these biases are not notably different than those presented in Table 3, which suggests that unrestricted $\rho_\beta \neq 0$ does not substantively affect the estimated coefficients α_1 , α_2 and β .

Second, when we compare the estimated variances σ_2^2 and σ_β^2 between Table 3 and Table 8, we can see that the estimated values in two tables share similar ranges. For instance, the estimated σ_2^2 in Table 8 ranges between 0.206 (0.010) and 0.503 (0.004). In Table 3, the estimated σ_2^2 ranges between 0.213 (0.005) and 0.500 (0.004). Similarly, the estimated σ_β^2 in Table 8 ranges from 0.340 (0.015) to 0.377 (0.004), while in Table 3 it ranges from 0.364 (0.011) to 0.380 (0.003). These results suggest that there may be slight differences in the estimated variances when we allow $\rho_\beta \neq 0$.

When we analyze the estimated ρ_β , we see that MSL produces biased results in some columns of Table 8. Specifically, in 5 out of 9 columns of Table 8, the estimated ρ_β is separated from 0 by more than 3 standard errors. For instance, when the true variances are $\sigma_1^2 = \sigma_2^2 = 0.15$ and $\sigma_\beta^2 = 0.05$, the estimated ρ_β is -0.029 (0.012), which is within 3 standard errors from its true value $\rho_\beta = 0$. However, when the true variances are $\sigma_1^2 = 0.05$, $\sigma_2^2 = 0.25$ and $\sigma_\beta^2 = 0.15$, the estimated ρ_β is -0.123 (0.020), which is

separated from its true value by more than 6 standard errors. Analysts may erroneously infer correlation due to over-specification.

Finally, although it seems like there are only slight differences in the estimated coefficients and variances of Table 3 and Table 8, those small differences add up and substantively affect the estimated WTP parameters in Table 8. Specifically, all 36 estimated WTP parameters in Table 3 are within 3 standard errors from their respective true values. However, in Table 8, 13 out of 36 estimated WTP parameters are separated from their true values by more than 3 standard errors, with an average distance of 5.83 standard errors between the true and estimated values. In other words, over-specification may cause biases in WTP estimates.

Like with Table 8, Table 9 presents the Model 3 estimation (i.e., unrestricted $\rho_\alpha \neq 0$) results of the Group 1. By comparing Tables 3 and 9, we infer the effects of unrestricting ρ_α on the mixed logit estimates apart from simulation bias. The results generally confirm that findings from Table 8 namely that allowing $\rho_\alpha \neq 0$ does not substantively affect the estimated coefficients α_1 , α_2 and β or the estimated variances σ_2^2 and σ_β^2 , but over-specification may cause biases in the WTP estimates.

Table 8: Over-specification Bias for Group 1 (Unrestricted ρ_β)

	$\sigma_1^2 = 0.05$ $\sigma_2^2 = 0.05$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.15$ $\sigma_2^2 = 0.05$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.25$ $\sigma_2^2 = 0.05$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.05$ $\sigma_2^2 = 0.15$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.15$ $\sigma_2^2 = 0.15$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.25$ $\sigma_2^2 = 0.15$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.05$ $\sigma_2^2 = 0.25$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.15$ $\sigma_2^2 = 0.25$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.25$ $\sigma_2^2 = 0.25$ $\sigma_\beta^2 = 0.15$
$\alpha_1 = 1$	1.115 (0.003)	1.108 (0.004)	1.112 (0.005)	1.112 (0.003)	1.104 (0.004)	1.107 (0.004)	1.117 (0.003)	1.108 (0.004)	1.108 (0.005)
$\alpha_2 = 1$	1.117 (0.003)	1.107 (0.003)	1.115 (0.003)	1.115 (0.004)	1.107 (0.004)	1.110 (0.004)	1.122 (0.005)	1.113 (0.005)	1.108 (0.005)
$\beta = 1$	1.121 (0.003)	1.113 (0.003)	1.114 (0.004)	1.120 (0.004)	1.112 (0.004)	1.109 (0.003)	1.123 (0.004)	1.113 (0.004)	1.119 (0.004)
σ_2^2	0.222 (0.003)	0.229 (0.006)	0.206 (0.010)	0.386 (0.003)	0.387 (0.004)	0.388 (0.003)	0.497 (0.004)	0.490 (0.004)	0.503 (0.004)
σ_β^2	0.374 (0.003)	0.376 (0.003)	0.356 (0.013)	0.367 (0.004)	0.377 (0.004)	0.376 (0.004)	0.340 (0.015)	0.374 (0.004)	0.374 (0.004)
$\rho_\beta = 0$	-0.060 (0.015)	-0.029 (0.012)	-0.058 (0.013)	-0.097 (0.020)	-0.042 (0.014)	-0.038 (0.012)	-0.123 (0.020)	-0.053 (0.015)	-0.042 (0.014)
True $E(WTP_{\alpha_1})$	-1.105	-1.162	-1.221	-1.105	-1.162	-1.221	-1.105	-1.162	-1.221
True $E(WTP_{\alpha_2})$	-1.105	-1.105	-1.105	-1.162	-1.162	-1.162	-1.221	-1.221	-1.221
True $Var(WTP_{\alpha_1})$	0.270	0.472	0.734	0.270	0.472	0.734	0.270	0.472	0.734
True $Var(WTP_{\alpha_2})$	0.270	0.270	0.270	0.472	0.472	0.472	0.734	0.734	0.734
Est. $E(WTP_{\alpha_1})$	-1.159 (0.004)	-1.158 (0.005)	-1.167 (0.007)	-1.160 (0.005)	-1.157 (0.006)	-1.163 (0.007)	-1.167 (0.005)	-1.160 (0.005)	-1.153 (0.008)
Est. $E(WTP_{\alpha_2})$	-1.097 (0.005)	-1.098 (0.005)	-1.105 (0.004)	-1.150 (0.006)	-1.154 (0.006)	-1.161 (0.006)	-1.213 (0.007)	-1.214 (0.007)	-1.208 (0.007)
Est. $Var(WTP_{\alpha_1})$	0.487 (0.010)	0.473 (0.009)	0.500 (0.012)	0.498 (0.012)	0.482 (0.010)	0.488 (0.012)	0.521 (0.013)	0.485 (0.010)	0.478 (0.014)
Est. $Var(WTP_{\alpha_2})$	0.255 (0.005)	0.266 (0.005)	0.266 (0.005)	0.442 (0.008)	0.461 (0.009)	0.466 (0.010)	0.701 (0.016)	0.699 (0.016)	0.719 (0.017)

Table 9: Over-specification Bias for Group 1 (Unrestricted ρ_α)

	$\sigma_1^2 = 0.05$ $\sigma_2^2 = 0.05$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.15$ $\sigma_2^2 = 0.05$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.25$ $\sigma_2^2 = 0.05$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.05$ $\sigma_2^2 = 0.15$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.15$ $\sigma_2^2 = 0.15$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.25$ $\sigma_2^2 = 0.15$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.05$ $\sigma_2^2 = 0.25$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.15$ $\sigma_2^2 = 0.25$ $\sigma_\beta^2 = 0.15$	$\sigma_1^2 = 0.25$ $\sigma_2^2 = 0.25$ $\sigma_\beta^2 = 0.15$
$\alpha_1 = 1$	1.118 (0.003)	1.111 (0.004)	1.118 (0.005)	1.116 (0.003)	1.107 (0.004)	1.110 (0.005)	1.122 (0.003)	1.112 (0.004)	1.112 (0.005)
$\alpha_2 = 1$	1.118 (0.003)	1.107 (0.003)	1.118 (0.003)	1.115 (0.004)	1.107 (0.004)	1.111 (0.004)	1.122 (0.005)	1.114 (0.005)	1.106 (0.005)
$\beta = 1$	1.121 (0.003)	1.114 (0.003)	1.116 (0.004)	1.121 (0.004)	1.115 (0.004)	1.112 (0.003)	1.125 (0.004)	1.115 (0.004)	1.121 (0.004)
σ_2^2	0.227 (0.004)	0.239 (0.004)	0.220 (0.008)	0.391 (0.004)	0.390 (0.005)	0.392 (0.004)	0.499 (0.005)	0.494 (0.005)	0.502 (0.004)
σ_β^2	0.376 (0.003)	0.375 (0.003)	0.370 (0.008)	0.374 (0.003)	0.380 (0.003)	0.376 (0.004)	0.357 (0.013)	0.378 (0.004)	0.374 (0.004)
$\rho_\alpha = 0$	0.059 (0.020)	0.057 (0.017)	0.076 (0.017)	0.052 (0.018)	0.036 (0.016)	0.056 (0.014)	0.029 (0.018)	0.031 (0.013)	0.001 (0.013)

True $E(WTP_{\alpha_1})$	-1.105	-1.162	-1.221	-1.105	-1.162	-1.221	-1.105	-1.162	-1.221
True $E(WTP_{\alpha_2})$	-1.105	-1.105	-1.105	-1.162	-1.162	-1.162	-1.221	-1.221	-1.221
True $Var(WTP_{\alpha_1})$	0.270	0.472	0.734	0.270	0.472	0.734	0.270	0.472	0.734
True $Var(WTP_{\alpha_2})$	0.270	0.270	0.270	0.472	0.472	0.472	0.734	0.734	0.734
Est. $E(WTP_{\alpha_1})$	-1.155	-1.155	-1.162	-1.153	-1.151	-1.158	-1.156	-1.156	-1.149
	(0.004)	(0.005)	(0.007)	(0.005)	(0.006)	(0.007)	(0.005)	(0.005)	(0.008)
Est. $E(WTP_{\alpha_2})$	-1.100	-1.099	-1.107	-1.153	-1.154	-1.161	-1.217	-1.216	-1.202
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.008)	(0.007)	(0.007)
Est. $Var(WTP_{\alpha_1})$	0.455	0.453	0.464	0.451	0.458	0.460	0.460	0.460	0.452
	(0.006)	(0.006)	(0.008)	(0.007)	(0.007)	(0.009)	(0.007)	(0.007)	(0.009)
Est. $Var(WTP_{\alpha_2})$	0.262	0.269	0.272	0.459	0.469	0.473	0.732	0.716	0.710
	(0.005)	(0.005)	(0.006)	(0.009)	(0.010)	(0.011)	(0.020)	(0.016)	(0.016)

4. Discussion

The evidence from this simulation study suggests that the MSL estimator will produce biased estimates of the mean coefficients, their variances and correlations irrespective of the underlying specification of the unobserved heterogeneity. These results suggest that the correct specification of the mixed logit model (i.e., simulation specification is the same as estimator specification) does not substantively mitigate biases in the estimated coefficients. However, a correctly specified mixed logit model can provide relatively accurate estimates of the WTP parameters. Restricting or unrestricted correlations erroneously can lead to further biases, particularly in the estimated variances (i.e., σ_2^2 and σ_β^2) as well as the estimated mean and variance of WTP.

Before estimating the mixed logit model based on the empirical DCE data, we highly recommend that choice modellers conduct simulations based on their intended specifications to assess the potential extent of simulation biases before relying on the MSL estimates of mixed logit parameters, particularly their variances and correlations. For example, some analysts may allow all parameters to be random and correlated; however, the findings from this paper may cast doubt about whether MSL estimators can reproduce results accurately.

To better illustrate the implication of these findings, we created a series of three plots comparing the true and estimated (WTP_{α_1}) for 500 Halton draws. In health preference research, often the primary aim is to estimate maximum acceptable risk (MAR) parameters or value sets on a quality-adjusted life year scale, instead of monetary units. These statistics are mathematically equivalent to $E(WTP_{\alpha_1})$.

Starting with Group 1 (No Patterns; Figure 1), the true $E(WTP_{\alpha_1})$ is represented with a black dot, and the estimated $E(WTP_{\alpha_1})$ for the Model 1, Model 2 and Model 3 methods are represented with blue, red and green dots, respectively. This plot illustrates that the correct specification of the mixed logit (i.e., Model 1) produces relatively better estimates of $E(WTP_{\alpha_1})$ than the two over-specified models (Models 2 and 3). In Group 2 (Figure 2), the correct specification of the mixed logit (i.e., Model 2) again produces relatively better estimates of $E(WTP_{\alpha_1})$ than the under-specified model (Model 1), depending on the extent of the correlation ρ_β .

The simulation biases may seem negligible in Groups 1 and 2, but the bias increases in Group 3 (substitution patterns), depending on the extent of the correlation ρ_α . Furthermore, erroneously restricting the correlation to zero (Model 1) leads to substantive under-specification biases in $E(WTP_{\alpha_1})$ regardless of the correlation. For analysts, we infer two additional lessons from Figure 3. Its simulation bias may be due to the binary nature of the data. Even when the model is correctly specified, controlling for the correlation between alternative specific constants may be challenging due to insufficient variability. Lastly, any choice task that includes an opt-out, status quo, or referent alternative likely has substitution patterns (e.g., a red bus may substitute for blue bus along the same route), and added care in the estimation of WTP is warranted due to potential simulation and under specification biases.

According to Palma et al. (2020), 72% among 150 papers indexed in the Research Papers in Economics (RePEc) produced in the period of 2008-2018 use less than 500 Halton draws, and only 5.6% use more than 2000 Halton draws. In this paper, we likewise used 500 Halton draws to be consistent with the levels used in the leading MSL applications.

Czajkowski and Budziński (2019) found that Halton draws are the second best performing quasi-random numbers after Sobol draws, and they suggested that approximately 2000 Halton draws (or 1170 Sobol draws) are necessary to have a Minimum Tolerance Level of 5% in a setting with 400 individuals, 4 choice tasks per individual, 3 alternatives, and 5 random parameters. Our DCE design consists of 200 individuals, 25 choice tasks per individual, 3 alternatives and only 2 random parameters. Therefore, for a robustness test, we believe that 3000 Halton draws would be sufficient. However, our analysis with 3000 Halton draws did not identify substantive differences, which may be an indicator of that our initial choice of the number of Halton draws (i.e., 500) is sufficient. The results with 3000 Halton draws can found in the Supplementary Materials.

In summary, Table 10 provides the average percentage distance between the true and estimated $E(WTP_{\alpha_1})$ and their ranges for 500 and 3000 Halton draws. On average, the simulation bias is less than 1 percent and misspecification biases ranged from 2.5 to 3.6 percent. In our previous simulation study (Jumamyradov et al., 2023) we showed that the conditional logit parameter estimates (α_1 , α_2 and β) may be biased depending on the structure of unobservable heterogeneity. Consequently, biases in the estimated parameters affect the estimated WTP, causing the WTP to be underestimated up to 20%. In combination with the results from this paper, some researchers may estimate the conditional logit model when their main interest is the coefficients, and estimate the mixed logit model when their main interest is WTP or similar parameters (e.g., quality-adjusted life years).

Table 10: Average Percentage Difference between the True and Estimated $E(WTP_{\alpha_1})$ and Their Ranges.

500 Halton draws			
	Group 1 No Patterns: $\rho_\beta = \rho_\alpha = 0$	Group 2 Taste: $\rho_\beta \neq 0, \rho_\alpha = 0$	Group 3 Substitution: $\rho_\beta = 0, \rho_\alpha \neq 0$
Model 1 No Patterns: $\rho_\beta = \rho_\alpha = 0$	Simulation bias: 0.795 [0.000, 1.979]	Under-specification bias: 3.315 [0.087, 8.881]	Under-specification bias: 2.496 [1.291, 3.873]
Model 2 Taste: $\rho_\beta \neq 0, \rho_\alpha = 0$	Over-specification bias: 3.593 [0.172, 6.540]	Simulation bias: 0.655 [0.087, 1.849]	
Model 3 Substitution: $\rho_\beta = 0, \rho_\alpha \neq 0$	Over-specification bias: 3.587 [0.090, 6.627]		Simulation bias: 0.833 [0.086, 2.065]
3000 Halton draws			
Model 1 No Patterns: $\rho_\beta = \rho_\alpha = 0$	Simulation bias: 0.678 [0.082, 1.201]	Under-specification bias: 3.201 [0.036, 9.269]	Under-specification bias: 2.421 [0.051, 3.886]
Model 2 Taste: $\rho_\beta \neq 0, \rho_\alpha = 0$	Over-specification bias: 3.634 [0.182, 7.135]	Simulation bias: 0.523 [0.001, 1.369]	
Model 3 Substitution: $\rho_\beta = 0, \rho_\alpha \neq 0$	Over-specification bias: 3.585 [0.014, 7.148]		Simulation bias: 0.768 [0.198, 1.518]

There are three primary limitations of this simulation study. First, the design of this DCE was selected for its simplicity and ability to mimic uninformative and informative heterogeneity, including taste and substitution patterns. The DCE is based on a single full-factorial design with three alternatives, 200 respondents and 25 choice sets. Its parsimonious model contained a constant term and a single variable for each alternative. It is unclear what would come out of a much larger study of a more complicated design and model, similar to the ones researchers ultimately intend to use. Nevertheless, the results suggest that the use of the mixed logit (when correctly specified) estimates WTP well, but not variances

and correlations. Further MSL implications along other purposes, such as market shares and substitution effects, are left for another day. Second, the data generation process in this study used only lognormal random parameters. Alternative distributions (e.g., symmetric or discrete) may lead to different results. Further work is necessary to investigate the implications of different distributions assumptions. Third, our specifications of unobservable heterogeneity allow only for interpersonal heterogeneity, however, some analysts may wish to account for both heterogeneity and intrapersonal variability (i.e., variability in a person’s parameters across tasks).

5. Conclusion

In this study, we illustrate the extent of biases in MSL estimates of mixed logit coefficients, variances and correlations using a standard DCE. We found that while estimating the mixed logit model with its correct specification may have little to none implications on the estimated WTP parameters, erroneously restricting or unrestricted correlations may lead up to 9% difference in the true and estimated WTP. Simulation and under specification biases are particularly detrimental to WTP estimates when alternative-specific constants are correlated (substitution patterns). Based on these findings, choice modellers may conduct similar simulations of their own DCE and assess the potential implications of these biases before relying on the MSL estimates of mixed logit parameters, particularly their variances and correlations.

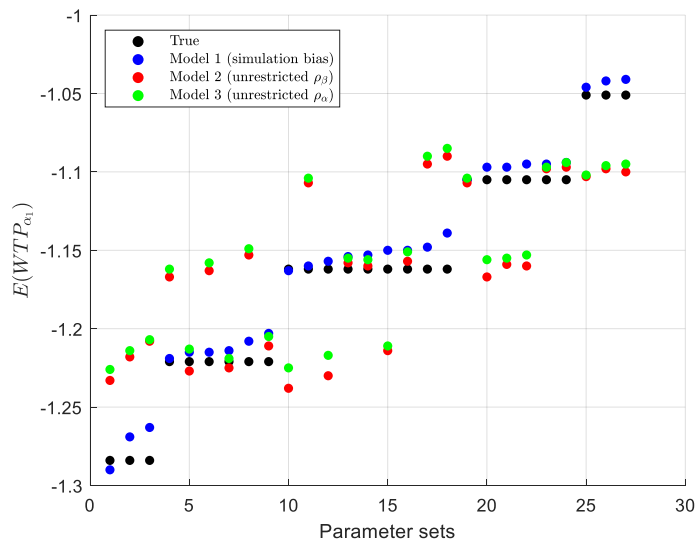


Figure 1: $E(WTP_{\alpha_1})$ by Model Specification in the Group 1 (no Pattern).

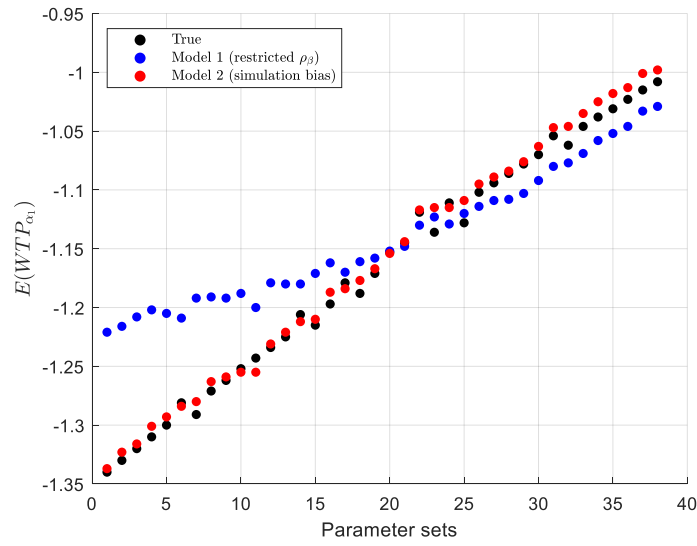


Figure 2: $E(WTP_{\alpha_1})$ by Model Specification in the Group 2 (Taste Patterns).

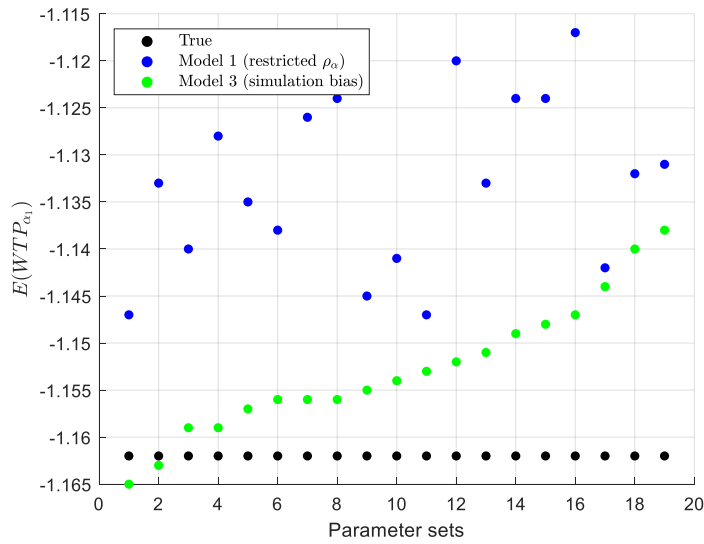


Figure 3: $E(WTP_{\alpha_1})$ by the Model Specification in the Group 3 (Substitution Patterns).

Appendix A: Experimental Design

The full-factorial design includes five out-of-pocket price levels {0.1, 0.4, 0.9, 1.6, 2.5}. Each individual $i = 1, \dots, N$ faced the same alternatives in each of the 25 choice situation $t = 1, \dots, T$.

Table A1: Out-of-pocket price

t	Alt. 1	Alt. 2	Alt. 3
1	0.1	0.1	0
2	0.1	0.4	0
3	0.1	0.9	0
4	0.1	1.6	0
5	0.1	2.5	0
6	0.4	0.1	0
7	0.4	0.4	0
8	0.4	0.9	0
9	0.4	1.6	0
10	0.4	2.5	0
11	0.9	0.1	0
12	0.9	0.4	0
13	0.9	0.9	0
14	0.9	1.6	0
15	0.9	2.5	0
16	1.6	0.1	0
17	1.6	0.4	0
18	1.6	0.9	0
19	1.6	1.6	0
20	1.6	2.5	0
21	2.5	0.1	0
22	2.5	0.4	0
23	2.5	0.9	0
24	2.5	1.6	0
25	2.5	2.5	0

References

- Brownstone, D. and K. Train (1999). 'Forecasting new product penetration with flexible substitution patterns.' *Journal of Econometrics*, 89, 109-129.
- Campbell, D. and S. Erdem (2019). 'Including opt-out options in discrete choice experiments: issues to consider.' *The Patient-Patient-Centered Outcomes Research* 12/1, 1-14.
- Clark, M. D., Determann, D., Petrou, S., Moro, D., and de Bekker-Grob, E. W. (2014). 'Discrete choice experiments in health economics: a review of the literature.' *Pharmacoeconomics*, 32(9), 883-902.
- Craig, B. M., Lancsar, E., Mühlbacher, A. C., Brown, D. S., and Ostermann, J. (2017). 'Health preference research: an overview.' *The Patient-Patient-Centered Outcomes Research*, 10, 507-510.
- Craig, B. M., de Bekker-Grob, E. W., González Sepúlveda, J. M., and Greene, W. H. (2022). 'A guide to observable differences in stated preference evidence.' *The Patient-Patient-Centered Outcomes Research*, 15(3), 329-339.
- Czajkowski, M., and Budziński, W. (2019). Simulation error in maximum likelihood estimation of discrete choice models. *Journal of choice modelling*, 31, 73-85.
- Daly, A., Hess, S., and Train, K. (2012). Assuring finite moments for willingness to pay in random coefficient models. *Transportation*, 39, 19-31.
- Gonzalez, J. M. (2019). 'A guide to measuring and interpreting attribute importance.' *The Patient-Patient-Centered Outcomes Research*, 12(3), 287-295.
- Gourieroux, C. and A. Monfort (1990). 'Simulation based inference in models with heterogeneity'. *Annales d'Economie et de Statistique*, 69–107.
- Greene, William H. (2012), 'Econometric analyses'. 7th edition, Prentice Hall, Upper Saddle River, NJ.
- Hajivassiliou, V. A. and P. A. Ruud (1994). 'Classical estimation methods for LDV models using simulation.' *Handbook of econometrics*, 4, 2383–2441.
- Hensher, D., Rose J., W. H. Greene (2005). 'Applied choice analysis: a primer'. Cambridge university press.
- Hossain, I., Saqib, N. U., and Haq, M. M. (2018). 'Scale heterogeneity in discrete choice experiment: An application of generalized mixed logit model in air travel choice'. *Economics Letters*, 172, 85-88.
- Jumamyradov, M., B. Craig, M. Munkin, and W. Greene (2023). 'Comparing the conditional logit estimates and true parameters under preference heterogeneity: a simulated discrete choice experiment'. *Econometrics*, 11 (1), 4.
- Jumamyradov, M., and Munkin, M. K. (2022). 'Biases in Maximum Simulated Likelihood Estimation of Bivariate Models.' *Journal of Econometric Methods*, 11(1), 55-70.
- Lee, L. (1995). 'Asymptotic bias in simulated maximum likelihood estimation of discrete choice models'. *Econometric Theory*, 437–483.
- Marschak, J., (1960). 'Binary choice constraints on random utility indicators', in K. Arrow, ed., *Stanford Symposium on Mathematical Methods in the Social Sciences*, Stanford University Press, Stanford, CA, 312-329
- McFadden, D. (1974). 'Conditional logit analysis of qualitative choice behavior', in P. Zarembka, ed., *Frontiers in Econometrics*, Academic Press, New York, 105-142.
- McFadden, D. and K. Train (2000). 'Mixed MNL models of discrete response'. *Journal of Applied Econometrics* 15, 447-470.

Palma, M. A., Vedenov, D. V., and Bessler, D. (2020). The order of variables, simulation noise, and accuracy of mixed logit estimates. *Empirical Economics*, 58, 2049-2083.

Quandt, E. Richard. 1970. *The Demand for Travel: Theory and Measurement*. Lexington, MA: D.C. Heath and Company.

Revelt, D. and K. Train (1998). 'Mixed logit with repeated choices'. *Review of Economics and Statistics* 80, 647–657.

Soekhai, V., de Bekker-Grob, E. W., Ellis, A. R., and Vass, C. M. (2019). 'Discrete choice experiments in health economics: past, present and future.' *Pharmacoeconomics*, 37(2), 201-226.

Train, K., (2009). *'Discrete choice methods with simulation'*. Cambridge university press.