OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG

WW FAKULTÄT FÜR WIRTSCHAFTSWISSENSCHAFT

Chair of Operations Management

Examiner : Prof. Dr. Sven Müller

Supervisor : Christoph Rippe

Scientific Project Paper

Master Scientific Project

Scientific Project: Mixed Logit Models

Date of Submission : 23.07.2021

Name: Pinto Bhusan Datta

Email: pinto.datta@st.ovgu.de

Study program: M.Sc. in Operations Research and Business Analytics

Contents

List of Abbreviations and Symbols

Abbreviations

MNL	Multinomial Logit
MXL	Mixed Multinomial Logit
i.i.d	Independent and Identically Distributed
IIA	Independent from Irrelevant Alternatives
ASC	Alternative Specific Constant
ASV	Alternative Specific Variance
LL	Log-Likelihood
SLL	Simulated Log-Likelihood
EV	Extreme Value Distributed
VTTS	Value of Travel Time Savings
WTP	Willingness to pay

Symbols

N	Normally Distributed
P	Probability

List of Tables

List of Figures

Chapter 1

Introduction

This scientific project explains the mixed multinomial logit model(MXL) different implementation approach with travel mode choice data. The multinomial logit model(MNL) is less performing as a heavily restricted assumption-based model in many practical situations. However, the MXL model is a less restricted assumption-based model for overcoming most basic MNL model limitations. The flexibility of the mixed multinomial logit model also brings many challenges to implement this model with the correct model specification and identification.

One of the earliest implementations of MXL has been successfully done by Boyd and Mellman (1980, p. 367), in the analysis of the market share of different types of automobiles concerning fuel economy improvement. McFadden and Train (2000, p.447-448, 458-463) have also established that with proper specification and mixing distribution, the mixed logit model is better than other discrete choice models to approximate a more complex utility model with the desired accuracy. We estimate the different types of MXL models with different interpretations. An overview of the discrete choice model, MXL model general form, random coefficient based MXL, simulation, error component structured MXL, substitution patterns, latent class logit model discuss in chapter 2. Overview of travel mode choice dataset, model specification and results of simple MNL model, error structure-MXL: nesting, cross nesting and alternative specific variance, random coefficient based MXL, and latent class logit model discuss in chapter 3. Finally, in chapter four, we discuss the conclusions.

Chapter 2

Mixed Logit Model

This chapter discusses an overview of the discrete choice model, MXL model general form, random coefficient based MXL, simulation, error component structured MXL, substitution patterns, latent class logit model.

2.1 An Overview of Choice Modeling

Discrete choice modelling is the process of establishing a choice model to predict the agent (i.e., person, firm, decision-maker) choice, based on their choice behaviour-related information, Train (2009, p.3). Following Koppelman and Bhat (2006, p.9), there are several factors associated with the choice decision of an individual decision process, and these processes can be generally explained by four components: decision-maker, available alternatives, attributes of alternative, and decision rule. According to the utility maximization decision rule, an individual n will choose an alternative i if and only if satisfy the condition of the equation 2.1,

$$U_{ni} \geq U_{nj} \tag{2.1}$$

where:

C_n**:** The choice set of available alternatives,

(i, j)**:** Alternatives from the choice set C_n,

U_{ni}: Utility of decision maker n for choosing alternative i from C_n,

U_{nj}: Utility of decision maker n for choosing alternative j from C_n.

It also needs to mention that the alternatives in the choice set C_n are mutually exclusive, exhaustive and finite. However, if these three conditions are violated, a more advanced logit model needs to be implemented in discrete choice analysis, which we discuss later.

2.1.1 Multinomial Logit Model

Following Train (2009, 34-41), in the choice decision process, every individual gets a certain amount of utility from the available alternatives. Generally, in a linear relationship, the amount of utility from a specific alternative depends on the weighting sum of alternatives attributes, individual/agent characteristics, and an unobserved amount of random utility. Mathematically, the utility of an alternative to an individual is the sum of the deterministic part of the utility and unobserved portion of utility which can express in the equation 2.2,

$$U_{ni} = V_{ni} + \varepsilon_{ni} \tag{2.2}$$

where:

U_{ni} is the utility of alternative i to decision maker n,
V_{ni} is the deterministic utility of alternative i to decision maker n,
i is the alternative from the available choice set,
ε_{ni} is the unobserved part of utility of alternative i to decision maker n,
n is the decision-maker or individuals.

As the researcher unaware about the ε_{ni}, the decision maker n, chosen an alternative i with a certain probability as in equation 2.3 which is a cumulative distribution, Train (2009, p. 19).

$$\begin{aligned}
P_{ni} &= P(U_{ni} > U_{nj}) \forall j \neq i \\
\implies P_{ni} &= P(V_{ni} - \varepsilon_{ni} > V_{nj} - \varepsilon_{nj} \forall j \neq i) \\
\implies P_{ni} &= P(V_{ni} - V_{nj} > \varepsilon_{nj} - \varepsilon_{ni} \forall j \neq i) \\
\implies P_{ni} &= P(V_{ni} - V_{nj} + \varepsilon_{ni} > \varepsilon_{nj} \forall j \neq i)
\end{aligned} \tag{2.3}$$

where:

P_{ni}: The choice probability under MNL of observation n for choosing alternative i,

C_n: The choice set of available alternatives,

(i, j): Alternatives from the choice set C_n,

The equation 2.3 provide the information that if the unobserved utility difference of two alternatives (ε_{nj} - ε_{ni}) is smaller than the observed portion of utility difference (V_{ni} - V_{nj}) then the decision maker n choose the alternative i, otherwise j. Though ε_{ni} is unobserved and unknown, it assumes that this term is independent and identically extreme value type 1 distributed and (ε_{nj} - ε_{ni}) logistic distributed. Moreover, if we have more than two alternatives in our choice set C_n then the choice probability under MNL derived from the following joint probability density as in equation 2.4,

$$P_{ni} = Pr\left(U_{ni} \geq \underset{j \in C_n/i}{max} U_{nj}\right) \tag{2.4}$$

where:

U_{ni}: is the utility of alternative i to decision maker n,

U_{nj}: is the utility of alternative i to decision maker n,

C_n: The choice set of available alternatives,

(i, j): Alternatives from the choice set C_n.

Thus, the equation 2.4 indicate that an individual chooses an alternative i from the available alternatives choice set C_n when the utility of U_{ni} is greater than or equal to the utility which has maximum utility among all other alternatives except the alternative i. In addition, this logic is the central concept to develop a multinomial logit model following a binary logit model. So, after the following equation 2.3 and property of unobserved random utility: the choice probability of choosing alternative i by the individual n under MNL can express as in equation 2.5:

$$P_{ni} = \frac{exp(\mu V_{ni})}{\sum_{j \epsilon C_n} exp(\mu V_{nj})} \tag{2.5}$$

where:

P_{ni}: The choice probability under MNL of observation n for choosing alternative i,

C_n: The choice set of available alternatives,,

μ: The parameter of the logistic distribution.

(i, j) are the alternatives from the available choice set,

V_{ni} is the deterministic utility of alternative i to decision maker n,

V_{nj} is the deterministic utility of alternative j to decision maker n,

However, the deterministic part of utility V_{ni} of equation 2.5 can express as in equation 2.6,

$$V_{ni} = \sum_{k=1}^{K} \beta_{ik} X_{nik} \tag{2.6}$$

where:

n is the decision maker or observations,
i is the alternative from the choice set,
k is the attribute related to alternative i,
V_{ni} is the deterministic utility of alternative i to decision maker n,
β_{ik} is the magnitude of utility that associated with the k^{th} attribute of the alternative i,
X_{nik} is the value of attribute k of the alternative i to the decision maker n.

Furthermore, since $\mu > 0$ is not identifiable, then μ is normalized to 1. So, the choice probability under multinomial logit of choosing alternative i by an individual n is expressed as in equation 2.7:

$$P_{ni} = \frac{exp(\sum_{k=1}^{K} \beta_{ik} X_{nik})}{\sum_{j \epsilon C_n} exp(\sum_{k=1}^{K} \beta_{jk} X_{njk})} \tag{2.7}$$

Some of the important limitations of the MNL model are as follows:

- Only the observed data could represent random taste heterogeneity,
- Restricted substitution pattern,
- The unobserved components of utility are independent and identically distributed(i.i.d),
- There is a limited application of panel data, due to i.i.d assumption.

2.2 Mixed Multinomial Logit Model

The MXL, one of the powerful advanced choice models other than the available discrete choice models, Hensher and Greene (2003, p. 133).
Following Train (2009, p.134-137), we discuss about MXL model general mathematical formulation. The MXL model can overcome most of the limitations of the MNL model. Also, the MXL model derivation with particular behavioural specifications provides a specific interpretation and can derive in various ways for different interpreting. Thus, a discrete choice model choice probability can express as in equation 2.8:

$$P_{ni}^{'} = \int P_{ni}(\beta) g(\beta|\Omega) d\beta \tag{2.8}$$

where:

$P_{ni}^{'}$ is the choice probability under MXL of the observation n for choosing alternative i,

$P_{ni}(\beta)$ is the choice probability under MNL of observation n for choosing alternative i,

β is the random parameter,

Ω is the vector or matrix of parameter of the random parameter β,

$g(\beta|\Omega)$ is the density function of random parameter β under the parameter Ω to be estimated.

The equation 2.8 is the general mathematical formulation of the MXL model, which is the choice probability that the observation n chosen alternative i.
In addition, from equation 2.8, the MXL choice probabilities are the result of integrals of MNL choice probabilities over the density $g(\beta|\Omega)$. If the observed portion of utility is dependent on and linear in random parameters β, then the mixed logit choice probability can express in the equation 2.9:

$$P_{ni}^{'} = \int \frac{exp(\sum_{k=1}^{K} \beta_{ik} X_{nik})}{\sum_{j \epsilon C_n} exp(\sum_{k=1}^{K} \beta_{jk} X_{njk})} g(\beta|\Omega) d\beta \tag{2.9}$$

In the equation 2.9, if the distribution of the random parameters β follows multivariate normal, then the parameter Ω represent the mean vector and variance-covariance matrix of parameters. The choice probability formulae of the MXL model in the equation 2.9 has no closed-form. Thus, the choice probability of MXL is the weighted average of MNL that evaluated at different values of β and weighted by the density $g(\beta|\Omega)$ at each value of β. In addition, the MXL choice probability in equation 2.9 is the function of Ω, indicate it does not depend on β because β is integrated out. Moreover, the parameters β are similar to the random term ε_{ni} where both are absent in choice probability because they are integrated out. So, the parameter Ω estimate from the MXL model but does not β.

2.2.1 Random Coefficients based MXL

Following Train (2009, p. 141), the most common use of the MXL model is based on random coefficient though it can derive in other ways for different interpretations. So, the random utility model with individual-specific coefficient is the most common way to implement the MXL model where the utility can be express as in equation 2.10:

$$U_{ni} = \beta_n X_{ni} + \varepsilon_{ni} \tag{2.10}$$

where:

U_{ni} is the utility of alternative i of observation n ($n = 1, ..., N; i = 1, ..., I$); N and I are total number of observations and alternatives respectively),

β_n is the individual-specific vector of coefficient or random parameter,

X_{ni} is the value of attribute of alternative i of observation n,

ε_{ni} is the i.i.d extreme value type 1 distributed.

In equation 2.10, the utility specification is the most general form of utility function under MXL to capture the individual-specific unobserved heterogeneity after considering random parameters for each individual, which is also called individual-specific coefficient β_n. This random coefficient can also be a non-linear function to estimate specific economic or socioeconomic effects from the data. So, the random parameter or individual-specific random coefficient β_n varies over decision-makers. So, the utility specification in the equation 2.10 can be any form of linear or non-linear function which determine by the researcher. Our discussion is limited to the most general form of model specification of random coefficient based MXL model.

The choice will be deterministic when the decision-makers are known to their own β_n and ε_{nj}. Generally, for any individual, β_n is unknown to them; that is why we need to consider β_n as a random parameter with density $g(\beta|\Omega)$. On the one hand, if β_n is known before to the researcher, then the logit model of choice probability will be as in equation 2.7. The known β_n means estimating the coefficient β_n solely on the observed data. On the other hand, If we assume that the coefficient β_n does not know before, then we have to integrate the deterministic MNL choice probability model over the density of the random coefficients to get the unconditional choice probability as in equation 2.8. In MXL model, considering alternative attributes and other consideration, the researcher determine the distribution of the random coefficients and estimate parameters of that distribution. A common way to implement MXL is to consider $\beta_n \sim N(\bar{\beta}, \sigma^2)$ which is similar to $\beta = \bar{\beta} + \sigma\alpha$ where $\alpha \sim N(0, 1)$. The mathematical formulation of random utility of the equation 2.10 are expressed as in equation 2.11,

$$U_{ni} = (\bar{\beta} + \sigma\alpha)X_{ni} + \varepsilon_{ni} \tag{2.11}$$

where:

U_{ni} is the utility of alternative i to observation n,
X_{ni} is the value of attribute of alternative i to n,
$\bar{\beta}, \sigma$ are the parameters to be estimated,
α is the random variable from $N(0, 1)$,
ε_{nj} is the i.i.d extreme value type 1 distributed.

Then, the MXL choice probability for choosing alternative i by observations n can be express in equation 2.12,

$$P''_{ni} = \int_{-\infty}^{+\infty} \frac{exp((\bar{\beta} + \sigma\alpha)X_{ni})}{\sum_{j=1}^{J} exp((\bar{\beta} + \sigma\alpha)X_{nj})} g(\beta|\bar{\beta}, \sigma)d(\beta) \tag{2.12}$$

where:
P''_{ni} is the choice probability under MXL of observation n for choosing alternative i,
β is the random coefficient,
$g(\beta|\bar{\beta}, \sigma)$ is the distribution of β and $\bar{\beta}, \sigma$ are the parameters to be estimated,
X_{ni} is the value of attribute of alternative i of observation n,

ε_{nj} is the i.i.d extreme value type 1 distributed.

The researcher is free to determine the distribution of $g(\beta|\bar{\beta}/\sigma)$ in the equation 2.12 based on the satisfaction of the expectation of the behaviour of the observations. There are some examples of probability distributions, random coefficient β can follow:

- Normal: $\beta \sim$ Normal(μ, Σ) and estimate $\Omega = \{\mu, \Sigma\}$
- Uniform: $\beta \sim$ Uniform(a, b) and estimate $\Omega = \{a, b\}$
- Triangular: $\beta \sim$ Triangular(a, b, c) and estimate $\Omega = \{a, b\}$

In addition, generally, the distribution of the random coefficient of an attribute depends on the distributional assumption of that attribute. For example, the travel time and the travel cost attributes parameter sign are generally negative because the increase of these attributes creates negativity towards most respondents. Thus, the negative lognormal distribution is more meaningful than the normal distribution for that attributes random coefficient in the MXL model. There are some properties of log-normal distribution are as follows:

- If a random variable X is log-normally distributed, then ln(X) is normally distributed
- If the random variables X and Y are log-normally distributed, then the ratio X/Y is also log-normally distributed
- If the random variables X and Y are log-normally distributed, then the product of X and Y also log-normally distributed

Moreover, the equation 2.12 has no closed-form, which means from this equation, we could not estimate the value of the parameters of the MXL model. So, the simulation technique needs to apply to estimate the parameters of the MXL model.

2.2.2 Simulation

Following Train (2009, p. 144), we discuss this subsection, the simulation method to estimate the parameters of the random coefficients of MXL. The utility function, as expressed in the equation 2.10, the random coefficient β_n probability density function is $g(\beta|\Omega)$ and parameter of this distribution is Ω. To estimate the parameter Ω, the

researcher determine the functional form of $g(\beta|\Omega)$. Due to the non-closed form of the equation 2.8, we could not estimate the parameter Ω. We can approximate the log-likelihood of MXL by using a numerical solution.
In the classical approach, the choice probability is approximated for any given value of Ω are as follows:

- Draw R random vector $\{\beta^1, \beta^2, ..., \beta^R\}$ of β from the probability density $g(\beta|\Omega)$.
- For each random vector of β , calculate logit choice probability $P_{ni}(\beta^R) = \frac{exp(\sum_{k=1}^{K} \beta_{ik} X_{nik})}{\sum_{j \epsilon C_n} exp(\sum_{k=1}^{K} \beta_{jk} X_{njk})}$,
- Over R random draws, taking an average of these choice probabilities as in equation 2.13,

$$\hat{P_{ni}} = \frac{1}{R} \sum_{R=1}^{R} P_{ni}(\beta^R) \tag{2.13}$$

where:

$\hat{P_{ni}}$ is the estimated logit choice probability under MXL of the decision maker n for choosing the alternative i,

R is the number of draws,

$P_{ni}(\beta^R)$ is the estimated logit choice probability under MNL of the decision maker n for choosing the alternative i in the R^{th} draw.

As the number of draws R increases, $\hat{P_{ni}}$ tends to the unbiased estimator of the true value of P_{ni} as well as the variance of $\hat{P_{ni}}$ will also decrease.
By inserting the simulated probabilities into the log-likelihood function, the simulated log-likelihood(SLL) calculate as in equation 2.14:

$$SLL = \sum_{n=1}^{N} ln(\sum_{i=1}^{I} h_{ni} \hat{P_{ni}}) \tag{2.14}$$

where:

N is the total number of observations,
I is the total number of alternatives,
$h_{ni} = 1$, if the alternative i is chosen,

$h_{ni} = 0$, otherwise.

The SLL value zero indicate the best model fit based on estimation of the parameters from the sample data.

2.2.3 Error Component Structured MXL

Following Hensher and Greene (2003, p. 135), in the MNL model, the error component of the utility function of the alternatives are independent and identically distributed(i.i.d.) extreme value type 1; for that reason, correlation does not allow over error component of alternatives. However, suppose the unobserved component of the utility of the alternatives is correlated. In that case, one of the ways to divide this error component into two parts where one part is i.i.d. but other parts allow correlation over alternatives and heteroskedastic as in equation 2.15:

$$U_{ni} = \beta_n X_{ni} + [\eta_{ni} + \varepsilon_{ni}] \tag{2.15}$$

where:

η_{ni} is the unobserved part of random utility with zero mean and its distribution depends on the other parameters and the data of the alternative i and individual n,

ε_{ni} is the remaining unobserved part of random utility, i.i.d. extreme value Type 1 distributed.

In MXL the distribution of η_{ni} can be any distribution such as normal, log-normal, and uniform etc.. The general form of probability density of η_{ni} is denoted by $g(\eta_{ni}/\delta)$ where δ is the fixed parameter. Thus, for any given value of η_{ni}, from the density $g(\eta_{ni}/\delta)$, the conditional choice probability of choosing alternative i by the observation n is express as in equation 2.16, since in the equation 2.15 ε_{ni} is extreme value type 1 distributed.

$$P_{ni}(\beta_n/\eta_{ni}) = \frac{exp(\beta_n X_{ni} + \eta_{ni})}{\sum_{j=1}^{J} exp(\beta_n X_{nj} + \eta_{nj})} \tag{2.16}$$

As, η_{ni} is taken from a probability distribution, so, unconditional choice probability of observation n for choosing alternative i is the integration of the equation 2.16

over the density $g(\eta_{ni}/\delta)$ for all possible values of η_{ni} can be express as in equation 2.17:

$$P'_{ni}(\beta_n/\delta) = \int P_{ni}(\beta_n/\eta_{ni})g(\eta_{ni}/\delta)d\eta_{ni} \tag{2.17}$$

The equation 2.17 is the general form of error component structured MXL where $g(\eta_{ni}/\delta)$ is the mixing distribution.

2.2.4 Substitution Patterns

Following Train (2009, p. 45 and 141), generally, in discrete choice analysis, we expect an increase of choice probability of an alternative due to ameliorating one or more attributes of that alternative. According to the statistical probability rules, the sum of choice probability of all available alternatives is one to an individual. So, if one alternatives choice probability increases, then all other or some alternatives probability will also be affected to keep the sum of probabilities of all alternatives to be one or the market share of all the alternatives are one or 100 per cent. It is important to know how an increase or decrease of one alternatives choice probability impacts other alternatives choice probabilities.

For example, in a city there exist three travel mode alternatives car, blue bus, and tram services with equal market share P_{car}, P_{blue_bus}, P_{tram} with value 1/3, 1/3, 1/3 respectively and the ratio of market share of car and blue bus is $\frac{P_{car}}{P_{blue\_bus}} = 1$ under the null MNL model. If a new alternative called red bus introduce, then the general expectation is that the market share of the blue bus will change more than tram and car. Since red bus and blue are different due to colour only, thus their market share will also be the same. As a result, now the four alternatives one possible new market share will be P_{car}, P_{blue_bus}, P_{red_bus}, P_{tram} with value 1/4, 1/4, 1/4, and 1/4 respectively, to keep constant the ratio of market share of car and blue bus as $\frac{P_{car}}{P_{blue\_bus}} = 1$ under the MNL model. However, blue bus and red bus alternative importance has increased together though these two are mostly same alternative. We could expect that the blue bus market share is greatly affected by the red bus, not others. This anomaly that means due to introducing the red bus, which is a close substitute to the blue bus, the market share of others alternatives also changes. So, it is important to know in discrete choice model analysis, attributes of one alternative improvement or introducing new alternatives, how to affect other

alternatives choice probability. The above issue may express in two approaches: first, restriction on the ratios of choice probabilities and/or second, restriction on the cross-elasticity of choice probabilities.
On the one hand, the ratio of choice probabilities for two alternatives i and l under MNL can be express as in equation 2.18,

$$\begin{aligned} \frac{P_{ni}}{P_{nl}} &= \frac{\frac{exp(V_{ni})}{\sum_{j\epsilon C_n} exp(V_{nj})}}{\frac{exp(V_{nl})}{\sum_{j\epsilon C_n} exp(V_{nj})}} \\ \implies \frac{P_{ni}}{P_{nl}} &= \frac{exp(V_{ni})}{exp(V_{nl})} \\ \implies \frac{P_{ni}}{P_{nl}} &= exp(V_{ni} - V_{nl}) \end{aligned} \tag{2.18}$$

where:

C_n**:** The choice set of available alternatives,

P_{ni}**:** The choice probability under MNL of observation n for choosing alternative i,

P_{nl}**:** The choice probability under MNL of observation n for choosing alternative l,

i, j, l are the alternatives from available choice set,

V_{ni} is the deterministic utility of alternative i to decision maker n,

V_{nl} is the deterministic utility of alternative l to decision maker n.

From the above equation 2.18, we have observed that the ratio of choice probabilities of alternatives i and l depends only on these two alternatives deterministic utility only that means this ratio is independent of other alternatives attributes change which is called independence from irrelevant alternatives or IIA. The MNL model follows the IIA property. On the other hand, the ratio of choice probabilities for two alternatives i and l under MXL can be express as in equation 2.19,

$$\frac{P_{ni}}{P_{nl}} = \frac{\int_{-\infty}^{+\infty} \frac{exp(V_{ni}(\beta))}{\sum_{j=1}^{J} exp(V_{nj}(\beta))} g(\beta|\bar{\beta}, \sigma) d(\beta)}{\int_{-\infty}^{+\infty} \frac{exp(V_{nl}(\beta))}{\sum_{j=1}^{J} exp(V_{nj}(\beta))} g(\beta|\bar{\beta}, \sigma) d(\beta)} \tag{2.19}$$

where:

P_{ni} is the choice probability under MXL of observation n for choosing alternative i,
P_{nl} is the choice probability under MXL of observation n for choosing alternative l,

β is the random coefficient,

$g(\beta|\bar{\beta}, \sigma)$ is the distribution of β and $\bar{\beta}, \sigma$ are the parameters to be estimated,

$V_{ni}(\beta)$ is the value of the deterministic utility of alternative i to the decision-maker n,

$V_{nl}(\beta)$ is the value of the deterministic utility of alternative l to the decision-maker n.

In equation 2.19, both choice probabilities P_{ni} and P_{nl} denominator also included in the integration parts, so they cannot cancel out. As a result, the ratio of choice probabilities of two alternatives depends on all other alternative's deterministic utility increase or decrease. So, the mixture of logit models does not follow the independence of irrelevant alternatives (IIA) property like MNL.

2.2.5 Latent Class Logit Model

Following Train (2009, p. 135-136), the MXL model is called latent class logit model if the random coefficient β in the equation 2.8 is a discrete random variable that means β take distinct finite value from the population. For example, In marketing, segments of customers are identifiable based on some observed characteristics of the customer. Thus, to implement MXL in marketing, we could think about the different values of β for different customer segments and estimate different MNL or MXL models for each segment. The researcher can determine or estimate the probability of a decision-maker belonging to a segment/class based on some characteristics of the decision-maker. Suppose, there exist m segment/class in the population where probability of an individual n belonging to a segment/class is denoted by Q_{nm} under the condition that $\sum_{i=1}^{m} Q_{nm} = 1$, $(m = 1, ..., M)$ where M is the total number of class/segment. So, the latent class choice probability can be express as in equation 2.20,

$$P_{ni}''' = \sum_{m=1}^{M} Q_{nm} \left(\frac{exp(d_m X_{nik})}{\sum_{j=1}^{J} exp(d_m X_{njk}} \right) \tag{2.20}$$

where:

P_{ni}''' is the choice probability of choosing alternative i by the decisiom maker n under latent class logit model,

d_m is the value of the parameter under of the m^{th} segment of k^{th} attributes or characteristics,

X_{njk} is the value of the j^{th} alternative of the k^{th} attribute,

J is the total number of alternatives.

Some of the parameters estimate for the latent class model only, and others are estimated for each segment/class MNL model separately. The log-likelihhod value also estimates for the latent class model as well as the separate MNL model of each segment. Following Greene and Hensher (2003, p. 682), many methods are available to estimate segment/class weight or probability of a decision-maker belonging to as segment/class. They have proposed the conditional choice probability under MNL to estimate Q_{nm} which can be express as in equation 2.21,

$$Q_{nm} = \frac{exp(\beta_m Y_{nm})}{\sum_{m=1}^{M} exp(\beta_m Y_{nm})} \tag{2.21}$$

where:

Q_{nm} is the probability/ weight that n^{th} individual belongs to m^{th} segment/class,
Y_{nm} is the vector of characteristics of individuals n that determine the class weight,
β_m is the estimable parameters vectors of m^{th} segment.

In practice, the MNL choice probability of the equation 2.21 estimated after considering decision-maker characteristics like age, gender, and income Etc., as explanatory variables. The MXL model with continuous random coefficient will fall in a deep chasm if the researcher fails to find out the proper distribution of random coefficients, Hensher and Greene (2003, p. 133). Thus, we can say that if there are identifiable segments or classes available, then the latent class logit model may be better than the random parameter based MXL model.

Chapter 3

Methods and Results Discussion

This chapter discusses an overview of travel mode choice dataset, model specification and results of simple MNL model, error structure-MXL: nesting, cross nesting and alternative specific variance, random coefficient based MXL, latent class model, and statistical test for IIA, and an overview of "apollo" R package.

3.1 Overview of Travel Mode Choice Data

In our discrete choice analysis, we use synthetic travel mode choice data from the apollo choice modelling website, where 500 travellers are considering journey options from four possible alternatives: car, bus, air, and rail. In this dataset, revealed preferences (RP) consider inter-city trips where at least two alternatives are available for each person. In addition, there are 14 stated preferences (SP) for each traveller after maintaining similarity with RP; available alternatives also exist. The SP alternatives are created artificially. The access time (except car alternatives), the travel time and travel cost variables for each alternative describe the trips where time and cost measure in minutes and pounds, respectively. The air and rail mode alternative also have a categorical variable for service attributes with three different labels named by no-frills, wifi available, or food available. Also, the information of each individual's gender, trip for business or not and income available in the data set, Palma (2021).

In our analysis, the synthetic data set of travel mode choice with 220 travellers with

3520 RP and SP observations have been taken from the original dataset. The travel time and travel cost for each alternative are considered as explanatory variables. In table 3.1, we show summary statistics of different alternatives choice situation.

Mode Choice Options	Chosen	Percent	Total Available
Car	1151	0.327	3520
Bus	147	0.042	3520
Air	810	0.230	3520
Rail	1412	0.401	3520
Total:	3520	1.00	14080

Table 3.1: **Summary of travel mode choice data**

In table 3.1, we observe that the decision-maker who choose rail and car travel mode alternatives are the highest and the second-highest in number, respectively. However, the number of decision-makers who choose bus travel mode alternative is a small portion of the total observations.

3.2 Statistical Programming Language R "apollo" package

Following the "apollo" package manual, Palma (2021), the apollo choice modelling package has different estimation steps to estimate the model parameters and other post-processing functions.
Generally, the following steps are follow in the estimation of any discrete choice model parameters:
Preparing Inputs Steps:.

- Initial Setting: In this step, the apollo library is loaded, and core controls like identification of individuals, mixing Etc., are needed to define.
- Data loading: Data is necessary to load with a specified named database.
- Model parameters definition: In this step, initialization of model parameters of all types of choice models and in the case MXL, mixing distribution of coefficients are needed to define.
- Input Validation: Previous sections defined inputs validation, check in this

section.

- Estimation Settings: Model specification settings for each alternative to calculate likelihood value of each observation.

Estimation and Output:

- Model estimation and report generation: If all the previous steps are executed correctly, then in this step, the parameter estimation process start with predefined settings, and the output shows after getting the optimum output value of the parameter at the final iteration.
- Post processing of the results: Prediction with existing and new data, model validation (testing), statistical test and many more task can perform after executing the model parameter estimation.

3.3 Test of Model Specifications

3.3.1 Informal Test

There are two common informal tests of model specifications stated below:

- Sign of the Coefficients
- Value of Trade-Offs

Sign of Coefficients

There are some expectations regarding model parameter estimated value in discrete choice analysis depending on the individual common response towards attributes. For example, generally, an increase in travel time and travel cost create negativity towards most respondents. Thus, we expect a negative parameter estimated value in a discrete choice model with travel time and travel cost attributes.

Value of Trade-Offs

One of the most common calculations of the value of trade-offs from the discrete choice model is the willingness to pay like the value of travel time savings(VTTS), Hensher and Greene (2003, p. 155). The VTTS define as the ratio of marginal utility of travel time and the marginal utility of travel cost. The VTTS value gives an idea that a decision-maker is willing to pay to equal the VTTS value for the one unit reduction of the travel time.

3.3.2 Statistical Test of IIA

Among many test procedure to test IIA property of a MNL model, one of them is Hausman-McFadden-Test are describe as follows:

- Estimate model with full set of alternatives attributes and characteristics parameters and variance-covarince matrix as: $\hat{\beta}_f$, $\sum_{\hat{\beta}_f}$
- Estimate model with subset of alternatives attributes and characteristics parameters and variance-covarince matrix as: $\hat{\beta}_s$, $\sum_{\hat{\beta}_s}$
- IIA exist: If the difference $\{\hat{\beta}_s - \hat{\beta}_f\} = 0$ is significant that is both model are different only by chance.

The test statistic of Hausman-Mc-Fadden Test express as in equation 3.1:

$$HM - Test\ Statisitc = A^{'}B^{-1}A \sim \chi^2_k \tag{3.1}$$

where:

$(\hat{\beta}_s - \hat{\beta}_f) = \mathbf{A}$

$(\sum_{\hat{\beta}_s} - \sum_{\hat{\beta}_f}) = \mathbf{B}$

k**:** is the number parameters of the sub-model.

IIA holds if the Hausman-Mc-Fadden Test statistic value is less than the critical value of χ^2_k.

3.4 Model 1: Multinomial Logit Model

Utility Specification of Model 1

The utility function of the four types of travel mode choice alternatives under the MNL model can be express in equation 3.2,

$$
\begin{aligned}
U_{car} &= \qquad\qquad \beta_{(time)}Tr.Time_{car} + \beta_{(cost)}Tr.Cost_{car} + \varepsilon_{car} \\
U_{bus} &= \beta_0\_bus + \quad \beta_{(time)}Tr.Time_{bus} + \beta_{(cost)}Tr.Cost_{bus} + \varepsilon_{bus} \\
U_{air} &= \beta_0\_air + \quad \beta_{(time)}Tr.Time_{air} + \beta_{(cost)}Tr.Cost_{air} + \varepsilon_{air} \\
U_{rail} &= \beta_0\_rail + \beta_{(time)}Tr.Time_{rail} + \beta_{(cost)}Tr.Cost_{rail} + \varepsilon_{rail}
\end{aligned}
\tag{3.2}
$$

where:

U_{car}, U_{bus}, U_{air}, and U_{rail} are the utility from the car, bus, air, and rail alternatives respectively,

β_0_bus, β_0_air, and β_0_rail are the alternative specific constant(ASC) of the bus, air, and rail alternatives respectively,

β_0_car, the alternative specific constant(ASC) of the car is set to zero for normalization,

$Tr.Time$ is the Travel Time attribute,

$Tr.Cost$ is the Travel Cost attribute,

$\beta_{(time)}$ is the generic travel time parameter,

$\beta_{(cost)}$ is the generic travel cost parameter,

ε_{car}, ε_{bus}, ε_{air}, and ε_{rail} are the unobserved random utility of car, bus, air, and rail alternatives respectively,

Estimated Result of Parameters of Model 1

Our discrete choice analysis starts from a simple MNL model with a generic travel time and a generic travel cost parameter. An alternative specific constant(ASC) for every four alternatives are considered, but the car alternative ASC is set to zero for normalization. The table 3.2 shows simple MNL model different parameter estimation results. The t-ratio is calculated by dividing the parameter estimated value by the standard error of the respective parameter. All the estimated parameter values are statistically significantly different from zero. In model 1, the bus alternative specific constant shows the highest negative value, indicating that individuals show the highest negativity towards the bus alternative compared to the car alternative if we consider other attributes and characteristics are remaining constant. The estimated parameter value of both the travel time and the travel cost is negative, indicating the individuals show negativity towards the increase in travel time and travel cost. Thus, we can say that utility of each alternative will increase with a decrease in travel time and travel cost of each alternative.

Model 1		
Parameter	Estimate	t-ratio(0)
$\hat{\beta}_0\_bus$	**-2.334178*** ** (0.097291)	-23.992
$\hat{\beta}_0\_air$	**-0.840535*** ** (0.156098)	-5.385
$\hat{\beta}_0\_rail$	**-0.663215*** ** (0.111678)	-5.939
$\hat{\beta}_{(time)}$	**-0.009903*** ** (0.00066)	-15.002
$\hat{\beta}_{(cost)}$	**-0.053783*** ** (0.001800)	-29.880
Model Description		
Number of individuals		220
Number of Observations		3520
Estimated Parameters		5
LL(final)		**-3679.413**
Adj. Rho-square(0)		0.245

Note: ***p<0.01, **p<0.05, *p<0.1,
$\hat{\beta}_0\_bus, \hat{\beta}_0\_air, \hat{\beta}_0\_rail$ are alternative specific constant estimated values for the bus, air and rail alternatives,
$\hat{\beta}_{(time)}$ is the travel time parameter estimated value,
$\hat{\beta}_{(cost)}$ is the travel cost parameter estimated value.

Table 3.2: **Model 1 Estimated Result of Parameters**

The standard errors of each estimated parameter mention in parenthesis. In the table 3.3, we estimate the parameters of the sub-model of model 1 for the test of the restricted substitution pattern of MNL, which is also called the IIA test. The null hypothesis of this test states that the original MNL model 1 and sub-model of model 1 estimated parameters are almost the same, and these values do not differ significantly. To create a sub-model from model 1, we drop the rail alternative from model 1 then estimate all parameter values under the MNL model. By following sub-section, 3.3.2, we calculate the Hausman-McFadden test statistic value, which is 55.76441. At one per cent level of significance, the critical value of χ^2 is 13.277 with four degrees of freedom, which is less than the calculated value of the test statistic. So, we could not retain the null hypothesis of the IIA test. Thus, we can say that our model 1 does not follow the IIA assumption of the MNL model, i.e. IIA fails.

Sub-Model of Model 1		
Parameter	Estimate	t-ratio(0)
$\hat{\beta}_0\_bus$	**-2.17427*****	-21.414
	(0.101534)	
$\hat{\beta}_0\_air$	**-1.37150*****	-6.777
	(0.202365)	
$\hat{\beta}_{(time)}$	**-0.01127*****	-12.632
	(0.00089)	
$\hat{\beta}_{(cost)}$	**-0.04773*****	-19.726
	(0.002420)	
Model Description		
Number of individuals		220
Number of Observations		2108
Estimated Parameters		4
LL(final)		**-1606.275**
Adj. Rho-square(0)		0.3047

Note: ***p<0.01, **p<0.05, *p<0.1,

Table 3.3: **Sub-Model of Model 1 estimation**

3.5 Mixed Logit Model: Error Component structured MXL

3.5.1 Model 2: Relaxing the Identical Distribution Assumption: Alternative Specific Variance Structured MXL

Utility Specification of Model 2

According to Greene and Hensher (2007, p. 620), "Importantly however, whereas the random parameters can account for differences across individuals and alternatives, the error components for alternatives and nests of alternatives focus is on the heterogeneity profile of additional unobserved effects associated with each alternative. The standard deviation parameters associated with each alternative capture this". Generally, to capture the mean effect of unobserved factors of an alternative, the alternative specific constant is introduced in the logit model. The estimation of the mean is not sufficient to capture the unobserved utility variability of each alternative. That is why in the MXL model, it is possible to estimate alternative specific variance (ASV) of each alternative unobserved utility to capture unobserved random heterogeneity, Walker et al. (2007, p. 1103).

In our analysis, one possible alternative specific variance structured MXL model utility specification can be express as in equation 3.3,

$$\begin{aligned} U_{car} &= V_{car} && + \varepsilon_{car} \\ U_{bus} &= V_{bus} + \sigma_{bus}\alpha_{bus} && + \varepsilon_{bus} \\ U_{air} &= V_{air} + \sigma_{air}\alpha_{air} && + \varepsilon_{air} \\ U_{rail} &= V_{rail} + \sigma_{rail}\alpha_{rail} && + \varepsilon_{rail} \end{aligned} \tag{3.3}$$

where:

V_{car}, V_{bus}, V_{air}, **and** V_{rail} are the deterministic part of utility of car, bus, air and rail alternatives respectively,

ASV parameter of the car alternative is set to zero for normalization,

σ_{bus} is the ASV parameter of the bus alternative,

σ_{air} is the ASV parameter of the air alternative,

σ_{rail} is the ASV parameter of the rail alternative,

$\alpha_{bus} \sim N(0, 1)$,

$\alpha_{air} \sim N(0, 1)$,

$\alpha_{rail} \sim N(0, 1)$,

$\varepsilon_{(i)} \sim EV$ i.i.d.

Following Walker, Ben-Akiva, and Bolduc (2007, p.1107), in alternative specific variance-based MXL model, ASV parameter normalization is not arbitrary. Generally, all possible models need to estimate to determine which alternatives ASV to be normalized. In practice, as a quick process, first estimate a model with the complete set of ASV parameters, then normalize the ASV parameter which one estimation value is the lowest that means minimum variance. They also showed, MXL model with two alternatives, none ASV parameters are identifiable. If the number of ASV parameters is three or more, then it is possible to identify the ASV parameter in the MXL model.

Result of estimated parameters of Model 2

In the table 3.4, all the estimated parameters are statistically significantly different

Model:	2-1 Unidentified (no constraints)	2-2	2-3	2-4	2-5
			Identified		
		σ car=0	σ bus=0	σ air=0	σ rail=0
Param.	Estim (SE)	Estim (SE)	Estim (SE)	Estim (SE)	Estim (SE)
$\hat{\beta}_0\_bus$	**-2.697*** **	**-2.701*** **	**-2.318*** **	**-2.758*** **	**-2.697*** **
	(0.16) ,	(0.16)	-(0.10)	(0.18)	(0.17)
$\hat{\beta}_0\_air$	**-1.001*** **	**-0.998*** **	**-0.994*** **	**-0.864*** **	**-1.0017*** **
	(0.18)	(0.17)	(0.18)	(0.17)	(0.18)
$\hat{\beta}_0\_rail$	**-0.717*** **	**-0.723*** **	**-0.713*** **	**-0.683*** **	**-0.718*** **
	(0.12)	(0.12)	(012)	(012)	(0.12)
$\hat{\beta}_{(time)}$	**-0.0108*** **	**-0.0106*** **	**-0.01079*** **	**-0.0104*** **	**-0.0108*** **
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\hat{\beta}_{(cost)}$	**-0.058*** **	**-0.0576*** **	**-0.0578*** **	**-0.0558*** **	**-0.0580*** **
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
$\hat{\sigma}_{car}$	**-0.507*** **	—	**-0.476*** **	**-0.618*** **	**-0.508*** **
	(0.07)	—	(0.07)	(0.063)	(0.06)
$\hat{\sigma}_{bus}$	**0.947*** **	**0.84862*** **	—	**1.116*** **	**0.9480*** **
	(0.16)	(0.15)	—	(0.16)	(0.16)
$\hat{\sigma}_{air}$	**0.762*** **	**-0.8754*** **	**-0.799*** **	—	**0.761*** **
	(0.08)	(0.08)	(0.076)	—	(0.08)
$\hat{\sigma}_{rail}$	**-0.0122**	**-0.291** **	**-0.003**	**-0.0603**	—
	(0.01)	(0.07)	(0.13)	(0.15)	—
SLL	**-3597.364**	**-3609.059**	**-3609.615**	**-3632.957**	**-3597.372**
Model description					
Estimated parameter		9			
Number ofindividuals		220			
Number ofobservations		3520			
Number of Draws		100(Halton)			

Note: ***p<0.01, **p<0.05, *p<0.1,
$\hat{\beta}_0\_bus, \hat{\beta}_0\_air, \hat{\beta}_0\_rail$ are alternative specific constant estimated values for the bus, air and rail alternatives,
$\hat{\beta}_{(time)}$ is the the estimated value of the travel time parameter,
$\hat{\beta}_{(cost)}$ is the the estimated value of the travel cost parameter,
$\hat{\sigma}_{car}$, $\hat{\sigma}_{bus}$, $\hat{\sigma}_{air}$, and $\hat{\sigma}_{air}$ are the estimated parameter value of the car, bus, and rail alternative ASV coefficient.

Table 3.4: **Model 2 Estimated Result of Parameters**

from zero at a 1 per cent level of significance except the rail ASV parameter value in models 2-1, 2-3, and 2-4. The estimated parameter result of model 2-1, the unidentified alternative specific variance-based MXL model, shows that the minimum variance alternative is rail. Due to illustration purposes, we have estimated four different models with a different base ASV of alternatives to find out a better fit model, Walker, Ben-Akiva, and Bolduc (2007, p. 1119). The results of the table 3.4 also shows that the arbitrary normalization of the ASV parameter does not support our models because the different base ASV parameters models simulated log-likelihood values are different. It is also found that the worst fit model is 2-4 when the base alternative is air, but in the unidentified model 2-1, the ASV parameter of the air alternative estimated value is the second-highest. Furthermore, in model 2-1, the alternative bus has the highest ASV estimated parameter value, but when it is set as the base alternative in the model 2-3 then it does not provide the highest loss fit model, which is contradictory to the principle of alternative specific variance model property as described by Walker, Ben-Akiva, and Bolduc (2007, p. 1117). In addition, identified and correctly specified model 2-5 SLL value is approximately equal to unidentified model 2-1, which interpretation is same as Walker, Ben-Akiva, and Bolduc (2007, p. 1119) " Note that while the unidentified model shows a slight improvement in fit over the identified models, it is not enough to estimate an additional parameter". So, this indicates that with less number of parameters, identified model 2-5 is better than unidentified model 2-1.

3.5.2 Model 3: Correlation among Alternatives: Nesting Structured MXL

Utility Specification of Model 3

The error component structured MXL, following the section 2.2.3, simple nested based MXL with non-overlapping two nests from the four types of travel mode choice alternatives can be express as one possible specification of utility as in equation 3.4,

$$
\begin{aligned}
U_{car} &= V_{car} + \sigma_{common\_transport}\alpha_{common\_transport} + \varepsilon_{car} \\
U_{bus} &= V_{bus} + \sigma_{common\_transport}\alpha_{common\_transport} + \varepsilon_{bus} \\
U_{air} &= V_{air} + \qquad\qquad\qquad\qquad\qquad\qquad + \varepsilon_{air} \\
U_{rail} &= V_{rail} + \sigma_{common\_transport}\alpha_{common\_transport} + \varepsilon_{rail}
\end{aligned}
\tag{3.4}
$$

where:

$\sigma_{common\_transport}$ is the nest parameter of the nest with car, bus, and rail alternatives,
σ_{air} is the alternative specific variance of the air alternative which set as zero for normalization,
$\alpha_{common\_transport} \sim N(0, 1)$,
$\alpha_{air} \sim N(0, 1)$,
$\varepsilon_{(i)} \sim EV$ i.i.d..

As part of the identification of the nest's parameter of the model with two nests based MXL, only one parameter is identifiable; that is why if we estimate both parameters, then both need to be identical, but in a non-overlapping nested model with more than two nests, all nest parameters are identifiable, and normalization is arbitrary, Walker, Ben-Akiva, and Bolduc (2007, 1109).

Results of estimated parameters of Model 3

The error component decomposition for identifying correlated alternatives is one of the tricks to implement the MXL model. On the one hand, we estimate the parameter of model 3 in table 3.5 after considering the same attributes as model 1. We consider two non-overlapping nests as one nest with air only and the other with car, bus, and rail alternatives. The results of table 3.5 show that all the estimated parameters are statistically significantly different from zero with a 1 per cent level of significance. The results also show that on the same set of data with two nest, with different normalization in model 3-1, 3-2 nest parameters estimated value $\hat{\sigma}_{commom_{transport}}$, $\hat{\sigma}_{air}$ are identical with the opposite sign but without normalization estimated value of the parameters $\hat{\sigma}_{commom_{transport}}$, $\hat{\sigma}_{air}$ in model 3-3 are not identical. However, all the three models simulated log-likelihood are almost the same. In two nest error component structured MXL model, one parameter is identifiable and possible to set arbitrarily for normalization. In addition, in the model with two nest parameter MXL, both parameters need to be equal in the unidentified model, Walker, Ben-Akiva, and Bolduc (2007, p. 1109).

So, we can say that our two nest error component structured MXL models as shown in the table 3.5 are not identically identified.

Model:	3-1	3-2	3-3
specification	(1, 0, 0, 0)	(0, 2, 2, 2)	(1, 2, 2, 2)
Parameter	Estim (SE)	Estim (SE)	Estim (SE)
$\hat{\beta}_0\_bus$	**-2.3612*****	**-2.3612*****	**-2.3562*****
	(0.0981)	(0.0981)	(0.0979)
$\hat{\beta}_0\_air$	**-1.000*****	**-1.0001*****	**-0.9985*****
	(0.1712)	(0.1712)	(0.1712)
$\hat{\beta}_0\_rail$	**-0.7083*****	**-0.7083*****	**-0.7088*****
	(0.1142)	(0.1142)	(0.1142)
$\hat{\hat{\beta}}_{(time)}$	**-0.0104*****	**-0.0104*****	**-0.0104*****
	(0.0006)	(0.0006)	(0.0006)
$\hat{\beta}_{(cost)}$	**-0.0570*****	**-0.0570*****	**-0.0570 *****
	(0.0019)	(0.0019)	(0.0019)
$\hat{\sigma}_{common\_transport}$	—-	**0.8521*****	**0.1391**
		(0.0741)	(0.4246)
$\hat{\sigma}_{air}$	**-0.8521*****	—-	**0.8397*****
	(0.0741)		(0.1024)
SLL	**-3621.789**	**-3621.789**	**-3621.899**
Model Description			
Number of individuals		220	
Number of Observations		3520	
Estimated Parameters		7	
Number of Draws		100(Halton)	

Note: ***p<0.01, **p<0.05, *p<0.1
$\hat{\beta}_0\_bus, \hat{\beta}_0\_air, \hat{\beta}_0\_rail$ are alternative specific constants for the bus, air and rail alternatives,
In specification row, 1, 0, 0, 0 indicate one nest with air alternative and others nest normalize to zero with car, bus, rail alternatives,
0, 2, 2, 2 indicate one nest normalized to zero with air and other nest with car, bus and rail alternatives,
1, 2, 2, 2 indicate one nest with air, and other with car, bus and rail,
$\hat{\sigma}_{common_{transport}}$ is the estimated value of the "common transport" nest parameter,
$\hat{\sigma}_{air}$ is the estimated value of the air alternative ASV parameter,

Table 3.5: **Model 3 Estimated Result of Parameters**

Utility Specification of Model 3a

Again, simple nested based MXL with non-overlapping three nests from the four types of travel mode choice alternatives can be express as one possible specification of utility as in equation 3.5,

$$\begin{aligned} U_{car} &= V_{car} + \sigma_{ma}\alpha_{ma} + \varepsilon_{car} \\ U_{bus} &= V_{bus} + \sigma_{ma}\alpha_{ma} + \varepsilon_{bus} \\ U_{air} &= V_{air} + \sigma_{air}\alpha_{air} + \varepsilon_{air} \\ U_{rail} &= V_{rail} + \sigma_{rail}\alpha_{rail} + \varepsilon_{rail} \end{aligned} \tag{3.5}$$

where:

σ_{ma} is the parameter of the nest called the "most available" nest with car and bus alternatives,

σ_{air} is the alternative specific variance parameter of the air alternative,

σ_{rail} is the alternative specific variance parameter of the rail alternative,

$\alpha_{ma} \sim N(0,1)$,

$\alpha_{air} \sim N(0,1)$,

$\alpha_{rail} \sim N(0,1)$,

$\varepsilon_{(i)} \sim EV$ i.i.d..

Results of estimated parameters of model 3a

In the table 3.6, we consider non-overlapping three-nested error structured MXL model, one nest with car and bus alternatives, another with air and the last one with rail alternative. We observe that without normalization of any nest parameter in mode 3a-1, all estimated nest parameter values are statistically significantly different from zero at a 1 per cent level of significance except $\hat{\sigma}_{rail}$. In model 3a-2, we normalize $\sigma_{rail} = 0$, and all the estimated parameters are statistically significantly different from zero at a 1 per cent level of significance. In the three nested error component MXL model, normalization is arbitrary, and all parameters are identifiable, Walker, Ben-Akiva, and Bolduc (2007, p. 1109). It is also observe from table 3.6 that unidentified model 3a-1 SLL value is almost equal to identified model 3a-2. Thus, with fewer parameters, identified model 3a-2 is better than unidentified model 3a-1. So, we can say that our nested error structured MXL model 3a-2 is identified and considering simulated log-likelihood value, model fitting is better than alternative specific variance structured MXL model 2 and two nested structured MXL models.

Model:	3a-1	3a-2
specification	(1,1,2,3)	(1,1,2,0)
Parameter	Estim (SE)	Estim (SE)
$\hat{\beta}_0\_bus$	**-2.360*****	**-2.361*****
	(0.098)	(0.098)
$\hat{\beta}_0\_air$	**-0.973*****	**-0.974*****
	(0.178)	(0.178)
$\hat{\beta}_0\_rail$	**-0.706*****	**-0.706*****
	(0.125)	(0.114)
$\hat{\beta}_{(time)}$	**-0.0108*****	**-0.0108*****
	(0.0006)	(0.0007)
$\hat{\beta}_{(cost)}$	**-0.0583*****	**-0.0583*****
	(0.002)	(0.002)
$\hat{\sigma}_{ma}$	**0.6002*****	**0.6001*****
	(0.07)	(0.07)
$\hat{\sigma}_{air}$	**-0.7250*****	**-0.7251*****
	(0.08)	(0.08)
$\hat{\sigma}_{rail}$	**0.00429**	—-
	(0.08)	
SLL	**-3595.260**	**-3595.261**
Model Description		
Number of individuals		220
Number of Observations		3520
Estimated Parameters		8
Number of Draws		100(Halton)

Note: ***p<0.01, **p<0.05, *p<0.1
item 1,1,2,3 indicate one nest with car, bus, and other nest with air and last one with rail,
1,1,2,0 indicate one nest with car, bus, and other nest with air and last one normalized to zero,
$\hat{\sigma}_{ma}$ is the estimated value of the "most available" nest parameter,
$\hat{\sigma}_{air}$ is the estimated value of the air alternative ASV parameter,
$\hat{\sigma}_{rail}$ is the estimated value of the rail alternative ASV parameter,

Table 3.6: **Model 3a Estimated Result of Parameters**

3.5.3 Model 4: Correlation among Alternatives: Cross Nesting structured MXL

Utility Specification of Model 4

Following Walker, Ben-Akiva, and Bolduc (2007, p. 1111, 1120), it is also possible to capture overlapping nested structured MXL which is called the cross nesting structured MXL. We show two overlapping nested structured MXL as in equation 3.6,

$$
\begin{aligned}
U_{car} &= V_{car} && + \sigma_{tc}\alpha_{tc} && && + \varepsilon_{car} \\
U_{bus} &= V_{bus} && && + \sigma_{mt}\alpha_{mt} && + \varepsilon_{bus} \\
U_{air} &= V_{air} && + \sigma_{tc}\alpha_{tc} && && + \varepsilon_{air} \\
U_{rail} &= V_{rail} && + \sigma_{tc}\alpha_{tc} && + \sigma_{mt}\alpha_{mt} && + \varepsilon_{rail}
\end{aligned}
\tag{3.6}
$$

where:

σ_{tc} is the parameter of travel comfort nest with car, air, and rail alternatives,
σ_{mt} is the parameter of mass transit nest with bus and rail alternatives,
$\alpha_{tc} \sim N(0, 1)$,
$\alpha_{mt} \sim N(0, 1)$,
$\varepsilon_{(i)} \sim EV$ i.i.d.

Following Walker, Ben-Akiva, and Bolduc (2007, p. 1111), in cross nesting MXL, no general rules need to follow for the identification and normalization.

Results of estimated parameters of Model 4

In model 4, we consider two overlapping error structured cross nesting where alternatives car, air and rail consider for the travel comfort nest, and bus and rail consider for the mass transit nest. We have taken 100 random halton draws from the normal distribution. We observe from the table 3.7, all estimated parameters are statistically significantly different from zero at a 1 per cent level of significance. In addition, the estimated value of the nest parameters show that in the travel comfort nest, there is a strong positive correlation among alternatives of that nest, and mass transit nest also shows a positive correlation among alternatives of that nests. The simulated log-likelihood value shows that cross nested MXL model 4 does not provide a better fit model than non-overlapping three nested MXL model 3a.

Model : 4				
Parameter	Estimates	t-ratio(0)	Rob. std. err.	Rob. t-ratio(0)
$\hat{\beta}_0\_bus$	**-2.7351***** (0.1619)	-16.893	0.1460	-18.727
$\hat{\beta}_0\_air$	**-0.8443***** (0.1572)	-5.370	0.1592	-5.303
$\hat{\beta}_0\_rail$	**-0.6701***** (0.1131)	-5.923	0.1170	-5.726
$\hat{\beta}_{(time)}$	**-0.00996***** (0.000609)	-14.940	0.000622	-14.560
$\hat{\beta}_{(cost)}$	**-0.0541***** (0.0018)	-29.74	0.00210	-25.65
$\hat{\sigma}_{tc}$	**0.970869***** (0.154774)	6.273	0.12557	7.732
$\hat{\sigma}_{mt}$	**0.158005** (0.101187)	1.562	0.10633	1.486
Model Description				
Number of individuals		220		
Number of Observations		3520		
Estimated Parameters		8		
LL(final)		**-3665.94**		
Adj. Rho-square(0)		0.2473		
Inter Individual Draws		100(Halton)		

Note: ***p<0.01, **p<0.05, *p<0.1,
$\hat{\beta}_0\_bus, \hat{\beta}_0\_air, \hat{\beta}_0\_rail$ are alternative specific constant estimated values for the bus, air and rail alternatives,
$\hat{\sigma}_{tc}$ is the estimated value of the travel comfort nest parameter,
$\hat{\sigma}_{mt}$ is the estimated value of the mass transit nest parameter.

Table 3.7: **Model 4 Estimated Result of Parameters**

3.6 MXL Model with Random Coefficient Technique

In the random coefficient based MXL, when the parameter value of an attribute shows the same sign to each individual generally, then the distribution of random coefficient follows a lognormal distribution. Consequently, in our analysis, we could expect travel time and travel cost parameter values as negative, indicating that one unit increase of those attribute values creates a negative utility towards the respondents. Thus, we try to fit that the negative of the travel time and travel cost random coefficients follows a lognormal distribution. On the one hand, in

Model 5, the travel time parameter is considered as random, the negative of random coefficient of travel time is lognormally distributed, but the travel cost parameter is fixed, utility specification can be express as in equation 3.7. On the other hand, in model 6, the travel cost parameter is considered as random, the negative of random coefficient of travel cost is lognormally distributed, but the travel time parameter is fixed, and utility specification can be express as in equation 3.8. Finally, in model 7, both parameters, travel time and travel cost, are considered as random, the negative of the random coefficients is lognormally distributed, and utility specification can be express as in equation 3.9.

3.6.1 Model 5: MXL model with travel time as random coefficient

Utility Specification of Model 5

$$U_{travel\_modechoice(i)} = \beta_{0}\_(i) + \left(\bar{\beta_{time}} + \sigma_{time}\alpha_{time}\right)Tr.Time_{(i)} + \beta_{cost}Tr.Cost_{(i)} + \varepsilon_{(i)} \tag{3.7}$$

where:

$Tr.Time$ is the Travel Time attribute,

$Tr.Cost$ is the Travel Cost attribute,

i**:** is chosen from the choice set {car, bus, air, and rail},

$\beta_{0(i)}$ is the alternative specific constant fot the i^th alternative,

$\beta_{0(1)}$**=0** , normalize the car alternative ASC,

$\bar{\beta_{time}}$**:** mean of normally distributed logarithmically of negative of travel time coefficient,

σ_{time}**:** standard deviation of normally distributed logarithmically of negative of travel time coefficient,

β_{cost}**:** fixed parameter of travel cost,

$\alpha_{time} \sim N(0, 1)$,

$\varepsilon_{(i)}$**:** random utility term.

Result of estimated parameters of model 5

As we have observed, Hausman-McFadden test also indicates that the IIA property fails in the case of our MNL model 1, so it is necessary to incorporate a more flexible assumption-based model.

Model: 5				
Parameter	Estimates	t-ratio(0)	Rob. std. err.	Rob. t-ratio(0)
$\hat{\beta}_0\_bus$	**-2.4272***** (0.09905)	-24.504	0.1044	-23.228
$\hat{\beta}_0\_air$	**-1.0753***** (0.1667)	-6.450	0.1632	-6.586
$\hat{\beta}_0\_rail$	**-0.7696***** (0.1182)	-6.509	0.1211	-6.353
$\hat{\hat{\beta}}_{(time)}$	**-4.5191***** (0.07411)	-60.977	0.0719	-62.816
$\sigma_{\hat{time}}$	**0.3894***** (0.0352)	11.047	0.0362	10.754
$\hat{\beta}_{(cost)}$	**-0.0582***** (0.0019)	-29.918	0.0022	-25.589
Model Description				
Number of Individuals		220		
Number of Observations		3520		
Estimated Parameters		6		
LL(final)		**-3569.334**		
Adj. Rho-square(0)		0.2673		
Number of Draws		100(Halton)		

Note: ***p<0.01, **p<0.05, *p<0.1,
$\hat{\beta}_0\_bus, \hat{\beta}_0\_air, \hat{\beta}_0\_rail$ are alternative specific constant estimated values for the bus, air and rail alternatives,

Table 3.8: **Model 5 Estimated Result of Parameters**

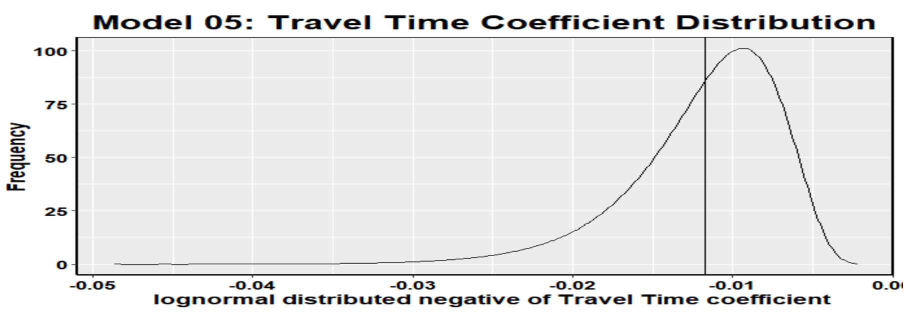

Figure 3.1: Coefficient Distribution in population(Unconditional) for Travel Time

In model 5, we consider the same attributes as model 1. This model considers the negative of travel time coefficient following lognormal distribution where travel cost is fixed. In the table 3.8, all the parameters estimated values are statistically signif-

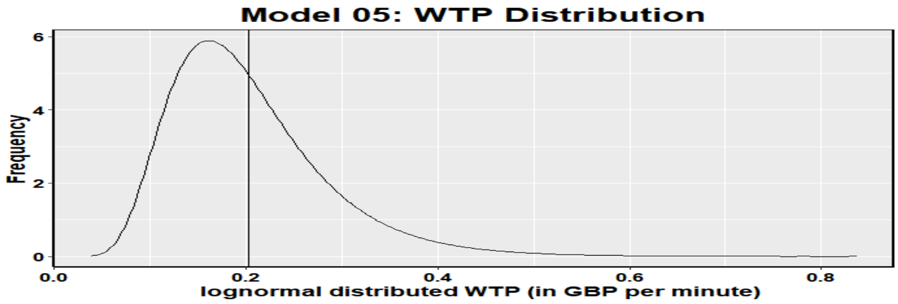

Figure 3.2: WTP distribution of the Model 05

icantly different from zero at a 1 per cent level of significance. The final simulated log-likelihood value is -3569.334. Also, in figure 3.1, we observe that the log-normally distributed negative of travel time coefficient distribution with mean -0.01173 and standard deviation 0.00476 are calculated from the unconditional distribution of travel time random coefficient.

Furthermore, in the figure 3.2, we observe that willingness to pay (WTP) distribution is the ratio of the log-normally distributed negative of travel time coefficient and non-random travel cost parameter value and also log-normally distributed. From the unconditional distribution of WTP of this model, we have found average willingness to pay is 12.1161 GBP for the one-hour reduction of existing travel time with a high standard deviation 4.8912 GBP per hour.

3.6.2 Model 6: MXL model with travel cost as random coefficient

Utility Specification of Model 6

$$U_{travel\_modechoice(i)} = \beta_0\_(i) + \beta_{time} Tr.Time_{(i)} + \left(\bar{\beta_{cost}} + \sigma_{cost}\alpha_{cost}\right) Tr.Cost_{(i)} \varepsilon_{(i)} \tag{3.8}$$

where:

i: is chosen from the choice set {car, bus, air, and rail},

$\beta_{0(1)}$=**0** normalize the car alternative,

$\bar{\beta_{cost}}$: mean of normally distributed logarithmically of negative travel cost coefficient,

σ_{cost}: standard deviation of normally distributed logarithmically of negative travel cost coefficient,

β_{time}: fixed parameter of travel time,

$\alpha_{cost} \sim N(0,1)$,

$\varepsilon_{(i)}$: is random utility term.

Result of estimated parameters of model 6

In model 6, we consider the negative of travel cost coefficient as random, follow a lognormal distribution. In table 3.9, we observe that all parameter estimated values are also statistically significantly different from zero at 1 per cent level significance.

Model: 6 Parameter	Estimates	t-ratio(0)	Rob. std. err.	Rob. t-ratio(0)
$\hat{\beta}_0\_bus$	**-2.7193***** (0.1094)	-24.839	0.1171	-23.219
$\hat{\beta}_0\_air$	**-1.1252***** (0.1659)	-6.779	0.1612	-6.977
$\hat{\beta}_0\_rail$	**-0.8101***** (0.1178)	-6.876	0.1183	-6.843
$\hat{\beta}_{(time)}$	**-0.0116***** (0.00071)	-16.252	0.00071	-16.347
$\hat{\hat{\beta}}_{(cost)}$	**-2.8467***** (0.0479)	-59.395	0.0415	-68.449
$\hat{\sigma}_{(cost)}$	**0.4806***** (0.0346)	13.888	0.0307	15.647
Model Description				
Number of Individuals		220		
Number of Observations		3520		
Estimated Parameters		6		
LL(final)		**-3544.456**		
Adj. Rho-square(0)		0.2724		
Number of Draws		100(Halton)		

Note: ***p<0.01, **p<0.05, *p<0.1,
$\hat{\beta}_0\_bus, \hat{\beta}_0\_air, \hat{\beta}_0\_rail$ are alternative specific constant estimated values for the bus, air and rail alternatives,

Table 3.9: **Model 6 Estimated Result of Parameters**

In addition, the simulated log-likelihood value of this model is -3544.456. Due to considering the randomness of the travel cost coefficient, the model fitting also improved than model 5.

Also, in figure 3.3, we observe that the log-normally distributed negative of travel

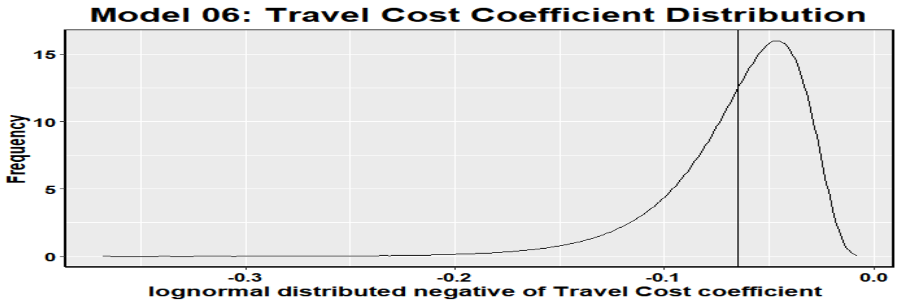

Figure 3.3: Coefficient Distribution in population(Unconditional) for Travel Cost

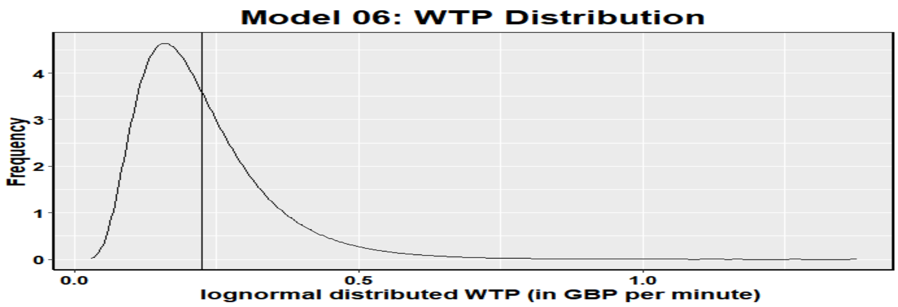

Figure 3.4: WTP distribution of the Model 06

cost coefficient distribution with mean -0.065111 and standard deviation 0.033083 are calculated from the unconditional distribution of that random coefficient. Furthermore, in the figure 3.4, we observe that willingness to pay (WTP) distribution which is the ratio of the non-random travel time parameter estimated value, and the log-normally distributed negative of travel cost coefficient is also log-normally distributed. From the unconditional distribution of WTP of our model, we have found average willingness to pay is 13.47565 GBP for the one-hour reduction of existing travel time with a high standard deviation value 6.8699 GBP per hour.

3.6.3 Model 7: MXL model with travel time and travel cost as random coefficients

Utility Specification of Model 7

$$\begin{aligned} U_{travel\_modechoice(i)} = \beta_0(i) + \left(\bar{\beta}_{time} + \sigma_{time}\alpha_{time}\right) Tr.Time_{(i)} \\ + \left(\bar{\beta}_{cost} + \sigma_{cost}\alpha_{cost}\right) Tr.Cost_{(i)} + \varepsilon_{(i)} \end{aligned} \tag{3.9}$$

where:

$\beta_{0(1)} = 0$: normalize the car alternative,

$\bar{\beta}_{time}$: mean of normally distributed logarithmically of negative travel time coefficient,

$\bar{\beta}_{cost}$: mean of normally distributed logarithmically of negative travel cost coefficient,

σ_{time}: standard deviation of normally distributed logarithmically of negative travel time coefficient,

σ_{cost}: standard deviation of normally distributed logarithmically of negative travel cost coefficient,,

$\alpha_{time} \sim N(0, 1)$,

$\alpha_{cost} \sim N(0, 1)$,

$\varepsilon_{(i)}$: random utility term.

Result of estimated parameters of model 7

Model: 7				
Parameter	Estimates	t-ratio(0)	Rob. std. err.	Rob. t-ratio(0)
$\hat{\beta}_0\_bus$	**-2.6648***** (0.1090)	-24.442	0.1146	-23.243
$\hat{\beta}_0\_air$	**-1.1767***** (0.1706)	-6.896	0.1663	-7.073
$\hat{\beta}_0\_rail$	**-0.8210***** (0.1206)	-6.803	0.1224	-6.704
$\hat{\bar{\beta}}_{(time)}$	**-4.4526***** (0.0676)	-65.867	0.0660	-67.421
$\hat{\sigma}_{(time)}$	**0.2686***** (0.0360)	7.445	0.0.0353	7.593
$\hat{\bar{\beta}}_{(cost)}$	**-2.8337***** (0.0445)	-63.661	0.0413	-68.591
$\hat{\sigma}_{(cost)}$	**0.4073 ***** (0.0345)	11.795	0.0295	13.774
Model Description				
Number of Individuals		220		
Number of Observations		3520		
Estimated Parameters		7		
LL(final)		**-3531.328**		
Adj. Rho-square(0)		0.2749		
Number of Draws		100(Halton)		

Note: ***p<0.01, **p<0.05, *p<0.1,

Table 3.10: **Model 7: Estimated Result of Parameters**

Finally, in model 7, we consider the MXL model with both negative of random coefficients travel time and travel cost are log-normally distributed and estimate parameters. In the table 3.10, we observe that the mean and standard deviation of normally distributed logarithmically of negative of travel time coefficient and logarithmically of negative travel cost coefficient and estimated all parameters are statistically significantly different from zero at 1 per cent level significance. The SLL value indicates that this model fit is better than model 5 and model 6. Also, in figure 3.5, we observe that the negative of travel time coefficient distribution is log-normally distributed with mean -0.0120738 and standard deviation 0.003297114 are calculated from the unconditional distribution of this random coefficient. In addition, in the figure 3.6, negative of travel cost coefficient distribution is log-normally distributed with mean -0.06385 and standard deviation 0.027072 are calculated from the unconditional distribution of this random coefficient. Furthermore, in the figure 3.7, we draw willingness to pay (WTP) distribution which is the ratio of the log-normally distributed negative of travel time coefficient, and the log-normally distributed negative of travel cost coefficient is also log-normally distributed. From the unconditional distribution of WTP of this model, we have found average willingness to pay is 13.39 GBP for the one-hour reduction of existing travel time with a high standard deviation value 6.945852 GBP per hour.

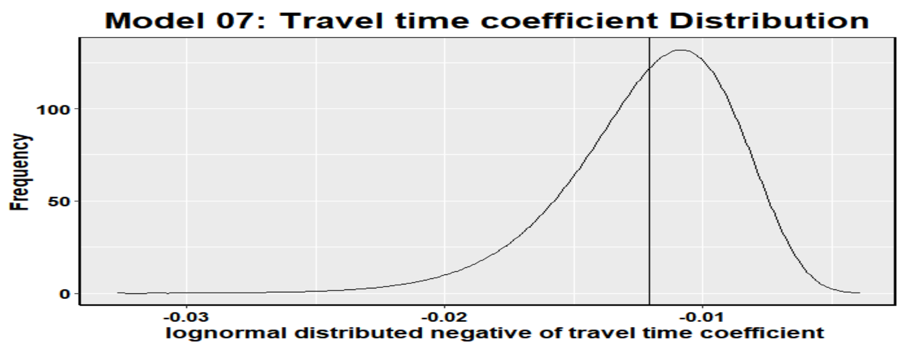

Figure 3.5: Coefficient Distribution in population(Unconditional) for Travel Time

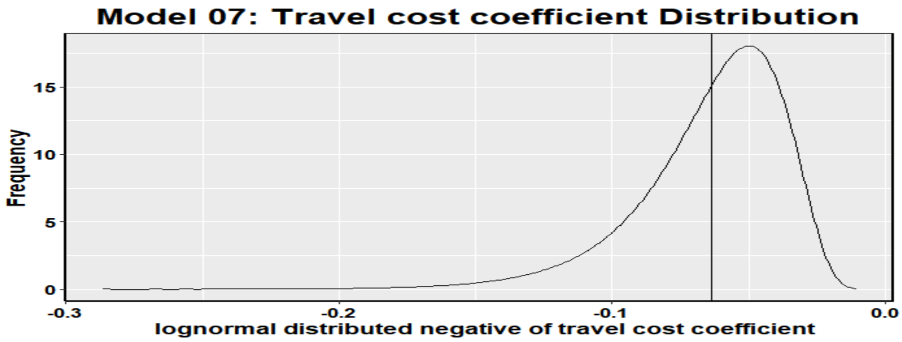

Figure 3.6: Coefficient Distribution in population (Unconditional) for Travel Cost

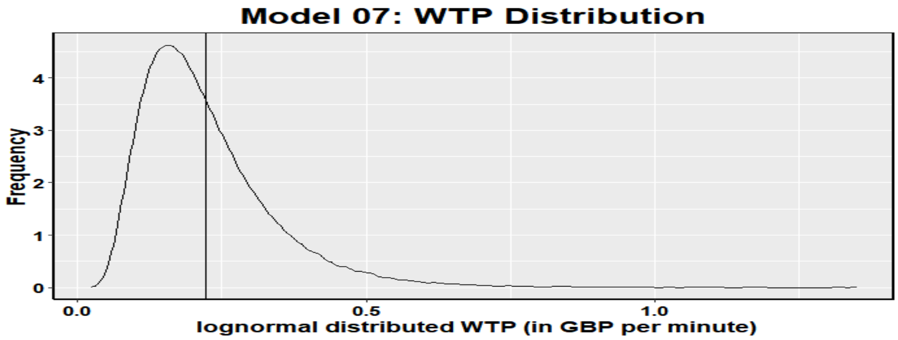

Figure 3.7: WTP distribution of the Model 07

Furthermore, in the figure 3.7, we draw willingness to pay (WTP) distribution which is the ratio of the log-normally distributed negative of travel time coefficient, and the log-normally distributed negative of travel cost coefficient is also log-normally distributed. From the unconditional distribution of WTP from our data, we have found average willingness to pay is 13.39 GBP per hour for the one-hour reduction of travel time with a high standard deviation of 6.945852 GBP per hour.

3.7 MXL Model with Discrete Random Coefficient: Latent Class Logit Model

3.7.1 Model 8: Latent Class Logit

Utility Specification of Model 8

Following the section 2.2.5, the simple latent class model with two classes will implement, class defining characteristics consider as an alternative specific constant. Utility specification under MNL to calculate the choice probability of belonging a decision-maker n in a class can express as in equation 3.10. In equation 3.11, we express the two utility specifications of two MNL models of class-A and class-B to build a latent class logit model where the class determining probability of individuals will calculate from the constant only MNL which utility specification is in equation 3.10.

Utility Specification of segment/class determining MNL model

$$\begin{aligned} Class_{(A)} &= 0 + \varepsilon_{(A)} \\ Class_{(B)} &= \delta_B + \varepsilon_{(B)} \end{aligned} \tag{3.10}$$

where:

δ_A is the ASC of the segment determining MNL model of the alternative $Class_{(A)}$ which is set to zero for normalization,

δ_B is the ASC of the segment determining MNL model of the alternative $Class_{(B)}$,

$\varepsilon_{(A)}$ is random utility term of class A,

$\varepsilon_{(B)}$ is random utility term of class B.

Utility Specification of Class A under MNL and Class B under MNL

$$\begin{aligned} U_{ModeChoice(i)} &= \beta_{0(i)} + \beta_{(time)A}Tr.Time_{(i)} + \beta_{(cost)A}Tr.Cost_{(i)} + \varepsilon_{(i)} \\ U_{ModeChoice(i)} &= \beta_{0(i)} + \beta_{(time)B}Tr.Time_{(i)} + \beta_{(cost)B}Tr.Cost_{(i)} + \varepsilon_{(i)} \end{aligned} \tag{3.11}$$

where:

i is indicate travel mode choice from choice set {car, bus, air, rail},

$\beta_{0(1)}$ ASC for car is set to zero for normalization,

$\beta_{(time)A}$ is generic fixed travel time parameter under MNL of class A,

$\beta_{(time)B}$ is generic fixed travel time parameter under MNL of class B,

$\beta_{(cost)A}$ is generic fixed travel cost parameter under MNL of class A,

$\beta_{(cost)B}$ is generic fixed travel cost parameter under MNL of class B,

$\varepsilon_{(i)}$ is random utility term.

Result of estimated parameters of Model 8

In model 8, we consider the same attributes of alternative as MNL model 1. In the table 3.11, we show the estimated value of the parameters. It is observed from the table that all estimated parameter values are statistically significantly different from zero at a 1 per cent level of significance.

Model: 8				
Parameter	Estimates	t-ratio(0)	Rob. std. err.	Rob. t-ratio(0)
$\hat{\beta}_0$_*bus*	**-2.53195*** ** (0.102635)	-24.670	0.107538	-23.545
$\hat{\beta}_0$_*air*	**-1.12960*** ** (0.166831)	-6.771	0.161759	-6.983
$\hat{\beta}_0$_*rail*	**-0.79850*** ** (0.118016)	-6.766	0.119175	-6.700
$\hat{\beta}_{(time)}$_A	**-0.01022*** ** (0.000741)	-13.785	0.000734	13.905
$\hat{\beta}_{(time)}$_B	**-0.01379*** ** (0.000942)	-14.621	0.001078	-12.792
$\hat{\beta}_{(cost)}$_A	**-0.07395*** ** (0.003523)	-20.993	0.004886	-15.135
$\hat{\beta}_{(cost)}$_B	**-0.04009*** ** (0.003716)	-14.389	0.003090	-12.975
$\hat{\delta}$_*A*	**0.0**	—	—	—
$\hat{\delta}$_*B*	**-0.46299** ** (0.204973)	-2.259	0.278359	-1.663
Model Description				
No of individuals		220		
No of Observations		3520		
Estimated Parameters		9		
LL(A)		**-3872.976**		
LL(B)		**-4070.982**		
LL(final)		**-3543.705**		
Adj. Rho-square(0)		0.2722		
Class_A Probability		**0.6137**		
Class_B Probability		**0.3863**		

Note: ***p<0.01, **p<0.05, *p<0.1
$\hat{\beta}_{(time)}$ is the estimated value of the travel time parameter,
$\hat{\beta}_{(cost)}$ is the estimated value of the travel cost parameter.

Table 3.11: **Model 8 Estimated Result of Parameters**

We consider the MNL model for each class to estimate the class-specific parameter's value. We have chosen for simplicity two latent classes, considering the alternative specific constant only MNL model to calculate the class probability/weight for each respondent. The class-A MNL model LL value is closer to the overall LL of the latent class logit model than class-B MNL. These results also indicate that there exist random taste preferences among the individuals. We also observe that alternative specific constant δ_A and δ_B are different in the segment/class determining MNL model. The MNL model of latent class A shows best fit than the MNL model of class B. The mean probability of class indicates a 61.37 per cent chance for an individual who belongs to class A, and 38.63 per cent belong in class B. The travel cost estimated parameter value indicates that class A respondents are more negative towards travel cost increase than class B. In respect of travel time, class B respondents are few more negative towards increase travel time.

Alternative specific constant for the bus alternative indicates that all the respondents are more negative towards bus alternative than car alternative considering that remaining other attributes and characteristics are constant.

Chapter 4

Conclusion

In this project, we have estimated MNL, MXL, and Latent class logit model. We have found that our basic MNL model with travel time and travel cost attributes fails to satisfy IIA property, leading us to explore a more flexible assumption-based discrete choice model. On the one hand, in our estimation, it is observed that the error component structured non-overlapping two nested MXL could not correctly fulfill the identification property. On the other hand, error component-based, non-overlapping three nested MXL shows a better fit. Also, we estimated the MXL model with travel time and travel cost where consider these attributes parameters jointly and separately as a random. MXL model with both parameters considering random shows better fit model than separately consider as a random. Our latent class logit model estimation shows that we can consider two latent classes to capture random taste heterogeneity. We have estimated different interpreting MXL models but did not get expected results in some models though we have the limited computational power to take random draws more than 100 draws.

We have compared all models based on willingness(WTP) to pay as shown in the table 4.1. The value within parenthesis indicates the standard deviation in the case of the MXL models.

We have observed that in MNL and Error structured MXL, WTP calculations are almost the same. In addition, random coefficient-based MXL models show higher WTP than MNL and error component-based MXL and show high standard deviations. In the latent class model, WTP for class A and class B are very different, but

	MNL	Error Component (2-5 - 4)			Random Coefficient (5-7)			LCM	
Model	1	2-5	3a-2	4	5	6	7	8 class-A	8 class-B
WTP	11.05	11.17	11.11	11.05	12.12 (4.89)	13.48 (6.87)	13.39 (6.95)	8.29	20.64

Table 4.1: **Model Comparison based on WTP**

class A value is within one sigma limit, and class B value within the two-sigma limit of random coefficient based MXL models WTP. In addition, the mean and standard deviation of WTP of both models 6 and 7 are pretty similar, where the travel cost parameter is considered random in both models.

Bibliography

J. H. Boyd and R. E. Mellman. The effect of fuel economy standards on the us automotive market: an hedonic demand analysis. Transportation Research Part A: General, 14(5-6):367–378, 1980.

W. H. Greene and D. A. Hensher. A latent class model for discrete choice analysis: contrasts with mixed logit. Transportation Research Part B: Methodological, 37(8):681–698, 2003.

W. H. Greene and D. A. Hensher. Heteroscedastic control for random coefficients and error components in mixed logit. Transportation Research Part E: Logistics and Transportation Review, 43(5):610–623, 2007.

D. A. Hensher and W. H. Greene. The mixed logit model: the state of practice. Transportation, 30(2):133–176, 2003.

F. S. Koppelman and C. Bhat. A self instructing course in mode choice modeling: multinomial and nested logit models. 2006.

D. McFadden and K. Train. Mixed mnl models for discrete response. Journal of applied Econometrics, 15(5):447–470, 2000.

S. H. . D. Palma. Apollo, 2021. `http://www.apollochoicemodelling.com/files/Apollo.pdf` [Last Accessed: 02.04.2021].

K. E. Train. Discrete choice methods with simulation. Cambridge university press, 2009.

J. L. Walker, M. Ben-Akiva, and D. Bolduc. Identification of parameters in normal error component logit-mixture (neclm) models. Journal of Applied Econometrics, 22(6):1095–1125, 2007.

Appendix A

Some Apollo R code

```
### "Scientific_Project_2021_Mixed_Logit

library("apollo")
library(ggplot2)
library(xtable)
#library(tidyverse)

###Model_07
###Random Parameter MXL
#Both Coefficient Of Travel Time
and Travel Cost are random
rm(list = ls())

library(apollo)### Initialise code
apollo_initialise()

### Set core controls
apollo_control = list(
  modelName ="Model_07new
␣␣(␣Mixed␣Logit__Both_Random_Coeficient__Vary)",
  modelDescr ="Mixed␣Logit___Random␣Coeficient,
␣␣intra-individual␣heterogeneity",
  indivID   ="ID",
  mixing    = TRUE,
  panelData = TRUE,
  nCores    = 5)
```

```
#### Load Data    ####
database = read.csv("ModeChoiceDataset.csv",
header=TRUE, sep = ";")

#### Define Model Parameters     ####

# Vector of parameters,including any that are kept
fixed in estimation
apollo_beta=c(ASC_car          = 0,
              ASC_bus          = 0,
              ASC_air          = 0,
              ASC_rail         = 0,
              mu_Travel_time              = -3,
              sigma_Travel_time           = 0,
              mu_Travel_cost              = -3,
              sigma_Travel_cost           = 0)

apollo_fixed = c("ASC_car")

####  Defien Random Components   ####

### Set parameters for generating draws

apollo_draws = list(
  interDrawsType = "halton",
  interNDraws    = 100,
  interUnifDraws = c(),
  interNormDraws = c("draws_sigma_Travel_time"
                   ,"draws_sigma_Travel_cost"),
  intraDrawsType = "halton",
  intraNDraws    = 100,
  intraUnifDraws = c(),
  intraNormDraws = c()
)

### Create random Coefficients

apollo_randCoeff = function(apollo_beta,
apollo_inputs){randcoeff = list()
```

```
randcoeff[["Travel_time"]] = -exp(mu_Travel_time
+ sigma_Travel_time * draws_sigma_Travel_time)
randcoeff[["Travel_cost"]] = -exp(mu_Travel_cost
+sigma_Travel_cost * draws_sigma_Travel_cost)

   return(randcoeff)
}


####  Group and Validate Inputs  ####

apollo_inputs = apollo_validateInputs()

####  Define Model and Likelihood   ####
apollo_probabilities=function(apollo_beta,
apollo_inputs,
functionality="estimate"){

#Attach inputs and detach after function exit
  apollo_attach(apollo_beta, apollo_inputs)
  on.exit(apollo_detach(apollo_beta,
  apollo_inputs))

  ### Create list of probabilities P
  P = list()

  ### List of utilities: these must use the
  same names as
  in mnl_settings, order is irrelevant
  V = list()
  V[['car']]  = ASC_car + Travel_time*time_car
  + Travel_cost * cost_car
  V[['bus']]  = ASC_bus + Travel_time*time_bus
  + Travel_cost * cost_bus
  V[['air']]  = ASC_air + Travel_time*time_air
  + Travel_cost * cost_air
  V[['rail']] = ASC_rail+ Travel_time*time_rail
  + Travel_cost * cost_rail
```

```
  ### Define settings for MNL model component
  mnl_settings = list(
    alternatives = c(car=1, bus=2,air=3,rail=4),
    avail        =
    list(car=av_car, bus=av_bus, air=av_air,
    rail=av_rail),
    choiceVar    = choice,
    V            = V
  )

  ### Compute probabilities using MNL model
  P[['model']] = apollo_mnl(mnl_settings,
  functionality)

  ###Average across intra-individual draws
  #P = apollo_avgIntraDraws(P, apollo_inputs,
  functionality)

  ### Take product across observation for
  same individual
  P = apollo_panelProd(P, apollo_inputs,
  functionality)

  ### Average across inter-individual draws
  P = apollo_avgInterDraws(P, apollo_inputs,
  functionality)

  ### Prepare and return outputs of function
  P = apollo_prepareProb(P, apollo_inputs,
  functionality)
  return(P)
}
################################
####   Model Estimation      ####

Model_07 = apollo_estimate(apollo_beta,
apollo_fixed, apollo_probabilities,
apollo_inputs)

####  Model Outputs   ####
```

```
apollo_modelOutput(Model_07)

apollo_saveOutput(Model_07)

#############################################
######Density Plot of Mixed Logit Model_07####

Model_07new.estimates <-read.csv("Model_07new
(Mixed␣Logit__Both_Random_Coeficient__Vary)
_estimates.csv")

xtable(Model_07new.estimates , digits = 4)

Model_07 = apollo_loadModel("Model_07new
(␣Mixed␣Logit__Both_Random_Coeficient__Vary)")

################ Unconditionals ################


unconditionals <- apollo_unconditionals
(Model_07, apollo_probabilities,
                            apollo_inputs)

#class(unconditionals)
df = as.data.frame(unconditionals)

dc <- density(as.vector(unconditionals
[["Travel_cost"]]))
mc = mean(unconditionals[["Travel_cost"]])
sc = sd(unconditionals[["Travel_cost"]])

rand_travel_cost <- as.vector(unconditionals
[["Travel_cost"]])
rand_travel_cost <- data.frame(rand_travel_cost)
mc<-mean(rand_travel_cost[,1])
sc<-sd(rand_travel_cost[,1])


ggplot(rand_travel_cost, aes(x=rand_travel_cost))
```

```
+  geom_density() +
  geom_vline(aes(xintercept = mc),
color="black", linetype="solid", size=0.1) +



 labs(title = "Model␣07:␣Travel␣cost␣coefficient
␣Distribution",x = "␣Negative␣lognormally␣travel
␣cost", y = "Frequency")+ theme(
# LABLES APPEARANCE
plot.title = element_text(size=20,
face= "bold", colour= "black", hjust = 0.5 ),
axis.title.x = element_text(size=14,
face="bold", colour = "black"),
axis.title.y = element_text(size=14,
face="bold", colour = "black"),
axis.text.x = element_text(size=12,
face="bold", colour = "black"),
axis.text.y = element_text(size=12,
face="bold", colour = "black"),
strip.text.x = element_text(size = 10,
face="bold", colour = "black" ),
strip.text.y = element_text(size = 10,
face="bold", colour = "black"),
axis.line.x = element_line(color="black",
size = 1),
axis.line.y = element_line(color="black",
size = 1),
panel.border = element_rect(colour = "black",
fill=NA, size=1)
)



#########################
####### WTP plot ########


WTP = as.data.frame(rand_travel_time
/rand_travel_cost)
```

```
WTP <- rand_travel_time/rand_travel_cost
WTP<-data.frame(WTP)

mw=mean(WTP[,1])
sw=sd(WTP[,1])

ggplot(WTP, aes(x=WTP[,1])) +
  geom_density() +
  geom_vline(aes(xintercept = mw),
    color="black", linetype="solid", size=0.1) +


 labs(title = "Model␣07:␣WTP␣Distribution",
        #subtitle = "Plot of length by dose",
        #caption = "Data source: ToothGrowth",
        x = "␣Log-normally␣WTP"
            ,y = "Frequency")+
  theme(
# LABLES APPEARANCE
plot.title = element_text(size=20,face= "bold",
colour= "black", hjust = 0.5 ),
axis.title.x = element_text(size=14,face="bold",
colour = "black"),
axis.title.y = element_text(size=14,face="bold",
colour = "black"),
axis.text.x = element_text(size=12,face="bold",
colour = "black"),
axis.text.y = element_text(size=12,face="bold",
colour = "black"),
strip.text.x = element_text(size = 10,face="bold",
colour = "black" ),
strip.text.y = element_text(size = 10,face="bold",
colour = "black"),
axis.line.x = element_line(color="black",
size = 1),
axis.line.y = element_line(color="black",
size = 1),
panel.border = element_rect(colour= "black",
fill=NA, size=1)
)
```