A Latent Class Model to Cluster Preferences

by

Krithika Raj Dorai Raj

5/13/2021

A thesis submitted to the
faculty of the Graduate School of
the University at Buffalo, The State University of New York
in partial fulfillment of the requirements for the
degree of

Master of Science
Department of Computer Science and Engineering

Abstract

In this study, we examine the heterogeneity in people's preferences to understand the different perspectives of fairness that exist, from survey responses. A latent class model was built to identify unobservable subgroups amongst the respondents to examine the specific preference patterns of each of the clusters formed. The latent class model computes the class membership probabilities of the responses and also estimates the preference parameters for each class using the expectation maximization algorithm. The model is used to fit the survey data i.e. the responses to survey questions requiring participants to choose the person they think is most fair to receive a particular service. The resulting clusters show how perceptions of fairness vary with the context of the service to be provided.

TABLE OF CONTENTS

List of Tables

List of Figures

Introduction

The motivation of this study is to understand people's perceptions of fairness in different contexts, by identifying subgroups amongst the responses that exhibit distinct patterns of preferences. A latent class model is built to cluster survey responses into subgroups where preference patterns within a subgroup are similar and the patterns across the subgroups differ.

Fairness is very subjective, different people have different ideas about what fairness is and this idea of fairness also varies with context [3]. It varies with who is asked the question, who the question is asked about and what the scenario is. Understanding people's perspectives of fairness can help inform decisions about ethical principles and guidelines [1]. This is crucial especially when countless decisions are constantly made by algorithms, that have far reaching implications.

Understanding preferences and examining the heterogeneity in preferences is very valuable in market analysis. Segmenting the customers based on their behavior and preferences allows businesses to personalize services or products to fit the needs of different types of customers. This type of analysis is also very popular in behavioral studies. It has been used extensively to study and classify child mental health [5], identify complex behavior patterns and variables that predict high-risk behavior patterns and identify the subgroups who are most at-risk for negative health consequences [6], study preferences of anglers with fishing license [2], etc.

This study uses survey responses where a respondent is given a particular scenario and a series of questions with two alternatives and is asked to choose the option that they think is most fair. By analyzing respondents' stated preferences, we can estimate the respondent's preferences and the importance associated with each of the attributes presented in the survey. The LCA model clusters these preferences and allows us to examine the various perspectives of fairness that exist amongst the population.

Related Work

The Moral Machine Experiment [1] is a study conducted by MIT to understand societal expectations about how autonomous vehicles should solve moral dilemmas. They estimate people's preferences by collecting survey responses where a participant is shown two scenarios that mimic the trolley problem and are asked to choose what they think the self-driving car should do. An example of the two alternatives presented is- an autonomous car with three passengers will kill the two pedestrians crossing the road ahead if it stays on course, and alternatively the car can swerve to avoid the two pedestrians but would hit a barrier and kill the passengers instead. Conjoint analysis is used to calculate the average causal marginal effects (AMCE) of each attribute. AMCE is a causal quantity that represents the increase in probability of sparing a character when the attribute value changes from a baseline value to another value of the attribute [4]. This quantity represents the population's moral preference to save a character with this attribute level over the other. For example, for the attribute- gender, it gives us the difference in the probability that a male character is spared and the probability that the female character is saved.

The two main results of this paper are the global preferences and cultural clusters. The global preferences consider responses from multiple countries to find the overall preferences. Cultural clusters are the formed by computing country-wise AMCE estimates- where all the responses from a country are used to compute a single AMCE value that represents its moral values. These AMCE values are then clustered using hierarchical methods. The paper reports three distinct cultural clusters with different moral preferences. This paper shows how surveys can be used to understand people's preferences. In the cultural cluster analysis, by assuming that the whole country can be summarized by a single AMCE value, the heterogeneity in preferences within the countries cannot be examined. Latent class methods allow us to examine the heterogeneity within a population.

Breffle et al. [2] proposes a joint latent class model that combines choice data and Likert scale preferences to examine heterogeneity in preferences. The latent choice model used in this thesis is the choice-only model described in this paper. This paper implements a joint latent class model on the survey that was conducted amongst anglers with a fishing license in the Green Bay waters. Two alternatives that varied by the type of fish available, the catch rate or contamination levels in fish were presented. They report four classes of preference patterns in the responses. The estimated class membership probability, the cluster size and the distinct preferences of the respondents in each class were also reported. Using the respondent's attributes (gender, income, etc.) as the clustering variables, the distribution of the respondents amongst the classes by their attributes was obtained. This enabled the authors to analyze the types of anglers in each class and analyze the features or services that were most important to them. In order to analyze the different perceptions of fairness in this study in a similar manner, a latent class model is developed from the choice-only model presented in [2].

This thesis uses the data collected by the Hannan et all. paper [3]. There are two research questions of this paper that are relevant to the study- how fairness perceptions vary based on the particular service being allocated and how it varies based on the demographics of the participants. In this thesis, we use latent class analysis on the data to uncover the latent classes amongst the preferences identifying distinct responses patterns and the preferences associated with them. The survey conducted by the authors [3] presents the respondent with a service allocation scenario and asks a series of choice questions. The respondent is asked to choose an individual they think it is most fair to allocate the service to. Marginal means are computed for the attributes varied in the alternatives, representing the probability of a service being allocated to a person with a particular attribute level relative to the other levels of the attribute [3].

The patterns of preferences clustered together by the latent class model in this thesis aligns with the preference patterns reported in [3]. An example of these results is- older people were the least likely to be given the services of tuition aid and life-saving devices, but were most likely to be given pain reduction medicine or housing services. It was also observed that people with mental health issues were less likely to be given life saving devices when compared to healthy people, whereas people with physical disabilities or diabetics were more likely to receive the same services.

Latent class analysis is a popular statistical method used in many market segmentation analysis and behavioral studies. It has been used to study citizens' preferences regarding public transport innovations [7], identify the factors that influence consumers' choices when buying mobile phones [8] and to understand student preferences for curriculum planning for college courses [9]. In all these papers, the latent class analysis is performed by licensed software like Sawtooth [7], Latent Gold [9]. Since all existing tools for LCA required a license, an open source latent class model for discrete choice experiments was implemented in R for this study.

Data

The data used in this study are survey responses [3]. Participants were presented with a scenario and were asked a series of 10 questions. A sample question [3] is shown in table 1. The scenario presented to the respondents were varied by service context and severity. The two service contexts used were social work and Covid-19 and these contexts varied with high or low severity. The four possible scenarios are housing availability for people without a place to live (social work – high severity), tuition assistance for courses at a local community college (social work – low severity), new treatment or ventilator that improves survival odds by 50% (Covid-19 – high severity) and medical device that alleviates mild discomfort (Covid-19 – low severity). The persons presented as alternatives in the survey are varied by these attributes- race, age, criminal history, health issues, occupation, upbringing and political affiliation. The levels of each of the attributes, as shown in the paper [3] is presented in table 2. The demographic attributes of the respondents used in this study is their political affiliation (Democrat, Republican and Independent/Other).

Table 1: An example of the questions asked in the survey [3]

You are a social worker that provides a small tuition stipend to individuals who want to take courses at the local community college. If you had to choose, which of the two individuals would you say it is most fair to provide stipend to?		
Attribute	**Person A**	**Person B**
Age	20	70
Children	Two kids	Two kids
Occupation	Doctor	Nurse
Criminal history	Prior history of non-violent crime	Prior history of violent crime
Race	Hispanic	Hispanic
Upbringing	Grew up poor	Grew up middle class
Political affiliation	Democrat	None specified
Health Issues	Generally healthy	Mental health issues

Table 2: The attributes allocated to hypothetical individuals in the survey [3]

Attribute	N	Attribute Levels
Race	4	White, Black, Asian, Hispanic
Age	3	20, 40, 70
Criminal history	3	None, Prior history of non-violent crime, Prior history of violent crime
Health Issues	4	Generally healthy, Mental Health Issues, Physically Disabled, Diabetic
Occupation	15	janitor, nurse, doctor, unemployed, firefighter, Instacart shopper, artist, politician, banker, scientist
Children	2	No kids, 2 kids
Upbringing	3	Grew up poor, Grew up middle class, Grew up upper class
Political affiliation	3	Democrat, Republican, None Specified

Methodology

Discrete Choice Models

The choices presented in the survey are modeled in this work by the Random utility model (RUM). The description of the RUM model below follows the notations used in the book by Kenneth E. Train [10]. RUM assumes that the decision maker makes a choice that maximizes the utility. The decision maker gets some utility from each alternative, and picks the alternative that provides the greatest utility. The utility (U) is not directly observed. Instead, it is inferred from the information that we know: the alternative that is chosen, the attributes of the alternatives that were presented, some attributes of the decision maker. It also includes some unobservable utility. The estimated utility along with the unobservable utility is used to predict what the total utility must have been, and the effect of each attribute on the total utility U. Train [10] notes that models derived from RUM are consistent with utility maximization, i.e. even if the user doesn't make the decision that maximizes the utility, the choices are still consistent with the concept of utility maximization.

The total utility (U_{ij}) a respondent *i* gets from alternative *j* can be decomposed into two components [10]:

1. Representative Utility (V_{ij}): is the utility of the observed factors i.e. the attributes of the alternatives and the decision maker.
2. The utility of the unobserved factors (E_{ij}): everything that affects the utility that is not included in V_{ij}. This is treated as the random variable in the RUM.

The expression for the total utility for i th respondent and alternate *j*:

$$\text{Eq 1.} \quad U_{ij} = V_{ij} + E_{ij}$$

The representative utility is specified as linear function and is represented as:

$$\text{Eq 2.} \quad V_{ij} = \beta' x_{ij}$$

where x_{ij} is the vector of attributes of alternative j and βs are the structural parameters that provide information about how the observable attributes relate to utility.

The total utility from alternative j, substituting from Eq 1:

$$\text{Eq 3.} \quad U_{nj} = \beta' x_{nj} + E_{nj}$$

With the assumption that representative utility is linear, the parameters β' are interpreted as marginal utilities. Since the utility is not directly observed, we use observed choices that are the results of utility to infer β'. The marginal utilities are interpreted as the true underlying preferences of the decision maker, and in this survey setup, these preferences are the respondent's perceptions of fairness.

The unobservable utility $E_n$is treated as random and the joint density of the random vector is represented as $f(E_{nj})$. Assuming the joint density of unobserved utility, f (E_n) is i.i.d. type I extreme value, we can express the choice probability as a logit model [10].

The probability of a decision maker *i* chooses the alternative j and the choice is represented as y_{ij}, the choice probability is given by the expression

$$\text{Eq 4.} \quad P_{ij} = \left(\frac{e^{Vij}}{\sum_j e^{Vij}}\right)^{y_{ij}} = \left(\frac{e^{\beta' x_{ij}}}{\sum_j e^{\beta' x_{ij}}}\right)^{y_{ij}}$$

If a choice set has two alternatives A and B, the choice probability that the respondent would choose alternative A conditional on being in class C is described in [9] as follows:

$$\text{Eq 5.} \quad P(A) = P(U_{Aj|c} > U_{Bj|c}) = \left(\frac{e^{V_{Aj|c}}}{e^{V_{Aj|c}} + e^{V_{Bj|c}}}\right)$$

From the above equations we see that the choice probability depends on the attributes that are presented in the alternatives and the choice of the respondent y_{ij}.

The probability that a respondent *i* belongs to class *C* and makes a choice y_i, is the choice probability of the response conditional on class *C*. This value depends on the attributes x_i and the structural parameters that define the class C, β'_c. i.e. $P(y_i|c) = P(y_i|x_i, \beta'_c)$

Latent Class Model

In the latent class model, the assumption is that the underlying preferences are latent/unmeasured. Assume that there are C classes or subgroups of respondents. Everyone with the same preference belongs to the same class and the preferences between the classes differ. The preference parameters that are estimated for each class are the marginal utility parameters β' described earlier. These parameters describe the preference patterns of a given class and are dependent only on the attributes of the alternative and the choice response to it [2]. The latent class model explained here is derived from the choice-only model in Breffle et al. [2], the notions used here also follow the same notations [2].

The class conditional probability is the probability that a decision maker *i* belongs to class C. It is represented as $\mathrm{P(c\colon t_i|y_i)}$ and indicates the class conditional probability depends on the type of the individual $\mathrm{t_i}$, and the respondent's choice $\mathrm{y_i}$**.** It depends on two probabilities- the choice probability and the unconditional class probability. The choice probability of a response conditional on a class is determined by the respondent's choice and the preference parameters that describe that class (β'_c).

The unconditional class probability of a response belonging to class C is a function of the decision maker's type (type meaning the subset of individuals that have the same attributes). It is called the unconditional class probability since it does not depend on the alternatives presented or the responder's choice. It only depends on the mean of the conditional class probabilities of all individuals of a type $\mathrm{t_i}$. The unconditional class probability allows us to model the clustering attributes or the decision maker's attributes into the latent class model. This probability averages over the class conditional probabilities of all the responses that have the same clustering attribute values. For example, in this study, our clustering attributes are context (0 or 1), severity (0 or 1) and respondent political preference (0, 1 or 2). There is an unconditional class probability for each combination of these attributes (example: context-0, severity-0 and respondent political preference- 0), and its value is the average of the class conditional probabilities of the responses that have the same clustering attributes- context-0, severity-0 and respondent political preference- 0. In this example, there are 12 unconditional class probabilities for each combination of context, severity and respondent political preference.

The class conditional probability is:

$$\text{Eq 6.} \quad \mathrm{P}(\mathrm{c}: \mathrm{t_i}|\mathrm{y_i}) = \frac{\mathrm{P}(\mathrm{y_i}|\mathrm{c})\mathrm{P}(\mathrm{c}:\mathrm{t_i})}{\sum_{c=1}^{C} \mathrm{P}(\mathrm{y_i}|\mathrm{c})\mathrm{P}(\mathrm{c}:\mathrm{t_i})}$$

where, $P(y_i|c)$ is the choice probability of the decision maker choosing y_i conditional on class C, $P(c: t_i)$ is the unconditional class probability.

The unconditional class probability is:

$$\text{Eq 7.} \quad P(c: t_i) = \frac{1}{N_{t_i}} \sum_{t_j \in t_i} P\left(c : t_j \middle| y_j\right)$$

where, N_{t_i} is the number of individuals of type *i* and $P\left(c : t_j \middle| y_j\right)$ is the conditional class probability.

The log likelihood function of the Latent Class discrete choice model is given by:

$$\text{Eq 8.} \quad L = \log \prod_i \left[\sum_{c=1}^{C} P(y_i|c) P(c: t_i) \right]$$

Expectation Maximization Algorithm

The log likelihood function is maximized using the Expectation Maximization algorithm. Initially, assign random class probabilities for each choice response, such that they sum up to 1 across the *n* classes. This indicates the probability that the response belongs to respective class. Next, find the average of the class probabilities for every type of decision maker using eq 7. Assign the calculated average as the unconditional class probability for the respective responses. With the unconditional class probability as the weight, estimate the parameters (β'_c) using a regression model for each class. This regression model estimates the preferences of the responses in each class. Next, calculate the choice probability for every response using the above estimated class parameters using eq 4. Calculate the log likelihood using eq 8. Update the class conditional probability using eq 6. This completes one iteration of the EM algorithm. Iterate these steps until the log likelihood reaches a maximum or remains constant.

Results

Model Selection

Since LCA is an unsupervised method, there is no way of defining exactly what the clusters should be. To determine if the resulting *n* clusters are the best fit for the data, Information Criteria (IC) is used as the method of model selection. IC is a measure of how well the model fits the data, it also takes into consideration the complexity of the model. The value of the IC is not relevant by itself, it does not convey any information about the fit of the model. The best fit model for the given data is determined by comparing the IC values across the models. The latent class model is applied to the data with a different number of classes (*n*) each time and the number of classes that has the lowest IC is chosen as the best fit model. In this study we estimate four ICs- AIC (Akaike's information criterion) [11], BIC (Bayesian Information Criterion) [11], CAIC (Bozdogan's Criterion) [11] and AICC (corrected Akaike criterion) [12]. The formulas for the ICs are presented in table 3.

In this study, the latent class model was applied for n ranging from n=2 to n=6. The maximum log likelihood and the IC values for each value of n are reported in table 4. Comparing the values of the maximum log likelihood and the ICs, the best fit model with the lowest values for every IC is the model with three classes. The model with two classes behaved oddly, the negative log likelihood increased with the iterations indicating that the model did not converge. This model was run with different initializations and all of the trials showed an increase in negative log likelihood.

Table 3: List of ICs used for model estimation and their formula. LL is the maximum log likelihood of the model, k is the number of parameters in the model, n is the number of datapoints in the dataset.

IC	***Formula***
AIC	$-2 * LL + 2 * k$
BIC	$-2 * LL + k * log(n)$
AICC	$-2 * LL + 2k * (n/(n - k - 1))$
CAIC	$-2 * LL + k * log(n + 1)$

Table 4: Maximum log likelihood and ICs for different number of clusters

No. of Clusters	**2**	**3**	**4**	**5**
Log Likelihood	-21742.8	-5930.93	-7726.72	-10059.4
AIC	43581.61	12005.85	15645.45	20358.83
BIC	43946.36	12552.97	16374.94	21270.69
AICC	43581.94	12006.57	15646.72	20360.81
CAIC	43994.36	12624.97	16470.94	21390.69

Class Distributions and Preference Patterns

Having chosen the best fit as three latent classes for the survey data, we examine the distribution of the grouping attributes amongst the clusters to identify distinct patterns present in each cluster. We can also compare the regression coefficients of each class to see how the different clusters' preferences vary.

The clustering variables considered in this study are the context, severity and responder's political preference. Examining the distribution of these attributes within the clusters shows that specific scenarios are grouped together indicating that they have similar preference patterns amongst respondents. This cluster wise distribution is shown in figure 1. The clustering results of this study showing that the preferences for the tuition and lifesaving context are similar and that of housing and pain-relief device are similar, match the preference patterns reported in the Hannan et all. paper [3]. These similarities are described later.

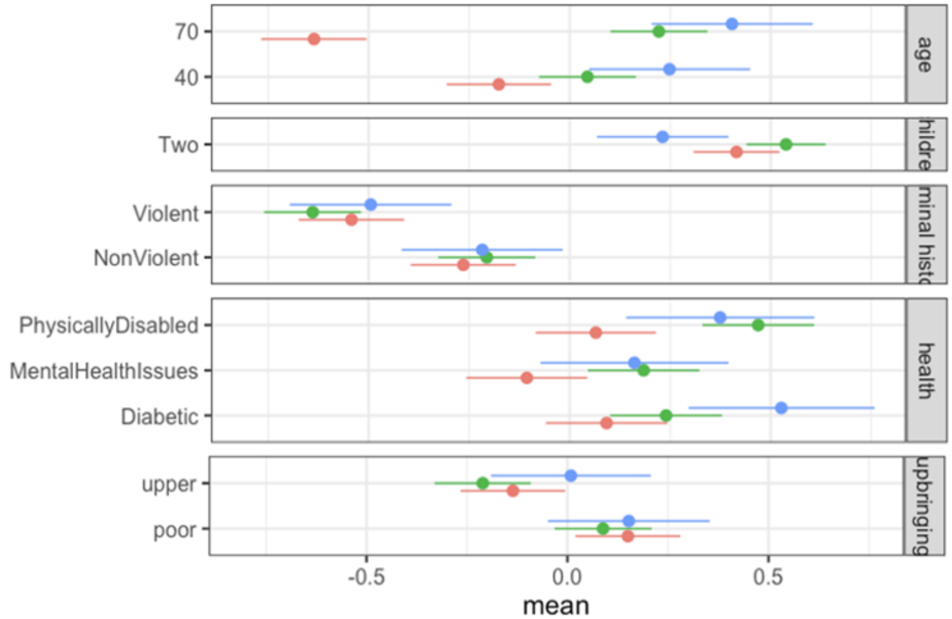
	Class 1	Class 2	Class 3
Membership in %	43.97%	42.36%	13.66%
Patterns	Social, Low (Tuition) Dem - 97.97% Rep, Oth/Ind - 100%	Social, High (Housing) All -100%	Covid, Low (Pain-relief) Dem - 9% Oth/Ind - 0%, Rep-100%
	Covid, High (Life-saving) Oth/Ind - 0.04% Dem, Rep - 100%	Covid, Low (Pain-relief) Dem - 91.07% Rep - 0%, Oth/Ind - 100%	Covid, High (Life-saving) Dem, Rep - 0% Oth/Ind - 95.9%
		Social, Low (Tuition) Dem - 2%	

Figure 1: Class-wise distributions of the clustering attributes

Class 1 is the largest class with 44% of responses. The only two scenarios present in this class are social- low and Covid-19- high, i.e. tuition assistance for courses at a local community college and ventilator or new treatment that increases survival odds by 50% in the case of high severity Covid-19. The distribution of the responses based on the responder's political preference, in Class 1, shows that almost all the responses are clustered in Class 1, indicating that they all share the same preference patterns. Since the number of responses of those who choose other/independent as their political preference is a very small fraction of the total responses, the distribution of these responses for the Covid-19-high scenario responses between class 1 and class 3 are ignored.

In Class 2, the two main scenarios present are social- high and Covid-19- low, i.e. housing availability for people without a place to live and medical device that alleviates mild discomfort in the case of low Covid-19 severity. A small fraction of responses in the social- low severity scenario by respondent's identifying as democrats (2%) is also present in this cluster. Examining the distributions according to political preferences, for the social high scenario, a 100% of the responses are clustered together in this class. For Covid-19-low however, 91.7% of the responses from democrats and a 100% of the responses from others /independent are present in this cluster, whereas all of the republican responses for this scenario are in class 3. This is the only scenario for which there is a difference in the preferences by partisanship.

Class 3 is the smallest of the three clusters with 13.66% of responses. It contains only Covid-19 scenarios- both low and high severity. If the responses by other/independent respondents is ignored due to it being a very small fraction of the responses, most of the responses in this class belong to Republican respondents of the Covid-19- low scenario. As mentioned earlier, this is the only scenario that shows a divide in preferences based on partisanship with only 9% of Democrats and 100% of Republicans of this scenario showing similar preferences for the allocation of pain relief devices.



Class 1 (43.97%)	Class 2 (42.36%)	Class 3 (13.66%)
Social, Low (Tuition)	Social, High (Housing)	Covid, Low (Pain relief)
Covid, High (Life-saving)	Covid, Low (Pain relief)	Covid, High (Life-saving)

Figure 2: Preference patterns of each cluster

Figure 2 shows a plot of the regression coefficients of the three classes, the red, green and blue lines representing class 1, 2 and 3 respectively. In this plot, the coefficient of a level of an attribute indicates how much it is preferred in comparison with the attribute level that is considered as the baseline. For example, for the age attribute, 20 years is the baseline and the positive coefficients (green and blue points) for the 40 and 70 age groups, indicate that in comparison to a person aged 20, higher preference is given to older people aged 40 or 70. Similarly the red points indicate that in comparison to the baseline of 20 years, less preference is given to the older people.

There are two main observations in this plot that match with the findings of the Hannan et all. [3] paper that show how the preferences vary with context. First, for the health attribute, the coefficient for mental health issues is negative in Class 1, whereas for physically disabled and diabetic levels, the coefficients are positive. This means that for tuition aid and life-saving devices, people give lesser preference to people with a history of mental health issues in comparison to a healthy person. Physically disabled or Diabetic people are given a higher preference for the same services. Second, for the age attribute, there is a clear preference for younger people to receive tuition aid and life-saving devices than compared to older people, whereas for housing and pain relief they are given the least preference. Older people (70 years) are given the most preference for housing and pain-relief services.

Other observations that also match the results of the Hannan et all. paper [3] are the preferences for the attributes of upbringing, children and criminal history. Irrespective of the context or scenario that is presented, all responders show high preference for persons with kids in comparison to persons with no children. Similarly, all responders gave importance to criminal history regardless of the service that was being allocated. Persons with a violent criminal history were given the least preferences for all the services. While persons with a nonviolent criminal history were also given lower preference when compared to a person with no history, they were given a higher preference than those with a violent history. In the case of the upbringing attribute, for all three classes, there is a clear preference given to a person of poorer upbringing when to compared to the middle-class baseline. For responses in Class 1 and Class 2, persons with an upper-class upbringing were given lesser preference to tuition, housing and life saving devices. For Covid-19- low scenarios however, Class 3 responses that comprises majority of republican responses for this scenario, show equal preference for persons with upper- or middle-class upbringing.

Conclusion

In this study, a latent class model was built to find latent subgroups amongst the survey responses [3], resulting in three distinct clusters of preferences as the best fit for the data. We examined these clusters to show that preference patterns were similar when the services to be allocated were tuition aid and life saving devices. The preferences for housing and pain relief services were clustered together in another class, showing similarity in preferences. We analyzed the regression parameters of each class to identify the attributes that showed differences in preferences with service context.

We observed that in all scenarios, context mattered more than partisanship. Differences in preferences with context was most evident for the age and health issues attributes- older people were least preferred for tuition assistance and life saving device but were given higher preference to pain relief and housing services. While people with mental health issues were less likely to be given life saving devices, healthy people were less likely to receive housing or pain relief services. The preferences for people with the children consistently higher compared to no children across all scenarios. Similarly, criminal history was also a very important attributed where people with no history were more likely to get the services.

The clustering variables used in this study included only one respondent demographic variable- political affiliation [3]. For future work, the other demographic attributes can be included to explore variations in preferences with age, education and race. The latent class model developed here can also be applied to any similar dataset to identify unobservable subgroups and examine the heterogeneity in preferences.

References

[1] Awad, E., Dsouza, S., Kim, R. et al. The Moral Machine Experiment. Nature, 2018. https://doi.org/10.1038/s41586-018-0637-6

[2] Breffle, W.S., Morey, E.R. & Thacher, J.A. A Joint Latent-Class Model: Combining Likert-Scale Preference Statements with Choice Data to Harvest Preference Heterogeneity. Environ Resource Econ 2011. https://doi.org/10.1007/s10640-011-9463-0

[3] Jacqueline Hannan, Huei-Yen Winnie Chen, and Kenneth Joseph. 2021. Who Gets What, According to Whom? An Analysis of Fairness Perceptions in Service Allocation. AAAI/ACM Conference on AI, Ethics, and Society, 2021. https://doi.org/10.1145/3461702.3462568

[4] Awad, E., Dsouza, S., Kim, R. et al. The Moral Machine Experiment – Supplementary Material. Nature, 2018.

[5] Kimberly J. Petersen et al. The Application of Latent Class Analysis for Investigating Population Child Mental Health: A Systematic Review. Front. Psychol., 2019. https://doi.org/10.3389/fpsyg.2019.01214

[6] Collins, L. M., & Lanza, S. T. Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences, 2010.

[7] Stelios Tsafarakis et al. Investigating the preferences of individuals on public transport innovations using the Maximum Difference Scaling method. Eur. Transp. Res., 2019. https://doi.org/10.1186/s12544-018-0340-6

[8] Liberato Camilleri et al. Market Segmentation through Conjoint Analysis using Latent Class Models. European Simulation and Modelling, 2011.

[9] T. Dillingham et al. A latent-class discrete-choice model to demonstrate how course attributes and student characteristics influence demand for economics electives: the challenge to increase enrollment. 2016.

[10] Kenneth Train, Discrete Choice Methods with Simulation (Book). 2009

[11] John J. Dziak et al. Sensitivity and specificity of information criteria. 2012.

[12] Paulo C.Emiliano et al. Information criteria: How do they behave in different models? 2013. https://doi.org/10.1016/j.csda.2013.07.032