Individual Preference Learning with Collaborative Learning Framework

Xi Zhu

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Cynthia Chen, Chair

Shuai Huang, Chair

Jeff Ban

Shan Liu

Program Authorized to Offer Degree:
Civil and Environmental Engineering

University of Washington

Abstract

Individual Preference Learning
with Collaborative Learning Framework

Xi Zhu

Co-Chairs of the Supervisory Committee:
Professor Cynthia Chen
Department of Civil and Environmental Engineering

Associated Professor Shuai Huang
Department of Industrial and Systems Engineering

Smart, personal devices that interact with individuals make it possible to trigger desired behavioral changes with personalized incentives. Personalized incentives are the incentives that suit an individual's preferences. In this dissertation, individual preferences refer to a set of parameters describing how the individual values each influential factor in a travel alternative. To trigger behavioral changes with personalized incentives, a model that can accurately and efficiently estimate an individual's preferences from his behavior data is required.

Two challenges exist in individual preference learning. For the first, the number of observations available from each individual for individual preference learning is limited. This issue causes difficulties in preference updating. For the second, the observability of the choices made is limited. This is because that it is not possible to directly observe the preference parameters – the only information that can be observed is an individual's choice-making behavior. The two challenges prevent the use of traditional preference-learning techniques such as advanced econometric models (e.g., discrete choice models) derived from Random Utility Maximization (RUM) [19].Other techniques such as machine learning also cannot be applied for similar reasons [116, 137]. New methods are needed for individual preference learning.

This dissertation contributes to the existing literature in travel behavior studies by proposing individual preference learning methods such that personalized incentives could be accurately estimated to trigger behavioral changes, and proposing a design of an online experiment to collect travel behavior data. Specifically, two research questions are of interest:

(1) What methodology could be used to learn an individual's preferences with only a few observations of choices made by him?

(2) How to collect individuals' choice data to test the method proposed in the dissertation in terms of triggering individual behavioral changes with personalized incentives? In the dissertation, the behavior data is collected via a carefully designed online experiment utilizing the AMT (Amazon Mechanical Turk) platform. Considering the validity and reliability of the data, the dissertation contributes to the travel behavioral study in:

(1) a full factorial design of a randomized experiment with two factors (commuting time and work flexibility, each with three levels) utilizing the online platform of AMT (Amazon Mechanical Turk) to collect individuals' travel choices on departure time in a sequence of hypothetical scenarios, and

(2) a design of data quality control strategies, which refers to the design of some methods to reduce and identify the low-quality data collected in the experiment.

These data quality control methods, such as understanding check, response consistency check, responding time record, and social desirability scale, can be applied to other online experiments and behavioral studies.

To learn an individual's preference from a few choices made by him, a model structure that integrates a time-varying model and the collaborative learning model is proposed in the dissertation. The time-varying model is used to replace the original constant preference parameter to a time-dependent function, allowing an individual's preferences to fluctuate in his choice-making process. The collaborative learning model can exploit the underlying canonical structure of individuals' preference variation in a heterogeneous population. Specifically, the collaborative learning model could identify several patterns of preference changes (known as "canonical models") that exist in the population. With the canonical models, each individual's preference change can be expressed by a linear combination of all those canonical models. Considering the model's computation time, an online updating strategy for the proposed model is also proposed, such that individual preferences could be learned accurately and efficiently. Detailed specifications of two different formulations of the time-varying model are presented in the dissertation, with some explorations on model properties with simulations. The models are also applied to the real-world dataset collected in the online experiment. Results show that the proposed models can achieve higher accuracy in parameter learning and behavioral prediction than traditional preference learning models such as the logit model and the mixed logit model.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Figure Number Page

ACKNOWLEDGMENTS

I would first like to thank my advisor, Dr. Cynthia Chen, whose expertise was invaluable in conceiving the research. You patiently guided me to the right path in research whenever I deviated from the track, and supported me in various aspects. It is a great honor to have Dr. Shuai Huang as my co-advisor. With your wealth of experience in methodology, you always provided me with helpful and practical suggestions. It is also my fortune to have a supportive dissertation committee. Dr. Jeff Ban pointed out a very important property of the methodology that I missed, which deepened my understanding of the methodology itself. Being well attuned to the methodology, Dr. Shan Liu provided invaluable insight into the study and shared her experiences and knowledge generously. Dr. Dan Abramson was very helpful in broadening my research horizons with his expertise, taking the time to talk with me about possible wide application prospects. I also very much appreciate Dr. Derek Reed for the professional suggestions in the design of the online experiment on the behavioral study.

I would like to acknowledge my colleague Feilong Wang for his contribution on proposing the "divide and conquer" strategy in the particle filter model presented in Chapter 3, and the assistance in the experimental design presented in Chapter 4. Staying in the same lab, you gave me a lot of inspiring ideas in our daily discussions. Thanks should also go to Jingshuo Feng, who helped a lot when I was exploring the methodology. I also thank all my friends in the lab, Xiangyang Guan, Yuanjie Tu, Katie Idziorek, Ni Dong, for listening to my rehearsal and providing valuable advice to the work.

Also, I would like to thank my parents for your unconditional love and supports. Special thanks to my counselor, Susan Busteed. Without you, I could not go through this process. Finally, there are my friends, who always made my life colorful.

Chapter 1
INTRODUCTION

1.1 Research motivation

1.1.1 Personalized alternatives in triggering behavioral changes

Let's see two stories first.

Loki drives to work every morning. He departs from his home at around $8:00$, and arrives at his office at $8:45-8:50$. In the afternoon, he leaves his office at $5:00$. Usually, it takes him about $30-35$ minutes to drive back home.

Kabie also drives to work every day. She leaves her home at $8:00$ in the morning, and leaves her office at $5:30$ in the afternoon. Typically, her commuting time is about $20-25$ minutes. Loki and Kabie are just two commuters in Seattle who drives alone to work and suffer from the bad traffic during peak hours every day. In fact, Seattle ranks the 2nd worst commuting time among 62 cities in the U.S., having 44% of its residents commuting to work by driving alone in a car, truck or van [151]. Since driving alone, or solo driving, is widely known as an unsustainable travel mode that produces congestion and pollution, it is expected that the traffic congestion could be relieved if we can persuade those solo drivers to change their travel behaviors, e.g., to switch from driving alone during peak hours to taking public transit or to travel at non-peak period.

Now, if you are asked to persuade drive-alone commuters to change their travel behaviors, what would you do? Or, let's consider a simpler question, how could you persuade Loki and Kabie to change their behaviors?

It is for sure that given the limited information above, we almost know nothing about the two commuters. This always happens when traditional generic incentives are provided in Transportation Demand Management strategies: we toll the freeway during peak hours or according to the traffic conditions; we set HOV and HOT lanes to encourage car-sharing; or we provide monthly pass of public transit to facilitate mode switch [151, 155, 43]. While, at the group level, we know that a certain percentage of the individuals would change their behaviors in response to these strategies, we have limited individual-specific knowledge about a specific driver, and could barely predict whether he would like to make the change if a certain strategy is provided. In our circumstance, we may not be able to tell whether Loki would be willing to take public transit if we provide a free pass to him, or whether Kabie would consider leaving home 50 minutes later than she usually does if we tell her that the traffic on the road could be much better at that time.

Let's try to get some more information about our two commuters. Loki's office is in downtown Seattle, which is 13 miles away from his home located in North Seattle. He has flexible working schedules, but every morning, he needs to drop his children off at a school near his home. It is his wife's responsibility to pick up the children in the afternoon, and the whole family always have dinner together. Kabie lives in a neighborhood which has a lake and several small parks. Her office is about 5 miles away from her home. She lives together with her dog. At weekends, she always walks her dog in the parks.

Getting a little bit more information about the commuters would help us provide more suitable alternatives, which are the alternatives promoted after taking individual-specific information, i.e., his constraints and preferences, into consideration. An individual's preferences here in travel choices refer to how the individual values different influential factors (such as travel time, travel cost) when he considers whether to select an alternative, which will further impact how attractive the alternative is for him. For example, we may know that Loki is likely to have low sensitive on departure time changes and the mode of transit, since he is responsible for sending his children to school and may not be willing to be absent from the dinner with his family. Therefore it is hard to persuade him to change behaviors by providing free transit pass or suggesting him to depart later. Route change may be possible for Loki if we provide some extra incentives as a compensation. Similarly, Kabie might be able to switch mode if incentives are provided, given her positive preferences on outdoor activities.

Assume that we have chosen another route for Loki, which also goes by his children's school (i.e., this new route won't impact him on sending children to school) but will be 2.5 miles longer than his original route. Let's further assume that we can use cumulative reward points as incentives, which could be redeemed to digital monetary credits and exchanged to services in many online stores. Now the question becomes: how many incentives should we provide to him? Similarly, how many points should we award to Kabie if we suggest her to bike to work?

1.1.2 Personalized alternatives, individual preferences, and personalized incentives

What we are doing to Loki and Kabie now is to shift individuals' travel behaviors via personalized travel alternatives and incentives, which is also the goal of an existing personalized system *Tripod* [9, 10, 165]. For system *Tripod*, personalization means identifying an alternative set unique to an individual according to his trip request and the utilities of alternatives are calculated with the preferences learned from his historical behaviors. The personalized system allocates incentives according to the contribution of the individual's behavioral change to the entire network that is predicted with micro-simulation of the network system [9, 10]. With *Tripod*, the personalized amounts of incentives provided to Loki and Kabie would be decided by the degree to which their behavioral changes can help relieve traffic congestion, reduce air pollution, or save energy.

It could be seen that for *Tripod*, behavioral consideration on whether the incentive is attractive enough for the individual to accept the promoted alternative(s) is not accounted for. This is exactly what this dissertation aims to achieve by determining the right amount of incentives such that the probability of the individual's accepting the promoted alternative is greater than a set threshold (e.g., 60%). In other words, by providing this type of personalized incentives, we are able to tell the probability of Loki's or Kabie's accepting and behavioral changing.

Ideally, both behavioral and system-level considerations shall be accounted for when providing personalized incentives. System-level considerations are not incorporated and they relate to, for example, deciding which promoted alternative(s) shall be presented to which individual, minimizing total reward points given subject to a minimum threshold improvement in system performance, or addressing equity concerns such that reward points are not doled out to favor those who may manipulate the system for more rewards. Achieving these system-level goals at the same time would require significant research that goes significantly beyond the current scope of the dissertation, which aims to address challenges at the individual level, or more specifically, learning individual preferences for personalized recommendations. Some preliminary results with system-level considerations are presented in paper [194].

Again, how many reward points are needed for the two commuters to change behaviors requires us to know their preferences. Let's give a more clear definition of the "individual preferences":

> The individual preference refers to a set of parameters β_r ($r \in 1, 2, \ldots, R$), each of which describes how the individual values an influential attribute x_r of a travel alternative (R gives the total number of influential attributes of an alternative, e.g., trip cost, trip time, reward points).

If we know how Loki values his travel time and the reward points, we may be able to provide him incentives that could make him feel the promoted alternative is attractive. For instance, we may be able to tell him: "Hi Loki! Would you like to have a one-month subscription of Disney+ for free? You only need to drive this route to your work for 10 times this month and gain 100 points!" Similarly, we may also be able to nudge Kabie: "Consider riding to your work for 5 days! You could gain 50 reward points, which equals to 5 coupon in PetSmart for your dog!"

It can be imagined that with these personalized incentives, the probability of Loki and Kabie's acceptance will be significantly higher than that with generic incentives. If we can provide incentives to all those drive-alone drivers, more people may switch to sustainable travel behaviors. The key task in providing personalized incentives is to learn each individual's preferences, namely, to know how each values different influential factors in his choice-making process. This task requires to consecutively interact with each individual, obtain his data, and learn his preferences with a model. In the past decade, the rapid proliferation of smart, personal devices has not only generated enormous data that allow us observe people's travel trajectories in time and space at an unprecedented scale [32], but also provided a medium that revolutionized the interactions between people and devices. Today, personalized interactions with individuals are a reality, and how to learn each individual's preferences from his data is an essential problem in the whole process.

1.2 Challenges in individual preference learning

Two challenges exist in individual preference learning. The first challenge is the limited number of observations available to the learning system, since a single individual can only generate a few observations [194]. This challenge prevents the use of traditional preference-learning techniques such as advanced econometric models (e.g., discrete choice models) derived from Random Utility Maximization (RUM) [19]. Preferences estimated from such models represent averages for a sample or a group, thus not satisfying the personalization goal. Even regressions with only an individual's data are difficult to apply due to the limited number of observations. Other techniques such as machine learning cannot be applied for similar reasons [116, 136, 137]. Also, machine learning techniques do not provide explanations for decision-making choices, even though their prediction accuracy can be high [128].

Additionally, this challenge is an obstacle to capturing preference changes. An individual's preference may not be stable or unchanged, but vary along with the choice scenarios or gradually evolve with personal experiences [21, 97, 81, 104], capturing preference changes may be the key point in preference estimation and behavior prediction. However, since an individual's preference can change even when he makes one choice, the issue of the limited observations also exists in the preference updating process. Regression models with only an individual's data are not applicable because the estimates of these models are the averages for a period rather than the preference at the latest time step.

The second challenge is the lack of observability: given two alternative choices one being the default choice (0) and the other being the promoted one (1), based on RUM, we can write the probability of choice (1) being selected as $p_1 = Pr(U_1 > U_0)$, where $U_i = V_i + \epsilon_i (i \in 0, 1)$ and is the utility associated with choice i, with V_i as the measurable systematic utility ($V_i = \sum_k \beta_k x_k$), and ϵ_i as the random utility [19]. Unfortunately, none of the key terms in this setup (p_1, U_i, or V_i) can be directly observed. The only piece of information that is directly observed is whether the individual accepts or rejects the promoted choice. This observable choice contains limited information, since if choice (1) is selected, it means $p_1 \in (0.5, 1)$, which can be the result of many possibilities for β_k.

This dissertation aims to propose methods that could learn an individual's preferences with only a few observations of his choices.

1.3 Research question

This dissertation aims to answer the question: How to learn and update an individual's preference with only a few observations of his choices available?

Two sub-questions are closely related to the research question:

1. What methodology could be used to estimate individual preference with only a few observations of choice made by him?

2. How to collect people's choice data to test the method proposed in the dissertation in terms of triggering individual preferences with personalized incentives?

1.4 Overview of the methodologies in the dissertation

1.4.1 Personalized control system and the particle filter model in individual preference learning

A personalized control system that can interact with each individual is designed to learn individual preference, such that it is possible to trigger the desired behavioral change by providing personalized incentives, rather than generic incentives implemented in traditional TDM strategies. The system is able to consecutively provide sustainable alternatives with personalized incentives to the individual, obtain his responses, and learn/update his preferences from his choices. To achieve this, three modules in the system work sequentially: (1) At each time step, when an individual submits a trip request, a promoted sustainable travel alternative with personalized incentives are presented to the individual in module PREDICTION, together with the default travel option. The personalized incentive is calculated based on RUM, given the attributes of the alternative and his preference estimated at the previous time step. (2) The individual's response, e.g., the choice he makes, is captured in module MEASUREMENT. (3) With the individual's choice and the alternatives presented to him, his preference is learned and updated in module UPDATE by a preference learning algorithm embedded in the system. Again, the updated preference parameters are sent to the PREDICTION module to estimate the personalized incentives at the next time step.

To tackle the challenge of learning with a limited number of observations, the particle filter is adopted in the preference learning algorithm in module UPDATE in [194], which is a bayesian approach in parameter estimation and can update the estimates with only one new data point. To tackle the challenge of lack of observability, an additional piece of information is solicited in module MEASUREMENT: besides the choice, the individual is also asked to estimate the level of attractiveness of the proposed alternative with a scale. Moreover, the "divide-and-conquer" strategy is also adopted in module UPDATE with carefully designed alternatives presented to the individual, such that it is possible to estimate only one parameter at each time the learning algorithm runs.

Though the preference learning algorithm based on the particle filter can tackle the challenge in individual preference learning, there is still plenty of room for improving the learning model in module UPDATE. For the first, soliciting additional information on the utility ratio increases the burden of the respondents. Ideally, we hope that the learning algorithm can learn an individual's preferences from his choices, which means that the individual only needs to make selections in the choice scenarios. For the second, the individual preference learning algorithm based on particle filter approach only utilizes "individual-level" information, while "group-level" information in the whole population is not considered. Preference learning with individual-level information means that the learning is based on a person's data, including the choices he has made and the contexts of those choices. It focuses on individual uniqueness, and the underlying hypothesis is that each individual has his peculiar tastes toward various factors that matter in a choice scenario. Preference learning with "group-level" information means learning combines many individuals' data, including their choices, the contexts of those choices, and individuals' characteristics and behavioral history. It focuses on the similarity between individuals, and the underlying hypothesis is that there is at least a certain degree of commonality in preferences shared by multiple individuals. In this dissertation, it is believed that both hypotheses are likely true to a certain extent, and possible models that can utilize information in both levels are proposed. It might help improve individual preference learning accuracy if knowledge from the massive individuals' data is also extracted in the preference learning process. For the third, the "divide-and-conquer" strategy requires carefully-designed promoted alternatives (at least in the first few choice-making scenarios) to make sure all the preference parameters can be learned and updated by the algorithm. It could be imagined that in the real-world situation, the promoted alternatives are likely to be decided by the real-time traffic conditions, rather than being deliberately designed. For the fourth, the preferences learned by particle filter model, or even by simple logit regressions, are highly unstable. The learned preferences can be seen as stable if they are in a reasonable range without many unexpected extreme values in the learning and updating process. Since the data that can be used in the learning process is limited, the influence of the noise becomes significant in models such as particle filter and logit regression, which may result in extreme values that are significantly different from other values even in order of the magnitudes. The particle filter model we used in the online experiment deals with this problem by manually setting reasonable ranges based on domain knowledge and a hypothetical budget, yet this may cause loss of the information and may not be applicable if the domain knowledge is not known.

Considering these four problems, other possible methodologies that could be used to learn individual preferences in module UPDATE are needed.

1.4.2 Online experiment

To test the system's performance in terms of triggering behavioral changes by learning individual preferences and providing personalized incentives, an online experiment is designed and conducted to collect individuals' stated choices in the real world. The performance of these models is evaluated by the percentages of the models' correct predictions on individuals' choice-making behaviors towards the promoted alternatives.

Considering the validity and reliability of the data collected in the online experiment, the design of the experiment includes two parts: (1) the design of the online experiment, and (2) the design of data quality control, which refers to a design of methods to reduce and identify the low-quality data collected in the experiment. The first part is related to data

Table 1.1: Table of notations.

Notation	Meanings
i	Individual i
j, J	Alternative j, alternative l, where $j \in (1, 2, \ldots, J)$
r, R	Can be in circumstances when referring to the rth attribute of an alternative, or rth dimension of the preference vector, where $r \in (1, 2, \ldots, R)$
k, K	The kth canonical model in a collaborative learning model, where $k \in (1, 2, \ldots, K)$
U_{ijt}	Utility of alternative j for individual i at time step t
$\boldsymbol{x}_{it}$	Vector of the attributes in choice scenario t (the binary choice at time step t) for individual i
$\boldsymbol{\beta}_i$	Vector of parameters/preferences of individual i corresponding to all attributes
ϵ	Random term drawn from a certain distribution
V_{ijt}	Systematic utility of alternative j for individual i at time step t
P_{ijt}	Probability for individual i choosing alternative j at time step t
y_{it}	The choice made to the proposed alternative in the binary choice at time step t by individual i. The individual accepts the alternative if $y_{it} = 1$
ΔV_{01}	Difference between V_0 and V_1

validity, which is the design of the scenarios and experiment groups. In the experiment, each individual is asked to make binary choices to a hypothetical scenario sequence, in which he needs to decide when to depart for work. Two factors are considered influential when deciding random groups: typical commuting time, work (arrival time) flexibility level. Each factor is categorized into three levels, forming nine groups in the full factorial design of the randomized experiment. The second part is related to the reliability of the data collected. Since the experiment is conducted online, and the respondents are recruited from AMT (Amazon Mechanical Turk) platform, several reasons may result in low-quality responses in the experiment:

1. Respondents do not understand in the hypothetical background setting.

2. Respondents are motivated by receiving payment for participating in the experiment.

3. Respondents tend to respond in a socially desirable way.

Considering these possible reasons, several checks are designed and applied in the experiment. These include questions checking the respondents' understanding of the background settings, the consistency in an individual's responses, the responding time, and a social desirability scale. The design of the behavioral experiment, especially the design of the checks, significantly helps reduce low-quality data in the experiment. Simultaneously, it contributes to behavior data collection by providing possible methods to control data quality in the behavioral experiment.

The the real-world dataset collected in online experiment is used to test the performance of the individual preference learning models proposed later in this dissertation. The performance of the model refers to the percentage of the individual accepting the promoted alternative (for the particle filter model which is already used in the system), or the percentage of the correct prediction on the individual's behavior (accepts or rejects) given the promoted alternatives (for other models).

1.4.3 Individual preference learning utilizing collaborative learning model

Regarding the limitations listed earlier in Section 1.4.1, this dissertation proposes to integrate a time-varying model and a collaborative learning model. The time-varying model integrated replaces the original constant preference parameter β to a time-dependent function $\beta(t)$, allowing an individual's preferences to fluctuate in his discrete choice-making process. Specifically, while in the traditional logistic model, the probability of the proposed alternative being chosen in the tth binary choice (the binary choice at tth time step) by individual i ($y_{it} = 1$) is $\frac{\exp(\boldsymbol{\beta}_i^T \boldsymbol{x}_{it})}{1+\exp(\boldsymbol{\beta}_i^T \boldsymbol{x}_{it})}$ ($\boldsymbol{x}_{it}$ is the differences of the attributes between two alternatives at time step t), in this dissertation, $\boldsymbol{\beta}_i$ is converted to a time-dependent model $\boldsymbol{\beta}_i(t)$, which may change with t. The basic idea of the collaborative learning structure is to exploit the underlying canonical structure in individuals' preference variation of a given population with heterogeneity [112, 111, 113]. The canonical structure can be seen as a system used to express and represent all the individuals' preference variations in the dataset. In a canonical structure, several preference variation patterns (called "canonical models") can be identified from the whole dataset, and each individual's preference variation is then represented by a combination of those common patterns. In the current study, collaborative learning model splits the learned preferences into two parts: (1) canonical models in the format of the time-dependent model $\beta(t)$, which represent the common time-varying patterns/types of preference identified from all the individuals in the population, and (2) a membership vector[1], representing the degrees of the resemblance of the individual's preferences to those canonical models. With the structure, each individual's preference is presented with a linear combination of all the canonical models, which means that it integrates the group-level similarities (represented by canonical models) and individual-level personality (represented by his membership vector).

The integrated model deals with the first challenge in individual preference learning by introducing group-level information into individual learning. From the perspective of a

[1]There could be several membership vectors for each individual if the preferences have multiple dimensions, i.e. if there are multiple attributes that matter in the decision-making process, as shown in this dissertation in Chapter 5. In general, the number of the membership vectors, or the formulation of the membership vector, is decided by the formulation of the canonical models.

single individual, more information is brought in his preference learning process. The model deals with the second challenge by utilizing the log-likelihood function of the binary logit model. The integrated model described in the previous paragraph is a general structure of the model. In this dissertation, the integration of $\beta(t)$ and the collaborative learning model is first preliminarily illustrated with $\beta(t)$ being a polynomial model. The purpose is to present how a time-varying preference model could be integrated with the collaborative learning model. Choosing $\beta(t)$ as a polynomial model in the illustration is because, for the first, the polynomial models allow to express time-dependent variations, while the only influential variable is t. This simple model is effective and can capture patterns in a range of applications. For the second, to further show that the polynomial model could capture preference changes in an individual choice-making process, some regressions have been run with individual preferences learned by the particle filter model embedded in the online experiment. The regression results show that polynomial models can fit an individual's preference data statistically significantly.

Though the polynomial model could capture the changes in preferences over time, it is not the best model to describe the preference changes. To select a better model of $\beta(t)$ that could describe the preference changes in the choice-making process, this dissertation further proposes to use a time-invariant model as $\beta(t)$. Literature suggests that an individual's preference may vary according to the choice scenarios or evolve gradually with personal experiences [21, 97, 81, 104]. Thus, in a proposed model of changing preferences following the formulation of a time-invariant system, $\beta(t)$ is impacted by both $\beta(t-1)$ and the attributes in the scenarios at the time step t.

The integrated model is able to (1) learn each individual's preference utilizing both individual-level and group-level information, (2) learn and update an individual's preferences given only one or a few observations available from him, (3) capture or allow an individual's preferences to vary over time in the individualized modeling, (4) have no requirements on scenario design in the learning process, and (5) only need individuals to make selections, i.e., no other responses required. This model structure meets all the needs. However, this method is computationally costly. The integrated model can be solved with a parameter estimation algorithm similar to the parameter estimation algorithm presented in [111, 112, 113], with which all the parameters in the canonical models and membership vectors could be estimated. When a few new observations are available, and an individual's preference needs to be updated, the new estimates can only be obtained by re-running the whole model and re-estimating all the parameters in canonical models and membership vectors. In other words, considering the computational cost, the algorithm can be inefficient when the population is large, but only several parameters might change and need to be updated. Thus, a two-stage (online and offline stages) parameter updating strategy is proposed in the dissertation, in which the online-updating algorithm can be used to just update the individual-specific parameters without the need to re-estimate the whole model. The assumption of the two-stage updating method is that the common patterns/types of individual preferences (i.e., the parameters in canonical models) shared by all individuals in the population would not change significantly in a short time, thus do not need to be updated frequently. We call the updating process "online updating" when only part of the parameters, i.e., an individual's membership vector, are revised while the canonical models are fixed. Accordingly, "offline updating" refers to the process when all parameters are updated with the iterative updating method used by the original collaborative learning model. In practice, the online updating process can take place several times each day for an individual, while the offline-updating may only happen periodically, e.g., once a week or so. In this dissertation, the two-stage updating strategy is applied to the parameter estimation process of both integrated models, which are further tested with simulations and the real-world dataset collected in the experiment.

1.5 Contributions

The main contribution of this dissertation includes:

1. The dissertation proposes a model structure integrating a time-varying preference model and the collaborative learning model such that it is possible to learn and update an individual's preference given only a few choices available from the individual.

2. To deal with the computation pressure in individual preference updating, the dissertation proposes an online-updating strategy for the learning model, such that an

individual's preferences can be estimated accurately and efficiently.

3. The dissertation proposes a design of an online experiment, showing a way to conduct and collect data for behavioral study. More importantly, the dissertation proposes a series of checks that can be used in the online experiment to control data quality, which is essential in online experiment and behavioral study.

1.6 Dissertation outline

The dissertation presents the work mainly in six chapters:

Chapter 2 delivers a literature review on relevant studies. The relevant studies include three main parts: (1) A brief introduction of RUM theory and the binary discrete choice model, which is based on the choice-making model we use in the dissertation. (2) A brief introduction of the Choice-Based Conjoint analysis, which is a commonly used preference learning method for choice-making behavior. A discussion on why the Choice-Based Conjoint analysis is not appropriate for the dissertation's problem is also presented. (3) The collaborative learning model adopted in the dissertation and the Latent Class Choice Model is compared, where the latter is also used in choice behavior analysis and has a similar structure with the collaborative learning model.

In Chapter 3, we describe a personalized control system, with which promoted alternatives could be presented to the individuals, the choice data could be obtained, and individual preferences could be learned and updated. We present the three modules of the system, along with the flowchart and the necessary details.

With the personalized control system, an experiment is designed and conducted to test the system's performance and collect a real-world dataset. The design of the experiment is elaborated in Chapter 4. The experiment mimics the personalized control system's interaction process, sequentially presenting choice scenarios to each individual and collecting his responses. After basic data cleaning and data processing steps, the dataset is used to test the performance of the models proposed in the following sections.

With some illustration of the collaborative learning model, Chapter 5 presents the integration of a time-varying preference model with the collaborative learning model to capture preference changes in the sequential choice-making process. In this chapter, the polynomial model with only one variable is used as the time-varying preference model to capture the changes of an individual's preferences. The polynomial model describes the trajectories of the preferences over time in the sequential choices, thus allows the preferences to change in the process. The model specification, parameter estimation algorithm, and the simulation results are presented in the chapter.

The model specification of integrating the collaborative learning structure and a time-invariant model is presented in Chapter 6. The time-invariant model describes how the attributes impact the preferences in each scenario and how the preferences evolve. The parameter estimation algorithms and the simulation results of both models are presented whereafter.

The last chapter provides a summary and a discussion of the works presented in the dissertation, including limitations, possible reasons, and other related issues. The discussion chapter also proposes future research directions.

Chapter 2

LITERATURE

2.1 Discrete Choice Model and Random Utility Maximization Theory

One of the most widely accepted models, the discrete choice modeling, is used to describe and predict the probability of selecting a choice from a finite choice set in a probabilistic fashion, based on the Random Utility Maximization (RUM) theory [174]. It is the basic assumption of the choice-making behaviors in the dissertation, and the basic choice-making model in the various preference learning models presented in later chapters. In this subsection, a brief introduction of RUM and the binary discrete choice model is presented.

2.1.1 Random Utility Maximization (RUM)

Assume that an individual is facing J alternatives in a choice-making scenario. For the individual, each alternative j has a certain level of utility, denoted as U_j. The utility is a concept quantifying the attractiveness of an alternative in a choice scenario. The utilities of all the alternatives are known to the individual but not by a researcher. The individual chooses the alternative that provides the greatest utility, i.e., the individual would choose alternative j if and only if $U_j > U_l$ $(\forall j \neq i)$ [174].

From the researcher's perspective, the individual's choice behavior among a finite set of alternatives may be influenced by a number of factors, including some socioeconomic and socio-demographic characteristics of the individual, the influential factors in each alternative, and the external environments. We denote the attributes of all these factors as $x_r,\ r \in (1, 2, ..., R)$. With a set of parameters β_r quantifying the effects of these factors, the researcher can specify a function to express the utility the individual obtains if he chooses alternative j: $V_j = \sum_r \beta_r x_{jr}$ (the systematic utility).

Notice that here the utility captured by the researcher V_j and the utility known to the individual himself U_j are denoted differently. Because that there are aspects of utility that the researcher does not or cannot observe, the two utilities are different and have a relationship of $U_j = V_j + \epsilon_j$, where ϵ_j captures the factors that affect the utility that is known to the individual but is not included in V_j [174].

Given $U_j > U_l$ ($\forall j \neq i$), the probability statements in the individual's choice-making process can be written as:

$$\begin{aligned}
P_j &\equiv P(\text{The individual choosing alternative } j \text{ over alternative } l) \\
&= P(U_j \geq U_l) \;\; \forall j \neq l \in (1, 2, ..., J) \\
&= P(V_j + \epsilon_j \geq V_l + \epsilon_l) \;\; \forall j \neq l \in (1, 2, ..., J) \\
&= P(V_j - V_l \geq \epsilon_l - \epsilon_j) \;\; \forall j \neq l \in (1, 2, ..., J)
\end{aligned} \tag{2.1}$$

As the researcher has no knowledge on $\epsilon_j, \forall j \in (1, 2, ..., J)$, he may assume that ϵ_j is random. Typically, ϵ_j is assumed to follow a Gumbel or normal distribution, respectively [174]. The logit model is derived under the assumption that ϵ_j follows the Gumbel distribution. This is out of consideration for the calculation to obtain a closed-form expression for the integral of the joint distribution of the random ϵ_j [174]. Probit model assumes that the distribution is normal [174].

2.1.2 Binomial logit model

The binary logit model is used when there are only two alternatives for an individual to select. Since all the choice scenarios are binary choices in the dissertation, a brief introduction of the binary logit model is presented in the following. The reason of using binary choice scenarios in the dissertation is that binary choice problem is believed to be the basic problem in choice-making behavior study, as multi-choice problems can all be seen as a combination of a sequence of several binary choices.

Suppose that the utilities of the two alternatives (0 and 1) in a binary choice scenario are U_0 and U_1, respectively, for an individual. Following Equation 2.1, the probability statement in binary choice scenario is [174]:

$$
\begin{aligned}
P_0 &= P(U_0 > U_1) \\
&= P(V_0 + \epsilon_0 > V_1 + \epsilon_1) \\
&= P(V_0 - V_1 > \epsilon_1 - \epsilon_0).
\end{aligned}
\tag{2.2}
$$

With logistic model specification, the probability for an individual to choose alternative 0 is [174]:

$$
\begin{aligned}
P_0 &= \frac{\exp(V_0)}{\exp(V_0) + \exp(V_1)} \\
&= \frac{1}{1 + \exp(V_1 - V_0)} \\
&= \frac{1}{1 + \exp(\boldsymbol{\beta}(\boldsymbol{x}_1 - \boldsymbol{x}_0))} \\
&= \frac{1}{1 + \exp(\boldsymbol{\beta}\Delta\boldsymbol{x})}
\end{aligned}
\tag{2.3}
$$

The utilities $V_0 = \boldsymbol{\beta}\boldsymbol{x}_0$ and $V_1 = \boldsymbol{\beta}\boldsymbol{x}_1$ are the constant parts of the utilities of the two alternatives. From Equation 2.2, it can be seen that only the differences of the utilities matter in the choice-making process [174]. In other words, in the binary logit model, the attribute of each influential factor is the difference between the attributes of the two alternatives, $\boldsymbol{x}_1 - \boldsymbol{x}_0 = \Delta\boldsymbol{x}$. Since all the choice scenarios are binary choice questions in the dissertation, let the two alternatives in each scenario be alternative 0 (default alternative) and alternative 1 (promoted alternative). For simplicity, the Δ notation in the probability function 2.3 will be omit hereafter. In other words, in the dissertation, all the attributes $\boldsymbol{x}$ in the models refer to the difference of the corresponding attributes in alternative 1 and alternative 0.

Given that the binary logit model is used in the whole dissertation, it is also necessary to explain how the estimates are obtained for the binary logit model. The binomial logit model's estimation uses the Maximum Likelihood Estimation (MLE), with which the estimated parameters for $\boldsymbol{\beta}$ have the greatest likelihood/probability to obtain the observed choices [174].

Let's use an example to illustrate the estimation method. Assuming that we have M observed binary choice behaviors $(y_1, \boldsymbol{x}_1)$, $(y_2, \boldsymbol{x}_2)$, ..., $(y_M, \boldsymbol{x}_M)$ for an individual, where

$y_m \in (0,1)$ refers to the alternative he selects, and $\boldsymbol{x}_m$ refers to the vector of the differences between the attributes of the two alternatives. We would like to predict his choice-making behavior with a binary logit model. For each observation m, assume that $P(y_m = 1) = \frac{\exp(\boldsymbol{x}_m\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}_m\boldsymbol{\beta})}$, and $P(y_m = 0) = 1 - \frac{\exp(\boldsymbol{x}_m\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}_m\boldsymbol{\beta})}$. The likelihood function could be formulated as [174]:

$$L(y|\boldsymbol{x};\boldsymbol{\beta}) = \prod_{m\in\mathcal{M}_+} \frac{\exp(\boldsymbol{x}_m\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}_m\boldsymbol{\beta})} \prod_{m\in\mathcal{M}_-} [1 - \frac{\exp(\boldsymbol{x}_m\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}_m\boldsymbol{\beta})}] \tag{2.4}$$

where $\mathcal{M}^+$ refers to the observations for which $y_m = 1$, and $\mathcal{M}^-$ to the observations for which $y_m = 0$.

The log likelihood function for the example is [174]:

$$\begin{aligned} \ln L(y|\boldsymbol{x};\boldsymbol{\beta}) &= \sum_{m=1}^{M} \left\{ y_m \ln \frac{\exp(\boldsymbol{x}_m\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}_m\boldsymbol{\beta})} + (1-y_m)[1 - \ln \frac{\exp(\boldsymbol{x}_m\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}_m\boldsymbol{\beta})}] \right\} \\ &= \sum_{m=1}^{M} \left\{ y_m[\boldsymbol{x}_m\boldsymbol{\beta} - \ln(1+\exp(\boldsymbol{x}_m\boldsymbol{\beta}))] - (1-y_m)\ln(1+\exp(\boldsymbol{x}_m\boldsymbol{\beta})) \right\} \end{aligned} \tag{2.5}$$

The MLE of $\boldsymbol{\beta}$ maximises the log-likelihood function of Equation 2.5. In the dissertation, the objective functions of the optimization problems of the two proposed integrated models are formed according to the log-likelihood function of the binary logit model. In the following chapters, how the log-likelihood function is deduced will no longer be presented.

2.2 *Comparison of the Latent Class Choice Model (LCCM) and the Collaborative Learning Model*

As briefly introduced in the first chapter of this dissertation, collaborative learning model has a canonical structure, with which a set of canonical models are identified from all the individuals' data representing the common preference patterns, and a membership vector is identified for each individual representing the degrees of the resemblance of the individual's preferences to those canonical models. This structure may look similar to some people familiar with the latent class choice model (LCCM), which is also used for choice behavior analysis and assumes that there is unobserved preference heterogeneity among the population. This subsection will present a brief comparison between the collaborative learning model and the latent class choice model.

The latent class choice model posits that an individual's discrete choice behavior depends not only on the attributes in a choice scenario, but also on the latent preference heterogeneity that varies with factors that can not be directly observed by the analyst [73]. Thus, LCCM assumes that there are group-level preference classes, and each individual belongs to one of them. When identifying the classes with the choice data, LCCM assigns a membership vector to each individual, representing the probabilities of the individual belonging to those preference classes. Though LCCM looks similar to the collaborative learning structure and even share the same name for the term "membership vector", the collaborative model is different from the latent class model of discrete choice in three aspects:

For the first, the assumptions of the two models are different. The latent class model of discrete choice assumes that there exist several classes or groups of individuals, and the individuals in the same class have the same preferences [26]. An individual's membership vector in the latent class model represents the probabilities for the individual to be in each class [69, 29]. In other words, the latent class model assumes that each individual can be seen as belonging to one class identified by the model, and his preference is the same as other members in the class. For the collaborative model proposed in this dissertation, we assume that several unique underlying preference patterns can be identified from the data. However, an individual's membership vector represents the degree of resemblance of the individual's preferences to each preference pattern. His preference is a linear combination of all the preference patterns identified by the model, where the weights are his membership vector.

For the second, the questions to be answered by the two models are different. LCCM tries to answer the question of how many classes/groups there are among all individuals. The functional relationship between the identified classes and some covariates are explored such that it is possible to predict which class/group an individual may belong to, or what response the individual would give, with probabilities. The proposed model aims to learn each individual's preferences on all the attributes in the choice-making process and to be able to update the preferences efficiently.

For the third, the estimates (i.e., the preferences) obtained are different. Though the individual's class may be predicted, his preferences conditional on a certain class are not personal preferences but estimated at an aggregated level, assuming that the individual has the same preferences as other people in the same class. The proposed model, however, aims to learn each individual's personal preferences with the canonical models and the individual's membership vector identified from the data.

Chapter 3

DESIGN OF THE PERSONALIZED CONTROL SYSTEM

In this dissertation, we design a personalized control system to learn an individual's preferences, such that we could trigger the desired behavioral change by providing personalized incentives, rather than generic incentives implemented in traditional TDM strategies. As briefly introduced in Chapter 1, personalized incentives in this dissertation refer to those incentives that are designed to suit an individual's preferences and constraints, and the individual's preferences are captured by a set of parameters $\boldsymbol{\beta}$ describing how the individual values different influential attributes of alternative x (e.g., trip cost, trip time). Given a sustainable alternative to be promoted and its attributes, the system uses the learned individual preferences to present his incentives based on the theory of Random Utility Maximization (RUM) [80, 121]. According to RUM, the probability of choosing among multiple alternatives depends only on the difference in their respective utilities, and an individual will select the alternative that provides the maximum utility. Here, the utility is a concept quantifying the attractiveness of an alternative in a choice scenario, and is assumed to be indirectly related to the various characteristics of the alternative, the individual and the surrounding environment (so called "indirect utility") [19, 79]. Therefore, the incentive presented is expected to add utility to the promoted alternative so that for the user, the probability of accepting it is no lower than that of the non-promoted or default alternative (e.g., single driving). While individual preferences are not being directly observed, the system is designed to learn preferences from interactions with the individual: (1) when an individual is to conduct a trip, the system presents an incentive for the promoted alternative relying on its previously estimated preferences along with other trip information associated with the alternative; (2) the individual responds by accepting either the promoted alternative or the default alternative and his decision is captured by the system so that the preference estimates are updated. In the dissertation, a binary choice scenario is presented to the individual as the binary comparison involved in the scenario is the basis for the choice-making behavior, and it could also be extended to a multi-choice scenario by regarding the latter as a combination or a sequence of several binary choices.

Two challenges exist in achieving personalized learning, as stated in the chapter of introduction in this dissertation. The first one is the limited number of observations that are available to the system. Since each individual could only generate a few observations, and the preference may vary during the choice-making process, traditional preference-learning techniques such as advanced econometric models derived from RUM [19] and machine learning methods can not be used here [116, 136, 137].

The second issue is the lack of observability: given two alternative choices one being the default choice (0) and the other being the promoted one (1), based on RUM, none of the key terms including the probability of choice (1) being selected as p_1, the utility associated with choice i $U_i = V_i + \epsilon_i (i \in 0, 1)$, the measurable systematic utility ($V_i = \sum_r \beta_r x_r$), can be directly observed. The only piece of information that is directly observed is whether the individual accepts or rejects the promoted choice.

To address the issue of limited sample size (which is also the core idea of personalization, i.e., learning based on limited observations from a single individual), we propose a particle filter approach that views β_r as the underlying states to be learned. The use of the particle filter approach however does not address the second issue, lack of observability, which has two dimensions. The first one, as noted above, is reflected in that the acceptance or rejection of an alternative by the individual conveys little information on the relative attractiveness of the two alternatives. To deal with this, we design an interface to solicit additional information. We assume that though p_1, U_i, or V_i cannot be directly observed, we can observe a utility ratio (R_u) that reflects the relative attractiveness of two alternatives. This ratio function is consistent with RUM as choice (1) is selected if and only if $R_u > 0.5$. It also conveys more information related to the individual's preferences as it continuously changes with U_1.

The use of the ratio R_u, however, does not address the second dimension of the lack of observability issue, which is related to the inherent trade-off nature of human decision making. The utility function is the weighted sum of various factors that come into the decision-making process (e.g., travel time and travel cost), or $V_i = \sum_r \beta_r x_r$. In other words, an infinite set of possible values for r s gives rise to a single R_u. It is thus important to devise mechanisms to identify a unique set of β_rs that most likely give the observed R_u. Two solutions are proposed to address it. The first one is the "divide and conquer" strategy, which decomposes a multi-dimensional problem into multiple conditional one-dimensional problems by considering responses from a pair of choice scenarios sharing similar attribute values. The second one is to add domain knowledge on travel behavior as constraints while learning the preferences.

In the following, a detailed introduction of the proposed personalized system is presented. After that, we simply introduce the limitations of the proposed model in preference learning, leading to the online experiment and other preference learning models elaborated in the following chapters.

3.1 System overview

Figure 3.1 provides an overview on how the system works. As noted earlier, the personalized control (i.e. promoting behavior changes by providing personalized incentives) requires individuals' preferences learned via interactions with a system over time. Each interaction can be triggered by a trip request from an individual. For clarity, an interaction triggered by the tth trip request is noted as the interaction at time step t. A request consists of trip information such as origin O, destination D, and departure time T. In each interaction, following the framework of the particle filtering approach, three modules work sequentially to update preference estimates, including PREDICTION, MEASUREMENT and UPDATE. Based on trip information contained in the trip request and the previous estimates on the individual's preferences, the PREDICTION module formulates an alternative that is to be promoted, and predicts the amount of personalized incentive needed for the individual to accept the promoted alternative. For the personalized incentive, a reward points system, which is widely used due to its quantitative nature [148, 88], is adopted and tested in the system. The reward points earned by accepting the promoted alternative can be exchanged for real rewards (e.g., a discount on transit fares). The MEASUREMENT module presents the individual the binary choice scenario (one being the default alternative and the other

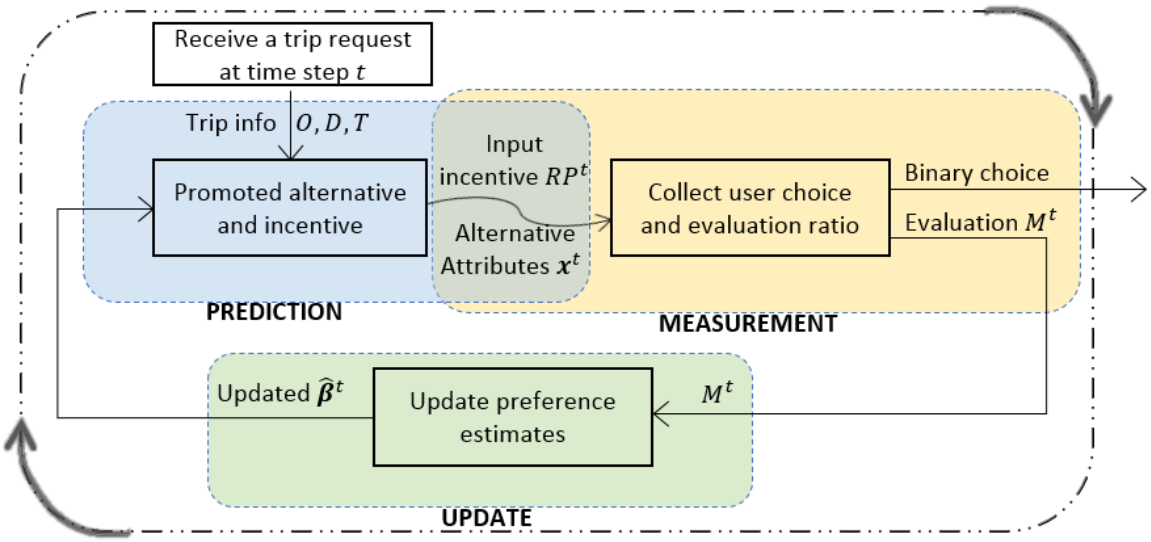


Figure 3.1: Flowchart of the control system.

being the promoted alternative formulated in the PREDICTION module), along with the information on trip attributes of each alternative such as travel cost and travel time as well as reward points. The individual makes the decision by either accepting or rejecting the promoted alternative, which is captured by the MEASUREMENT module. Lastly, by utilizing the response collected, the UPDATE module updates preference estimates (βs). The updated preference estimates are then fed into the next interaction. Each module is described with more details as follows.

3.2 PREDICTION Module

As a response to an individual's trip request, the PREDICTION module works to formulate a binary choice scenario where (1) the promoted alternative is decided and (2) the number of reward points associated with the promoted alternative is predicted. In this study, we assume that the alternative to promote is known (e.g., when is a better time to depart), while in real-world implementation, we may rely on real-time knowledge on the transportation system to find one. The system predicts the required quantity of reward points by modeling the individual's choice-making behavior.

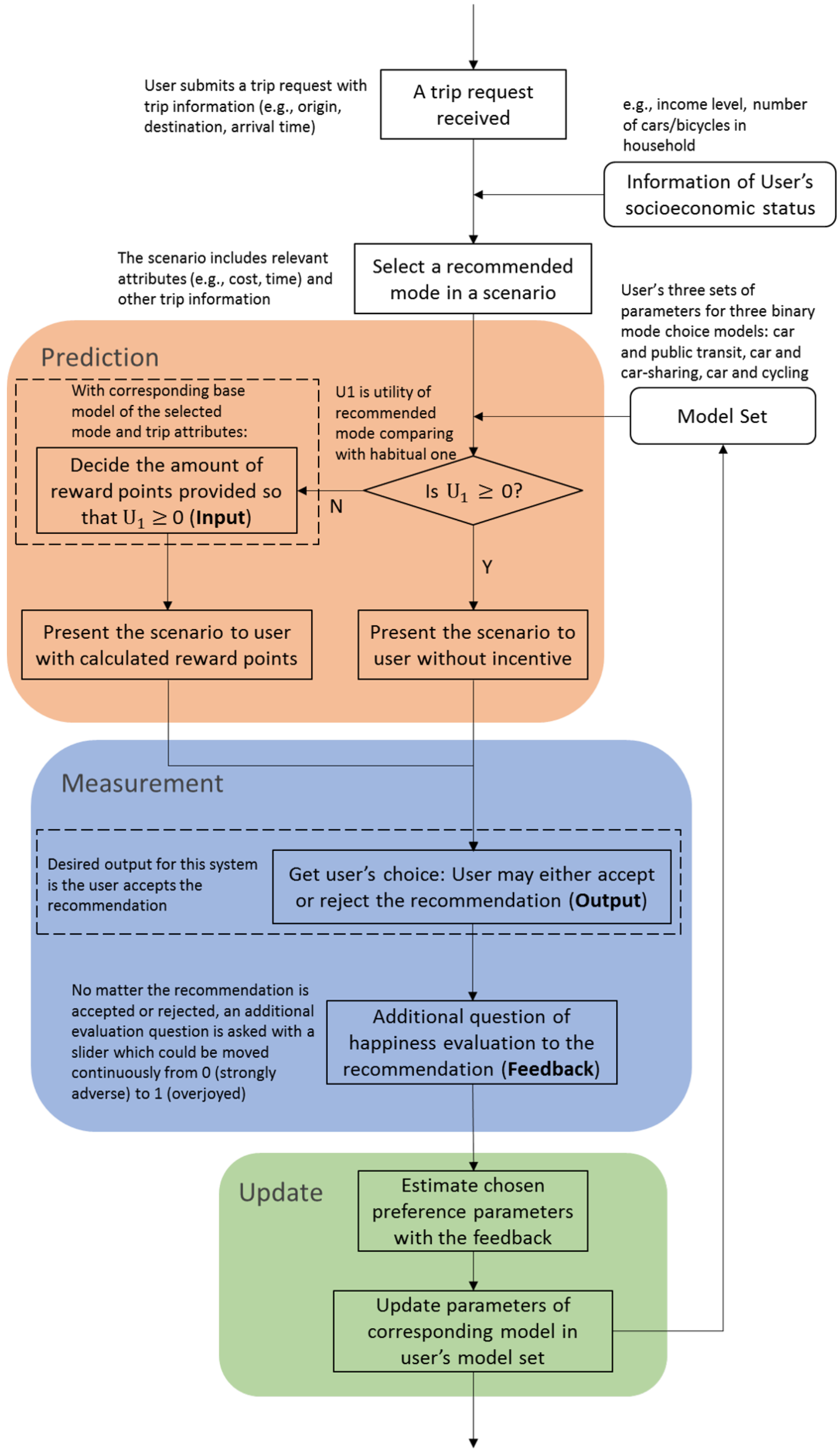


Figure 3.2: System flowchart in detail.

The estimation of the individual's assessment towards the promoted alternative is modeled via the predicted systematic utility $\hat{V}_1^t$:

$$\hat{V}_1^t(RP^t) = \sum_k \hat{\beta}_k^{t-1} \Delta x_k^t + \hat{\beta}_{RP}^{t-1} RP^t \tag{3.1}$$

Here, given that the individual's true preferences are not observable, $\hat{V}_1^t(RP^t)$ is estimated using previously learned preferences including $\hat{\beta}_k^{t-1}$ and $\hat{\beta}_{RP}^{t-1}$, which represent estimated/predicted preferences on trip attribute differences Δx_k^t and on reward points RP^t respectively.

According to RUM, the individual would select the alternative with the maximum utility. Therefore, with an expectation that she would accept the promoted choice, RP^t is calculated such that

$$p_1 = \frac{\exp \hat{V}_1^t(RP^t)}{1 + \exp \hat{V}_1^t(RP^t)} > 0.5 \tag{3.2}$$

Here, p_1 is the probability that an individual is expected to accept the promoted alternative given RP^t points. In our system, we could adjust our expected probability of the alternative to be accepted by pre-specifying p_1 ($p_1 \in (0.5, 1.0)$). Given a specific value of p_1 (e.g., $p_1 = 0.6$), RP^t can be solved as:

$$RP^t = \frac{\log(\frac{p_1}{1-p_1}) - \sum_k \hat{\beta}_k^{t-1} \Delta x_k^t}{\hat{\beta}_{RP}^{t-1}} \tag{3.3}$$

It is possible that Equation 3.3 returns a negative value, suggesting no reward point is needed to promote the alternative. In such a case, we let $RP^t = 0$.

3.3 MEASUREMENT Module

MEASUREMENT module presents the individual the formulated binary choice scenario and in turn, the individual makes his choice. For the particle filter model used in [194], the system needs to solicit additional information (asking respondents to evaluate the attractiveness of the proposed alternative) other than the choice data to learn and update individual preferences. Thus, Figure 3.3 provides an example of the interface with a binary choice scenario and an additional question on how the individual evaluates the relative attractiveness of the promoted choice compared with the default choice. By asking the additional question, the utility ratio R_u is assumed to be observed:

$$R_u = \frac{\exp(U_1)}{\exp(U_1) + \exp(U_0)} \tag{3.4}$$

The individual answers the additional question by adjusting the slider between two ends with one being "not attractive at all" and the other being "absolutely attractive" (see Figure 3.3). Consequently, R_u can be measured based on the position of the slider.

There are errors accounting for omitted factors in U_1 (factors that are important to the individual's decision-making process but not accounted for in the utility function), and the error in adjusting the slider (which translates into R_u) by the individual. We term both types of error as measurement noise in U_1. The measurement obtained through the slider is denoted as M and is expressed as:

$$M^t = \frac{\exp(v_m + \boldsymbol{\beta}^t \Delta \boldsymbol{x})}{\exp(v_m + \boldsymbol{\beta}^t \Delta \boldsymbol{x}) + 1} \tag{3.5}$$

where for simplicity, t is used as the preference vector containing both preferences on trip attributes $\beta_r^t (r \in 1, 2, \ldots, R)$ and the preference on reward points RP^t at time step t, x is used as the attribute vector containing both trip attribute differences and the number of points, and m is referred as the measurement noise (accounting for both types of errors) following a certain distribution, e.g., $v_m \sim N(0, m^2)$. Equation 3.5 serves as the measurement function in the particle filter approach.

3.4 UPDATE Module

Following a measurement on the choice made (or a measurement on the utility ratio acquired by the slider), the UPDATE module works to update the preference estimates with a preference learning and updating model embedded in the system.

In [194], the module works With the particle filter model and updates the preference estimates via a Bayesian estimation scheme. Given the non-linearity in the measurement function (Equation 3.5), the estimation is realized with a particle filter approach. The particle filter approach is a sequential Monte Carlo sampling method and designed to infer

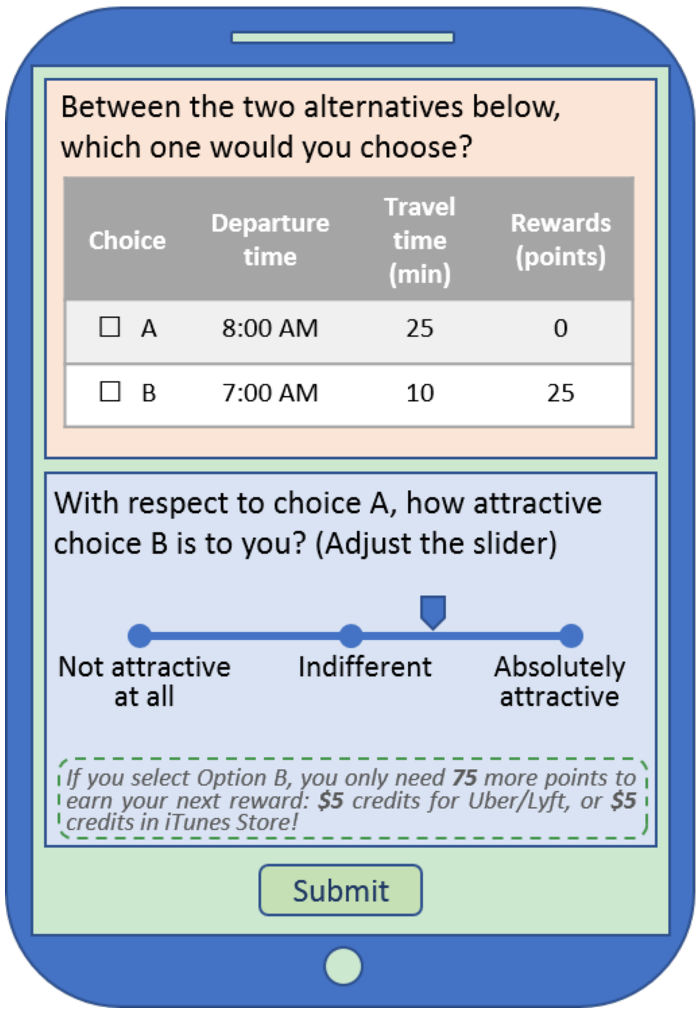


Figure 3.3: An example choice scenario and an additional question.

the state of a system from noisy measurements via a recursive predict-update scheme [34]. In the following, we first describe how the UPDATE model works and then give a summary on its implementation in our system [54].

The preferences $\boldsymbol{\beta}^t$, although not observable, are estimated by utilizing the available information contained in the noisy measurements, such as the one given by Equation 3.5. More specifically, Bayesian estimation calculates a conditional probability density function $p(\boldsymbol{\beta}^t|M^{1:t})$ to represent $\boldsymbol{\beta}^t$ by obtaining all the measurements up to time t (i.e. $M^1, \ldots, M^t$). The conditional pdf is updated according to [54]:

$$p(\boldsymbol{\beta}^t|M^{1:t}) = \frac{p(M^t|\boldsymbol{\beta}^t)p(\boldsymbol{\beta}^t|M^{1:(t-1)}}{\int p(M^t|\boldsymbol{\beta}^t)p(\boldsymbol{\beta}^t|M^{1:(t-1)}d\boldsymbol{\beta}^t} \propto p(M^t|\boldsymbol{\beta}^t)p(\boldsymbol{\beta}^t|M^{1:(t-1)} \tag{3.6}$$

Essentially, to acquire the conditional pdf, we update the prior pdf $p(\boldsymbol{\beta}^t|M^{1:t})$ by taking the new measurement Mt into consideration, with the likelihood function $p(M^t|\boldsymbol{\beta}^t)$ following the measurement function (Equation 3.5, where $m \sim N(0, m^2)$) and rewritten as:

$$p(M^t|\boldsymbol{\beta}^t) = \frac{1}{\sqrt{2\pi}m} \exp\left\{ \frac{-[-\log(\frac{1-M^t}{M^t}) - \Delta\boldsymbol{x}^T\boldsymbol{\beta}^t]m^{-2}[-\log(\frac{1-M^t}{M^t}) - \Delta\boldsymbol{x}^T\boldsymbol{\beta}^t]}{2} \right\} \tag{3.7}$$

And the prior $p(\boldsymbol{\beta}^t|M^{1:(t-1)}$ in Equation 3.6 is

$$p(\boldsymbol{\beta}^t|M^{1:(t-1)}) = \int p(\boldsymbol{\beta}^t|\boldsymbol{\beta}^{t-1})p(\boldsymbol{\beta}^{t-1}|M^{1:(t-1)})d\boldsymbol{\beta}^{t-1} \tag{3.8}$$

where $p(\boldsymbol{\beta}^t|\boldsymbol{\beta}^{t-1})$ is the probability of preference evolution.

We assume that an individual's preference evolution is in the following form:

$$\boldsymbol{\beta}^t = \boldsymbol{\beta}^{t-1} + \boldsymbol{u}, \quad \boldsymbol{u} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}) \tag{3.9}$$

Provided that preferences are stable in the very short term [63, 128], the evolution function (Equation 3.9) suggests that the system takes values from the previous time step $t-1$ for prediction but adds a process noise u. The process noise u here is to account for possible fluctuations of individuals' preferences over time and is also essential in learning dynamic preferences (i.e. individuals' preferences change over time). Equation 3.7 and Equation 3.8 define a recursive way to update $\boldsymbol{\beta}^t$, with the initial pdf available (i.e. $p(\boldsymbol{\beta}^0|M^0)$, M^0 means no measurement available). Due to the non-linearity in our measurement function, there is no analytical closed form solution. An approximate solution is obtained using the particle filter approach, with Monto Carlo sampling.

The key idea of the particle filter is to approximate the posterior pdf with a set of random samples with weights. These samples are particles. When the sample size is large enough, these particles approach an approximate representation to the posterior pdf. Suppose $\boldsymbol{\beta}^t(\boldsymbol{\beta}_j^t, w_j^t)$ denotes a collection of J particles, where $\boldsymbol{\beta}_j^t (j \in 1, 2, \ldots, J)$ is a preference sample and w_j^t is its corresponding weight. A weighted approximation to the posterior pdf is given in Equation 3.10, with weights update given in Equation 3.11 [54].

$$p(\boldsymbol{\beta}^t | M^{1:t}) \approx \sum_{j=1}^{J} w_j^t \delta(\boldsymbol{\beta}_j^t) \tag{3.10}$$

$$w_j^t \propto w_j^{t-1} \frac{p(M^t | \boldsymbol{\beta}_j^t) p(\boldsymbol{\beta}_j^t | \beta_j^{t-1})}{q(\boldsymbol{\beta}_j^t | \boldsymbol{\beta}_j^{t-1}; M^t)} \tag{3.11}$$

Here, $\delta(\boldsymbol{\beta}_j^t)$ is a delta function centered at $\boldsymbol{\beta}_j^t$ and $q(\boldsymbol{\beta}_j^t | \boldsymbol{\beta}_j^{t-1}; M^t)$ is the importance density, which is a known pdf to generate particles. Note here we assume that the importance density only depends on the state and measurement at the previous time step $t-1$. It is often convenient to set the importance density to be the same as the prior $p(\boldsymbol{\beta}_j^t | \boldsymbol{\beta}_j^{t-1})$, such that the weight update Equation 3.11 is further simplified as [54]:

$$w_j^t \propto w_j^{t-1} p(M^t | \boldsymbol{\beta}_j^t) \tag{3.12}$$

With the updated preferences at UPDATE Module, the preferences are prepared for the PREDICTION Module in the next scenario.

3.5 A problem in multi-dimensional learning

The problem of lack of observability arises when there are two or more parameters to be learned, i.e, an infinite set of parameters could give rise to a single M. Two solutions are proposed to address this problem: one being the “divide and conquer” strategy to modify the particle filter as described above and the other imposing domain knowledge on travel behavior as constraints in Monte Carlo sampling.

3.5.1 Solution 1: Divide and conquer

In the usual setup of the particle filter approach, every update utilizes the current observation at time step t after the individual makes a choice on a scenario presented to him. The likelihood function used in the update process is $L = Pr(M^t|\boldsymbol{\beta}_j^t)$, representing the likelihood of observing M^t, given the preferences represented by particle $\boldsymbol{\beta}_j^t$, where $\boldsymbol{\beta}_j^t = \boldsymbol{\beta}_j^{t-1} + \boldsymbol{u}$, with $\boldsymbol{u}$ representing the process noise from $t-1$ to t. In the proposed divide and conquer strategy, a multi-dimensional problem is decomposed into multiple, conditional one-dimensional problems. The idea is that, when different scenarios possess alternatives sharing similar attribute values, the relative attractiveness across alternatives only hinges upon that attribute with different values. In summary, it takes two modifications in the previously introduced update process.

1. Instead of using measurements in Equation 3.5 directly, we define a new measurement function by: (a) identifying scenarios with similar attribute values, leaving only one attribute whose values are different across scenarios; for example, two choice scenarios could be $(\Delta x_1, \Delta x_2, M)$, $(\Delta x_1', \Delta x_2', M')$ where M and M' are two measurements on the individuals' responses to scenarios $(\Delta x_1, \Delta x_2)$ and $(\Delta x_1', \Delta x_2')$, with $\Delta x_1 = \Delta x_1'$; and (b) recalculating a modified measurement on the difference between the two measurements $\bar{M} = M - M'$.

2. Accordingly, we modify the likelihood function (e.g., $L = Pr(M - M'|\Delta x_2 - \Delta x_2', \beta_2)$) such that the likelihood of obtaining the modified measurement is only conditioned upon a single parameter (e.g., β_2), and the attribute it associates with, or the difference between two different attribute values from different scenarios (e.g., $\Delta\bar{x} = \Delta x_2 - \Delta x_2'$). Specifically, the likelihood function in Equation 3.12 is modified as:

$$p(M^t|\beta_j^t) = \frac{1}{\sqrt{2\pi m}} \exp\left\{ \frac{-[-\log(\frac{1-\bar{M}^t}{\bar{M}^t}) - \Delta\bar{x}\beta_j^t] m^{-2} [-\log(\frac{1-\bar{M}^t}{\bar{M}^t}) - \Delta\bar{x}\beta_j^t]}{2} \right\} \quad (3.13)$$

Given that attribute values are available, the one-dimension parameter is thus solvable.

3.5.2 Solution 2: Sampling with domain knowledge

Domain knowledge refers to existing knowledge about travel behavior in existing studies. Past studies in travel behavior have revealed many insights on (1) what factors tend to matter in what types of behavior scenarios? For example, in departure time choice scenarios, only a limited number of factors (e.g., scheduled delay and travel time) are found to matter [64, 119, 132, 49]; (2) how various factors may affect travel behavior choices? For example, the parameters for travel time and travel cost must be negative and a rough range could be identified for each parameter from previous studies. In some cases, not only ranges are available but also their approximated distribution forms. Such knowledge can be useful in the re-sampling scheme within the particle filter process in providing us the information on what types of distributions to be re-sampled from.

3.6 Limitations of the proposed preference learning model

The model utilizing particle filter approach described in this chapter tackles the two challenges in individual preference learning, and succeeds in updating an individual's preferences at each time he makes a choice. However, from the description of the model presented in foregoing paragraphs, there are at least three limitations that can be improved.

For the first, the particle filter model deals with the challenge of the limited observability issue in choices by soliciting additional information on the utility ratio in the MEASUREMENT Module. This allows the system to obtain more information besides the choice behavior of rejection or acceptance. However, it also increases the burden on the respondents, which may cause fatigue and may bring in noises in the observations. Specifically, from the equations, we could see that the preference estimation process of the particle filter model does not utilize the information of the individual's behavior (acceptance or rejection) but only the response of the utility ratio, which means that the individual's inaccurate evaluation will directly impact the preference learning accuracy. It is expected that the model in the UPDATE Module could learn individual preferences from the observations of an individual's choice data.

For the second, the particle filter model does not fully utilize all the information of the data collected in individual preference learning. As we stated in the chapter of introduction, the particle filter algorithm only utilizes the information in the data generated by the individual himself, but the information in the data generated by all the individuals in the whole population is not considered. In the dissertation, we believe that while each individual has his peculiar tastes toward various factors that matter in a choice scenario, there is at least certain degree of commonality in preferences shared by multiple individuals. By exploring the responses from all the individuals, possible common preference patterns or structures may be identified, revealing the similarity and commonality in the heterogeneous population. Integrating the individual- and group-level information may help increase the preference learning accuracy of the individual preference learning algorithm, or at least help when the number of observations from an individual is very limited.

For the third, the particle filter model has a requirement on the travel scenario design. The ‘divide-and-conquer” strategy applied in the particle filter model solves the problem in multi-dimensional learning, however it requires the model to search for and find a scenario that (1) shares similar attribute values with the current scenario, and (2) only has one influence factor that has different values from the current scenario, from the individual's choice scenarios in history. If there are no eligible scenarios found in his historical data, his preferences could not be updated. Moreover, considering the preference changes over time, we might have to add constraints to the temporal intervals in which the eligible scenarios are searched. Thus, to make sure that all the preference parameters can be learned and updated by the model, a carefully-designed set of scenarios is required, which may not be possible in a real-world application. It is expected that the impact of the design of the scenarios would have limited influences on the performance of the individual preference learning model.

For the fourth, the stability of the estimates obtained from the particle filter model is low. The stability of the learned preferences refers to whether the estimates can be in a reasonable range and won't result in extreme values due to one noisy observation. It is easy to obtain extreme or unrealistic values that are significantly different from other estimates with particle filter model and logit models when the number of observations is very small, which is the case here in individual preference learning. The extreme values obtained by the preference learning model can cause problems in the personalized control system. For example, it may result the inaccurate estimation of other dimensions of the individual preferences, extreme values in calculated attributes of the proposed alternative, or the breakdown of the system.

With the personalized control system with particle filter model embedded in UPDATE Module, we are able to conduct an online experiment and collect stated preference data from respondents in real-world. The experiment enables us to learn the system's performance in terms of persuading behavioral changes with personalized incentives, and to obtain a real-world choice dataset such that the performance of other proposed preference learning models can be tested. In the next chapter, we will elaborate on the design of the online experiment in detail. Meanwhile, to solve the three issues in the particle filter model, a new model structure is proposed in later chapters (Chapter 5 and 6).

Chapter 4
THE DESIGN OF AN ONLINE EXPERIMENT

The personalized system introduced in the previous chapter describes an iterative process to collect behavioral data, learn and update an individual's preferences, and provide personalized incentives to trigger desirable behavioral changes. In this chapter, an online experiment is designed and conducted to mimic the personalized control system such that we could learn the performance of the system in terms of persuading individuals for travel behavior changes by learning preferences, and obtain a real-world choice dataset such that the performance of other proposed preference learning models can be tested.

The online experiment is based on Amazon Mechanical Turk (AMT) platform. The primary questions for this AMT experiment are related to the performance of the algorithm and system proposed in [194]: What's the performance of the system in terms of persuading individuals for travel behavior changes by learning preferences? Thus, in the data collection process during the experiment, each individual's preferences are learned via the proposed algorithm in paper [194]. Moreover, the experiment is conducted to explore questions that could only be answered in interactions with individuals in the real world. For example, influential factors of the system performance could be investigated in this process, as information of these factors including trip attributes (e.g., travel time) and individual's socio-demographic characteristics (e.g., age) can only be obtained and tested in the experiment.

The data collected is further used to test the performance of the algorithms proposed in this dissertation.

4.1 Research questions in the experiment

When interacting with an individual (in this AMT experiment), we could never know the true values of his preferences. This means that we could no longer use the accuracy of estimation or percent error to evaluate the performance of the learning algorithm, and we need to develop measures for the performance of the system in the experiment. One direct observation for the performance of this system is that whether the respondents accept the suggested alternatives. Therefore, a metric "acceptance ratio" ("AR", the ratio that the individual accepts the promoted alternative in all choices) is adopted as direct observation and measure on the system performance in this experiment. Specifically, the metric could be written as:

$$AR = \frac{N_{\text{Times of accepting the promoted alternatives}}}{N_{\text{Total times of choice made}}} \tag{4.1}$$

Considering research objectives mentioned in the previous paragraph, there are several research questions in this experiment:

1. As a test on our system in promoting sustainable travels, what is the acceptance rate (AR) of the promoted alternatives?

   By answering this question, we investigate how likely people would accept the alternative promoted by our system.

2. Comparing with what we expect (i.e., the expected probability of an individual accepting a promoted alternative, p_1, is pre-specified when calculating the personalized incentives), we may observe an unsatisfactory acceptance ratio (AR) in the experiment. What are the possible reasons and how we could identify them? What response checks and techniques could we use to reduce the risk?

   This question focuses on anticipating potential reasons/obstacles leading to an unsatisfactory AR, designing checks to identify them, and giving possible solutions. The main obstacles may come from two parts: (a) low quality of collected data (bad data) and (b) the system design itself. We first discuss these two obstacles and come up with checks and possible solutions in the following sub-section.

   (a) Unreliable data from the respondents

When the system interacts with people in our AMT experiment, the quality of survey data might be a crucial influential factor deciding whether the results will be grim or not. Because of this, in the survey, quality control and developing checks to examine response quality are very important. One reason causing the low quality of data may be that respondents are not paying enough attention when completing the survey, which includes two aspects:

i. Not paying attention to understanding the question itself, i.e., understanding what the scenario is and understanding what the questions are asking for

   Our experiment relies on a description of our hypothetical background setting for scenarios. Though we try to make the description concise, it is still possible that respondents might not understand it well and move to answer questions in scenarios. Understanding the hypothetical background setting is very important in our experiment and might directly impact data quality as people's reactions towards the following questions might be different. We need to make sure that the respondents understand the hypothetical background and understand what we are asking for in each question.

ii. Not paying attention giving responses to the questions

   However, understanding the hypothetical background setting could not guarantee that the respondents answer the survey questions carefully and seriously after thinking. It is possible that the respondents make their choices randomly, or move the slider thoughtlessly. For example, we may observe an inconsistency issue—the mismatch between the two questions in a scenario (Figure 4.1): acception-or-rejection question suggests choice B is preferred while the attractiveness evaluation question suggests that choice B is not attractive. The challenge here is how to identify those respondents who answer the survey questions thoughtlessly.

(b) System design

Assuming respondents participate in the survey with attention, some factors in the system design itself would still lead to unsatisfactory AR:

i. Measurement noise

Larger measurement noises may impact AR, though the magnitude of measurement noises is unclear in interacting with respondents. Here, measurement noises from a respondent could be reflected by the uncertainty in the observed utility ratio R_u. For example, even for two identical scenarios, we may observe two different utility ratio. A large measurement noise leads to low accuracy in leaning preferences, resulting in unsatisfactory performance in providing personalized incentives. We may observe another inconsistency issue: though the utility of the promoted choice in a scenario is expected to increase (decrease) compared with which in another scenario, the utility ratio R_u in the former scenario evaluated by an individual does not correspondingly increase (decrease). For example, since the rewards contribute positively to the utility, we expect to observe a higher R_u given more reward points with other trip attributes not changed. However, we may observe a decreased R_u. This issue may be serious in cases where two scenarios have similar attribute values, which easily confuses respondents as it is not easy to give responses by perceiving differences among them.

ii. The list of reward items

Respondents would perceive the value of reward points by checking the reward list, which shows items that points can be exchanged for. The system will fail to persuade respondents if the reward list has not items that are attractive to them.

3. How characteristics of travelers and trip scenarios play a role in affecting the acceptance ratio?

   As we concern individual-level preference learning in the decision-making process, factors including some trip attributes (e.g., travel time, trip schedule flexibility, travel cost, trip purpose) and individual's socio-demographic characteristics (e.g., age, gender, education, household characteristics) could be observed with real interactions with people. We would like to explore how those factors and the performance of the system are related. By exploring the extent of these relationships, we could understand the behavior of the system and assess its potentials in affecting different individuals and in different situations. With AR being a measure of the system performance, the relationship between these factors and AR could be investigated via regressions. These factors are classified into several categories and are to answer sub-questions as listed below:

(a) Factors relate to travel attributes (including main factors in the model), such as travel distance/time, work schedule flexibility (in flextime or fixed-time work schedule), travel cost, trip purpose (travel from or to home; commuting or non-commuting), etc.

(b) Factors relate to personal attributes/attitudes. Personal attributes, such as age, gender, income, profession, household structure, etc. Personal attitudes, such as whether an individual has a regular/habituated travel behavior or has developed a reliance on driving.

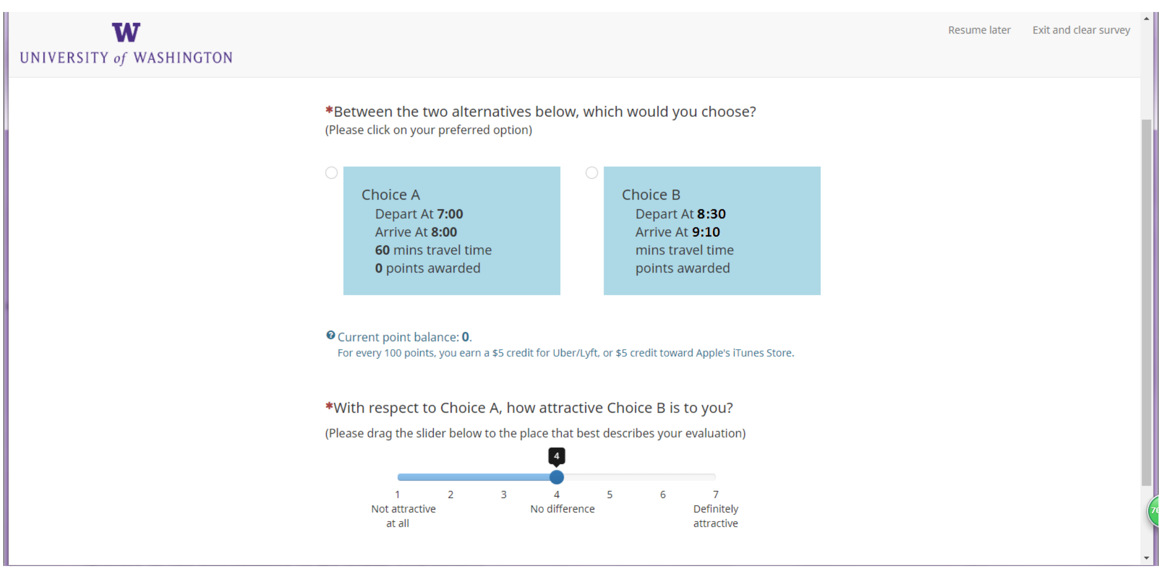


Figure 4.1: Screenshot of a scenario presented to an individual.

4.2 Overall design of the experiment

4.2.1 The general plan

The current AMT experiment focuses on studying departure time change. We consider this type of behavioral change because that departure time change is more applicable in most cities, as they rely less on building new facilities compared with persuading mode change, such as expanding the coverage of transit networks or constructing a ride-sharing matching platform.

To answer those research questions as we discussed earlier, there will be several independent sub-experiments, each of which is conducted using an online survey. Each survey will be completed by a random group of respondents. According to the research questions, each survey may be different from other surveys in typical commuting time or types/levels of flexibility in scenario setting. Steps designing and implementing a sub-survey are discussed in detail in the following subsections.

4.2.2 Survey design

To answer the questions raised earlier, the survey contains four sections (Figure 4.2). Section A aims to investigate the individual's travel behaviors. Section B will be a series of hypothetical choice scenarios, which is the most important part of the survey. Section C collects some socio-demographic information from the respondent. Section D is a short social desirability scale evaluating whether a respondent is more likely to behave in a socially desirable way. In the following, we first brief some preparation works for survey design/implementation, then talk about each section in the survey, focusing on Section B.

4.2.2.1 The beginning of the survey

1. Sample screening

   Focusing on promoting sustainable alternatives, the targeted population in our study is the group of people who usually take single driving trips. With the filtering feature on AMT, we limit our sample framework to those AMT users who are adults and own

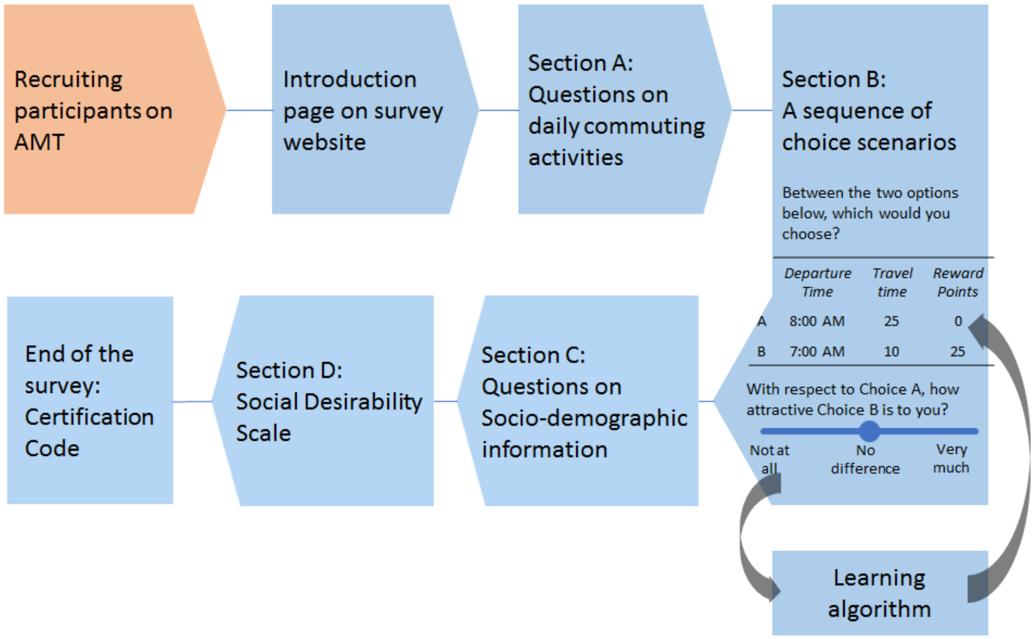


Figure 4.2: The flowchart of the online experiment.

at least one car in their households.

2. Webpage design for online survey

Due to the complexity of conducting our interactive survey and the limited function of the AMT platform, we hire a professional programmer to design an independent survey platform (a web page). Therefore, the planned HIT (Human Intelligent Task) on AMT will be a one-page website showing a brief introduction on the task, a link to our independent web page, and an input box for certification code. Each individual could access our survey via the link on the HIT page, and after the individual finishes the survey, he will be assigned a unique certification number. The individual will be required to enter the assigned code to the AMT page so that we could match his work on our survey web page and his ID on AMT.

4.2.2.2 Survey Section A

Section A collects data on the individual's travel habits and other activity-related information, such as work schedule flexibility, single car trip frequency, etc. Some questions in Section A are conditional, thus may not be applicable for every respondent.

4.2.2.3 Survey Section B

Section B contains a series of hypothetical choice scenarios. In each scenario, information provided includes trip purpose, trip-related attributes, reward points, and a list of rewards that could be exchanged with a certain amount of reward points. We place the list together with the scenarios instead of the beginning of the survey to ensure that the individual is aware of the value of the reward points when he makes choices. In the following, we describe each step to formulate choice scenarios.

1. Model selection: decide the travel choice-making model (the utility model)

   An individual schedules her travels by balancing traveling in congested traffic and a desired arrival time. To encourage an individual to depart at a time out of the rush hour which might contribute congestion (according to the literature review, the default departure time is assumed within the rush hours), we provide a certain amount of incentives so that the scheduled time delayed using the promoted alternative is compensated.

   Since preference estimates are learned using data from one individual only, only alternative-based attributes contribute to the utility model. Travel time and schedule delay (early or late) are selected as relevant alternative-based attributes based on existing studies, which report that these two factors are the primary ones in departure time choices for commuting trips [64, 162, 119, 132, 49, 91]. Consequently, our utility model involving travel time, schedule delay and reward points can be expressed as:

$$U_1 = \beta_{TTS} \times TTS + \beta_{SDE} \times SDE + \beta_{SDL} \times SDL + \beta_{RP} \times RP + \epsilon \qquad (4.2)$$

Here, SDE and SDL are minutes scheduled earlier and later than the desired arrival time, respectively, and for any individual, only one of them is used. TTS is travel time savings in minutes, and RP is the number of reward points. Coefficients of SDE and SDL are expected to be negative, while those of TTS and RP are positive.

2. Choice scenario design

   Our goal is to design a set of scenarios from which we could get as much information from the individual as possible so that the preference learning would be efficient. This means there should be enough variations in the combinations of attributes in the scenarios. We also need to make sure that all the scenarios are reasonable and realistic.

   Considering the utility function and our research questions, information related to 5 attributes/factors may be presented to respondents in scenarios, including SDL/SDE, travel time saved, travel time, schedule flexibility and the type of incentive strategy.

   The table below gives a summary of attributes with their levels considered in the experiment.

   The design of the levels of each attribute refers to the existing studies and reports.

   (a) SDE/SDL

      [91] suggests that for most of the commuters, the earliest acceptable arrival time is one hour ahead of the working start time. Unfortunately, no schedule delay late (SDL) is reported. No statistical description of the specific range of schedule delay is reported.

      With a survey dedicated to studying commuters' time choice, [163] reports that a regular time of arrival 42.5 minutes early and 17.5 minutes late.

      With a stated preference survey, [49] reported that commuters are willing to trade 1.23 min of travel time for every 1 min of late schedule delay and 0.83 min of travel time for early schedule delay. Therefore, the late schedule delay is, on average, more onerous than the early schedule delay.

Table 4.1: Attribute levels in the experiment.

Attributes	Levels
Attributes in utility function	
SDL/SDE	3 levels for SDE: 10, 30, 60 min 3 levels for SDL: 10, 30, 60 min
TTS (percentage of original travel time)	L1 slight congestion - 10% L2 severe congestion - 60%
Factors that may be considered in randomized group experimental design	
Flexibility	L1 no flexibility - "your company requires all employees to arrive no later than 8:00 am." L2 low flexibility - "your company encourages all employees to arrive no later than 8:00 am." L3 high flexibility - "your company has no requirements for arrival time, but most of the staff arrive between 7:00 am and 9:00 am."
Original Travel Time	L1 short commute - 10 min L2 moderate - 25 min L3 long commute - 60 min

Based on previous reviews, the levels for scheduled delay in this experiment are set as: 3 levels for SDE (10, 30, 60 min) and 3 for SDL (10, 30, 60 min).

(b) TTS

For delay analysis, referring to TTI 2015 Urban Mobility Scorecard [158], in peak travels, trips would experience extract travel time: about 10% with low congestion level and 60% with severe congestion level. Therefore, these two levels will be reflected in attribute Travel Time Saved, meaning in our hypothetical scenarios, by changing departure time, a commuter could save his travel time by either 10% or 60%.

(c) Average travel time

Referring to the National Report on Commuting Patterns and Trends from AASHTO in 2013 [135], average travel time for people who drive alone is 24.19 minutes. For carpooling, the average travel time might be 38.89 minutes. The report uses 20 minutes and 60 minutes as dividing points for short commutes and long commutes. Thus, 10 minutes, 25 minutes, and 60 minutes are selected to represent three commute time levels in the scenarios, where about 10% of people have commuting time shorter than 10 minutes, and about 10% of people have commuting time longer than 60 minutes.

An example of the scenario background setting description is shown in the following. Then the two alternatives presented to the respondent are shown in Figure 4.1.

> Imagine that recently your company moved to a new place, and *your company encourages all employees to arrive no later than 8:00 am*. With your experiences during the last week, you find that if leaving home at *7:00 am*, it would take you 60 minutes to reach your office by driving due to the heavy traffic.
>
> Also imagine that now your local Department of Transportation launches a program, which could suggest you a departure time (might be earlier or later than *7:00 am*) according to the real traffic situation. By adopting the

suggestion, you could likely spend less time on road. As an incentive, you earn a reward if you change your departure time following the suggestion. Of course, you can reject the suggestion, still choosing to depart at *7:00 am* as what you did in last week.

Each of the following scenarios contains a description of a suggested departure time to travel to work. Remember that the "default alternative" (the commuting choice you find, namely, departure at *7:00 am* and arrive at *8:00 am*) is always an option, so that you could compare it with the suggested alternative (depart at *8:30 am* and arrive at *9:10 am*) and make your choice based on the information given in each alternative.

(Contents in italic could be replaced by the information of other Flexibility levels of Travel Time levels)

3. Reward points: Assign the value of reward points, and design a list of rewards that the respondents can exchange to

   The individual shall know the value of the reward points he earns from a travel behavioral change, so that he can evaluate the utility of the promoted alternative together with the incentive we provide. The worth of reward points is reflected by the rewards that the individual could be used to exchange to, which is indicated by placing a note (i.e., a sentence) together with the scenario.

   The sentence in the scenario could be:

   "For every 100 points, you earn a $5 credit for Uber/Lyft, or $5 credits towards Apple's iTunes Store."

   Since the individual may view trivial the points earned at one time, to enhance the influence of the incentive, besides indicating the points earned at one time, we display the total points accumulated right above the sentence. See Figure 4.1 for an example.

4.2.2.4 Survey Section C

After finishing the sequence of binary choices, in Section C, the individual is requested to answer some questions related to her socio-demographic characteristics, so that we investigate how these characteristics would affect the system performance.

4.2.2.5 Survey Section D

Social desirability is a common type of bias that impacts the validity of survey data [131]. We use a Social Desirability Scale (SDS) to test whether participants' responses are resulted from their tendency of behaving in a socially desirable way, rather than reflecting their own preferences. Participants were asked at the end of the experiment to answer a short SDS survey with ten items [57, 171]. This is one of the data quality control methods used in the online experiment.

4.2.3 Randomized experimental design

Since the experiment also aims to study the impacts of travel time and flexibility, in the online experiment, the randomized experimental design is conducted.

In our preliminary consideration for Flexibility and Travel Time, respondents could be randomly distributed into 9 groups using a full design. Different groups are different in terms of the combination of these three factors, reflected by the scenarios presented to respondents in each group: the description of the background setting (e.g., the flexibility requirements of the company) and the default choice are different. As we will utilize the same platform (AMT) and the same sample screening method to select both groups of respondents, we assume that each group of respondents represents the same population on the AMT platform.

Moreover, since the experiment is conducted online on the AMT platform, several strategies are designed and used in the survey to control the response quality in the data collection process. Details of the quality control strategies are introduced in the next subsection.

Table 4.2: Randomized experiment design

	Flexibility	Travel Time
Group 1	L1	L1
Group 2	L1	L2
Group 3	L1	L3
Group 4	L2	L1
Group 5	L2	L2
Group 6	L2	L3
Group 7	L3	L1
Group 8	L3	L2
Group 9	L3	L3

4.3 Response quality control and data cleaning

4.3.1 Response quality control strategies

AMT has been a useful tool and applied to various types of research and experiments across disciplines, with comparisons between AMT and other tools suggesting the reliability of AMT for data collection [28, 44, 70]. However, the same as traditional offline surveys, data collected from AMT may contain data/responses of low quality (e.g., careless responses). To minimize its effects, we design the experiment using standard quality-control techniques and identify data of low quality [90, 126]. In our experiment, reasons getting low-quality data include (1) since the experiment, as a stated preference survey, relies on participants' understanding in the hypothetical background setting, those who do not pay attention to the introduction of our background setting will produce unqualified responses; (2) motivated by only the payment for participating the experiment, some participants may give responses randomly and speedily; (3) some participants may respond in a way that is socially desirable. With all these concerns, we apply several techniques to designing the experiment and develop metrics to examine the data quality of a total of 926 completed surveys we collected [20, 120, 126].

Several strategies are applied to control the quality of the responses. In the following, these strategies are briefly discussed together with the outcomes.

1. Check participants' understandings in background settings

   After introducing the background settings and before proceeding to scenario choice questions, participants are required to answer three simple and straightforward questions on the background settings. Those who fail any one of those three questions are forced to quit the experiment. Our database shows that besides those 926 participants who successfully finished the experiment, about 300 participants failed to check questions and quit the experiment, suggesting that these check questions effectively-identified participants who did not pay attention to understanding the background settings.

2. Consistency check

   We assume those who provide inconsistent responses were not paying attention to the experiment. Questionnaires containing a large number of inconsistent responses should be removed. The inconsistency could be checked in two ways: (1) response is inconsistent if a participant rejects (accepts) a recommendation while evaluating it as attractive (not attractive) using the slider; (2) a response to the current recommendation is inconsistent, if, compared with the historical response to a previous recommendation, the current one is of higher utility but is given a lower evaluation in terms of its attractiveness (we assume a higher RP, higher TTS, or lower SDE/SDL gives a higher utility). We observe that while the maximum number of inconsistent responses is 20 for a participant, more than 60% of the participants have fewer than 5 inconsistent responses.

3. Repeated question

We repeat the first scenario at the end of Section B, and the differences within responses on the two scenarios are used to qualify whether the participant is paying attention throughout the experiment. In our sample, 84.6% of the participants make consistent selections (i.e., among the 84.6% of questionnaires, the rejection (acceptance) choices are consistent between the first and the last scenarios). And 79.9% of the participants give two evaluations on the attractiveness of the same recommendation in the two scenarios with a difference no larger than 0.2 ($\Delta M \leq 0.2$).

4. Timing

   For each participant, the time spent responding to each choice scenario is recorded. We assume that too long or too short timing can be a potential indicator that participants did not respond to presented scenarios thoughtfully. In our sample, 60.0% of participants finished the sequence of hypothetical choice scenarios between 110 to 260 seconds (i.e., 8.46 to 20 seconds per scenario), with a median at 195 seconds.

5. Social Desirability Scale

   As introduced earlier, Section D in the survey is a short Social Desirability Scale, working on testing the existence of the social desirability bias in the responses. The replies from respondents who have been flagged by the scale may need further investigation in the data analysis stage.

4.3.2 Dataset after cleaning

In the data cleaning process, each questionnaire is assigned with a score as an indicator of its quality. Specifically, a questionnaire is ranked based on measurement from each check technique and the score is calculated as the sum of all ranks. Figure 4.3 gives the cumulative distribution of scores for all participants, with a higher score indicating a lower quality. Based on the Elbow rule, it is shown that a score at 3 is a good threshold classifying questionnaires (Figure 4.3). Based on this threshold, we remove 98 surveys (about 10.6%).

Finally, we have 826 eligible questionnaires in total, 387 of which are completed by females and 439 by males. About 70% of the participants are 40 or younger. The median

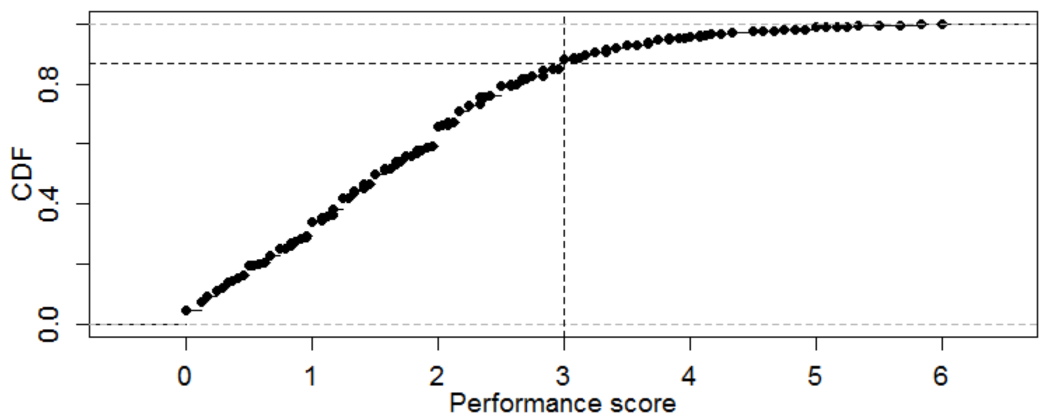


Figure 4.3: Cumulative distribution for the performance score.

annual income falls between \$50,000 and \$75,000. According to their stated commuting time in their daily lives, the average commuting time is about 31.58 minutes. The correlation between the SDS score and AR is -0.003, which indicates that social desirability does not play a significant role in affecting participants' different choice-making in our survey. The gender, age, and income level distributions are shown in Figure 4.4.

This cleaned dataset of 826 individuals is used to test the performance of the proposed models in the following two chapters.

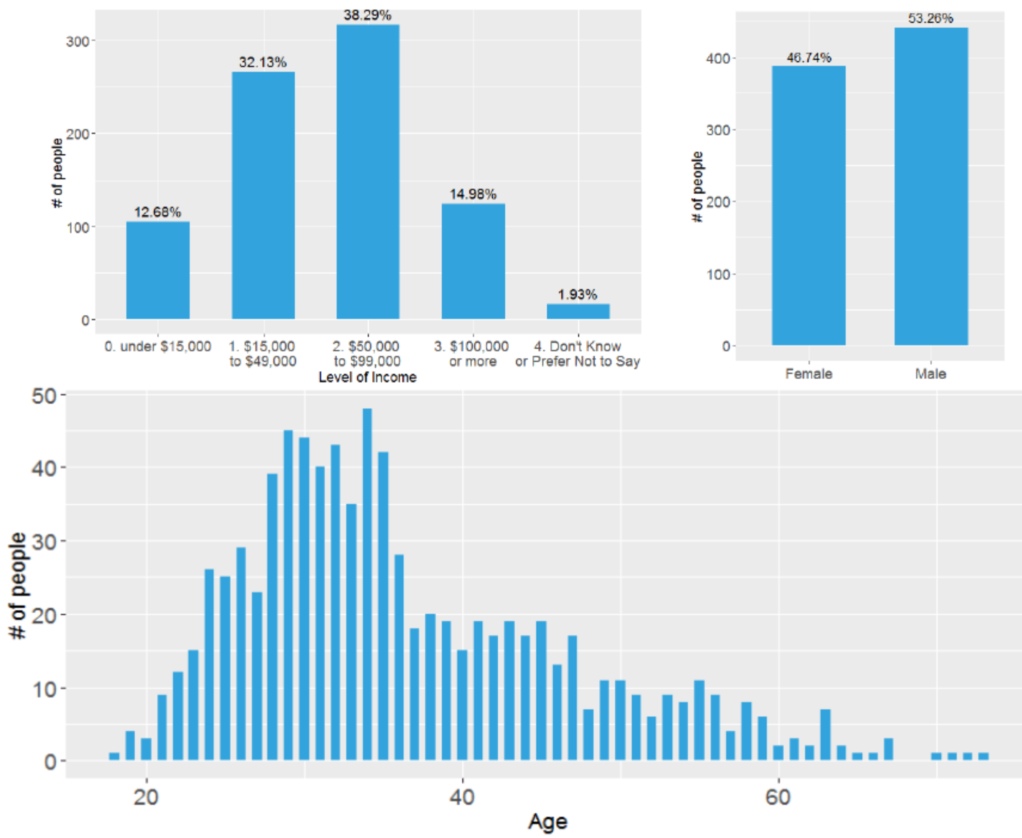


Figure 4.4: The distributions of respondents' age, gender, and income levels.

Chapter 5

INTEGRATING COLLABORATIVE LEARNING MODEL WITH A SINGLE-VARIABLE POLYNOMIAL

5.1 Personality, commonality, and stability in the collaborative learning model

Assume that there is a group of 300 individuals, each has his own preferences on Scheduled Delay Late β_{SDL} (to arrive at the destination later than original scheduled) and Reward Point β_{RP}. Also assume that we know the preferences of all these people, and could draw the distribution of the preferences as shown in Figure 5.1a. We could roughly tell that there are three groups of individuals in the population, which can be seen as three different types of individuals: Individuals of Type A dislike Scheduled Delay Late very much, and also don't feel the Reward Point can compensate; individuals of Type B also don't enjoy Scheduled Delay Late either, but feel that Reward Point is very attractive; individuals of Type C enjoy both Scheduled Delay Late and Reward Point. The preferences of the three different types of individuals, and the individuals in each group could be identified by applying any clustering method to this preference dataset. From the figure we could also see that β_{SDL} and β_{RP} of all the 300 individuals have boundaries: $\beta_{SDL} \in [-0.4, 0.2]$, and $\beta_{RP} \in [-0.1, 0.6]$, and the preferences are roughly distributed within a bounded triangle area.

However, in common cases, we do not know individuals' preferences, and need to estimate with some methods. As we discussed in Section 3.6, the stability of the estimates obtained from methods such as logit model or particle filter is low, meaning that the estimates can be extreme or unrealistic values due to noisy observations. The stability issue could be severe when the number of observations is small. To increase the stability of the estimates, the preference learning model is expected to be able to add constraints, or restrict the estimates in certain ranges in the learning process. For example, in the current case with the 300 individuals, if we could identify the triangle area and restrict the estimates in the area at each time step, the estimated preferences will be more stable without extreme values.

This is how the collaborative learning model provides stable estimates in individual preference learning: identifying the preference area with all individuals' preferences, and formulating each individual's preference in a way such that it would be in the area. We'll briefly illustrate how this works in collaborative learning model with our example.

To identify the area, we first identify the vertexes. Assumes that there are three individuals, who may or may not really be in the population. The first individual i_1 with preferences $(\beta^{i_1}_{SDL}, \beta^{i_1}_{RP})$ is an extreme person of Type A. The "extreme person" of Type A means that for any individual with preferences $(\beta^{i}_{SDL}, \beta^{i}_{RP})$ in the population, he would either dislike SDL less than individual i_1, or dislike RP less than i_1, i.e., $\beta^{i}_{SDL} \leq \beta^{i_1}_{SDL}$ or $\beta^{i}_{RP} \leq \beta^{i_1}_{RP}$. Similarly, the second individual i_2 with preferences $(\beta^{i_2}_{SDL}, \beta^{i_2}_{RP})$ is an extreme person of Type B. Any individual i with $(\beta^{i}_{SDL}, \beta^{i}_{RP})$ would either dislike SDL less than individual i_2, or like RP less than i_2, i.e., $\beta^{i}_{SDL} \leq \beta^{i_2}_{SDL}$ or $\beta^{i}_{RP} \leq \beta^{i_2}_{RP}$. For the third individual, we have that any individual i with $(\beta^{i}_{SDL}, \beta^{i}_{RP})$ would either like SDL less than individual i_3, or like RP less than i_3, i.e., $\beta^{i}_{SDL} \leq \beta^{i_3}_{SDL}$ or $\beta^{i}_{RP} \leq \beta^{i_3}_{RP}$. Since the three "extreme persons" are the extreme cases of the three types of preferences in the population, they represent three types of preferences that exist in the population and shared by all the individuals - Let's call them the three "common types of preferences".

In Figure 5.1b, three red points are added to represent the three common types of preferences. The preferences $(\beta^{i}_{SDL}, \beta^{i}_{RP})$ of any individual i could be written as $(\beta^{i}_{SDL}, \beta^{i}_{RP}) = c^i_1(\beta^{i_1}_{SDL}, \beta^{i_1}_{RP}) + c^i_2(\beta^{i_2}_{SDL}, \beta^{i_2}_{RP}) + c^i_3(\beta^{i_3}_{SDL}, \beta^{i_3}_{RP})$, where $i \in (1, 2, \ldots, 300)$, $\forall c^i_r \geq 0$ and $\sum_{r=1}^{3} c^i = 1$.

Now we have a new system to express each individual's preferences. Remember that each vertex represents a common type of preference in the population - we obtain the commonality from all the individuals' preferences. With the parameters c^i, we know that each individual i's preferences resemble the preferences of each common type of preferences to different degrees, which reflects each individual's personality (we could also see that as vertexes, the three common types of preferences do not resemble the other two at all, they should be significantly different from each other). Since each individual's preferences are generated by combining the three common types of preferences, the estimates would always be constrained by the preferences of the three common preference types and have upper and lower boundaries. In other words, they will always be stable and no extreme values would generated given an noisy observation.

From this simple case, we may have a feeling about how the collaborative learning model works. For a certain objective (e.g., preferences – as shown in the previous example, preference changing patterns – as shown in the following sections) that we are exploring, collaborative learning model identifies the common types of the objective in the population, and use these common types of objectives to express that of each individual with a set of individual-specific parameters c_r^i. The common types of objectives are called canonical models in the collaborative learning model, while the vector of the parameters c_r^i is called the membership vectors. Because the collaborative learning model successfully solves the problems we raised in previous, in the dissertation, the collaborative learning model is utilized for individual preference learning and capturing preference changes over time.

Not like the example in this section, in reality, we could not have a figure showing the distribution of the preferences of all the individuals, and we may not know that there should be three vertexes/common types. Typically, different numbers of vertexes are tested, and the most suitable number of vertexes will lead to the best performance of the model.

5.2 Collaborative learning model in individual preference learning

5.2.1 Time-varying preference $\beta(t)$ in Logistic Collaborative Model

The focus of the current chapter is to elaborate on an online preference learning and updating model whose basic structure is integrating the canonical structure and a time-dependent preference model. The proposed model, the Logistic Collaborative Model with Time-varying Parameters (LCM-T), allows an individual's preferences to vary over time in the individualized modeling. In other words, while in the traditional logistic model for binary choices, the probability of an alternative being chosen by individual i at tth time step ($y_{it} = 1$) is

$$Pr(y_{it} = 1|\boldsymbol{x}_{it}) = \frac{\exp(\boldsymbol{\beta}_i^T \boldsymbol{x}_{it})}{1 + \exp(\boldsymbol{\beta}_i^T \boldsymbol{x}_{it})} \tag{5.1}$$

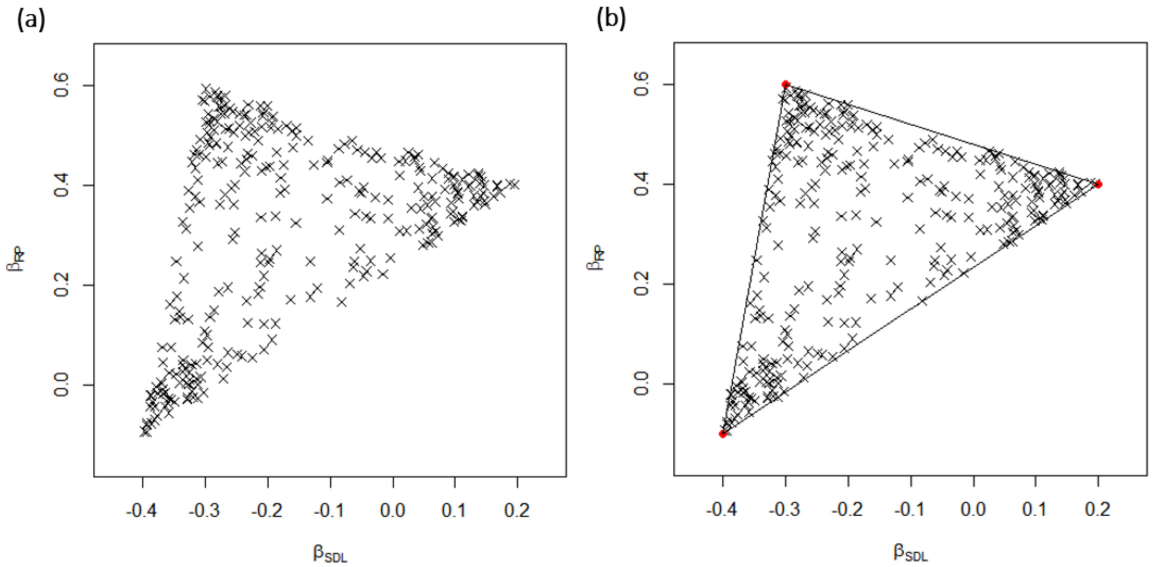


Figure 5.1: Distribution of the (a) two dimensional preferences of 300 individuals, and (b) three vertexes of the triangle area.

where $\boldsymbol{x}_{it}$ is the differences of the factor variables (attributes) between two alternatives at time step t. In this paper individual i's preference vector $\boldsymbol{\beta}_i$ is converted to $\boldsymbol{\beta}_i(t)$, which means that the vector $\boldsymbol{\beta}_i$ may be different when time t is different. An individual's preferences may vary according to specific scenarios/attributes, or evolve gradually with personal experiences [21, 97, 81, 104]. Considering the possible influential factors, there can be various time-dependent models for $\boldsymbol{\beta}(t)$. In the current study, we use polynomial models as the formulation of the time-dependent model $\boldsymbol{\beta}(t)$. The polynomial models allow us to express time-dependent variations, while the only influential variable is the time step t. To further prove that the polynomial model could capture preference changes in individual choice-making process, we have also run some regressions with individual preference data. In our previous study [194], we have collected responses for a sequence of 13 travel choice scenarios from 826 individuals and learned their preferences during the process with the particle filter approach. We ran polynomial regression models with the learned preferences. The results show that either the linear, quadratic or cubic polynomial model can describe the preference variation statistically significantly. Hereafter from Section 5.4.2 though Section 5.5, we use a formulation of cubic model $\beta(t) = q_0 + q_1 t + q_2 t^2 + q_3 t^3$ as an example to construct our model and introduce our method.

5.2.2 *The model formulation of logistic collaborative learning with time-varying parameters (LCM-T)*

5.2.2.1 *The collaborative learning framework with time-dependent preference: illustration with a simple example*

We will first describe the idea of the collaborative learning structure and the proposed model with the following example.

Assume that a group of people has made their travel choices for morning commuting for several days. When deciding whether to bike or not, they all concern three factors: total commuting time, their personal preferences on biking, and the weather. Different individuals may value each factor differently when making their commuting choices. We could say that an individual's preference on each factor forms a dimension of an individual's preferences, i.e., the preferences have 3 dimensions: preference on commuting time, preference on biking, and preference on the weather. It can be imagined that when the attributes of these factors in choice scenarios are different, people's preferences may change. For example, given that the weather is cloudy, some people's preference on biking may fall due to a concern of rain. Some other people, however, may feel it a good temperature and humidity to bike to work. Thus, when the weather changes every day, an individual's preference on biking may also vary accordingly, e.g., the preference rises from 0.5 to 0.8 on the second day which is cloudy, falls to -0.3 on the third day which is a rainy day, and so on. We define the curve that describes how the preference value changes over time or in a sequence of scenarios as "preference changing pattern". It can also be imagined that the changing patterns of different factors may not be the same: on a cloudy day, when an individual's preference on biking rises, his preference on commuting time or commuting cost may not rise correspondingly. Thus we could see that the changing patterns of each preference dimension may be discussed separately.

To simplify in the explanation, let us only focus on one preference dimension (i.e., the preference on biking) as an example, and assume that the changes of the preference on biking will only be impacted by weather. Also, assume that we have only two types of common preference changing patterns in the group: for preference changing pattern $\beta_A(t)$, an individual's preference on biking will be 1 on a sunny day, be 1 on a cloudy day, and be 0 on a rainy day; for preference changing pattern $\beta_B(t)$, an individual's preference on biking will be 0 on a sunny day, be -1 on a cloudy day, and be -1 on a rainy day. We can see that these two types are significantly different from each other: $\beta_A(t)$ represents a bike-lover since the values of preference on biking in the choice-making process are always non-negative, while $\beta_B(t)$ may represent someone who never bikes as the preference curve keeps staying in the non-positive zone. These people exist, yet most people, as we could imagine, are likely to be in between: in some choice scenarios, their preference on biking is positive, and in some others, the preference is negative.

According to the collaborative learning framework, the two common preference changing patterns $\beta_A(t)$ and $\beta_B(t)$ identified from this group of people are called "canonical models", which form the basic references for all the individuals' preferences on biking. Specifically, each individual's preference on biking is assumed to be obtained with a combination of $\beta_A(t)$ and $\beta_B(t)$. For example, an individual's preference on biking can resemble $\beta_A(t)$ at a degree of α. Since we assume that there are only two types of preference changing patterns here in the group, this individual's preference changing pattern would then resemble $\beta_B(t)$ at a percentage of $1-\alpha$. The individual's preference on biking would be $\alpha\beta_A(t) + (1-\alpha)\beta_B(t)$. The vector of the parameters $(\alpha, 1-\alpha)$ is called this individual's "membership vector" of his preference on biking. Since each individual may have different values of α, the preferences of each individual would be different.

With the example, we briefly describe the collaborative learning framework and the preference changing patterns we mentioned in previous sections. We interpret the identified "canonical models" as a set of "common types of preference changing patterns", meaning that these changing patterns are shared by all the individuals in the group, as each individual's preferences are composed of these canonical models with weights. Now, we can move to the general description and mathematical formulation of the proposed Logistic

Collaborative Model with Time-varying Parameters (LCM-T).

5.2.2.2 Mathematical formulation of Logistic Collaborative Model with Time-varying Parameters (LCM-T)

As we state in the introduction section, the collaborative learning framework exploits the underlying canonical structure of the preference changing patterns in the population with heterogeneity. With the framework, several common preference changing patterns are identified from all individuals' data, which are assumed to be the basic structural elements of the individual preferences for all individuals. Specifically, each individual's preferences are generated by linearly combining all the common preference changing patterns, in which the weights towards these common preference changing patterns are identified from his own choices. For each individual, the weights form his "membership vector", which describes to what degree his own preference changing pattern resembles each of the common patterns.

Now we write this model in a mathematical way. Assume that an individual's preferences have R dimensions, meaning that there are R attributes in each choice scenario, and $\boldsymbol{\beta}_i$ in Equation 5.1 is an R-dimensional vector. That means the individual i's preference vector is:

$$\boldsymbol{\beta}_i = [\beta_{i1}(t), \beta_{i2}(t), ..., \beta_{iR}(t)]^T \tag{5.2}$$

Also assume that for individual i, the changing pattern of the rth dimension of his preferences can be described by a cubic polynomial function $\beta_{ir}(t) = q_{ir,0} + q_{ir,1}t + q_{ir,2}t^2 + q_{ir,3}t^3$ ($r = 1, 2, ..., R;\ i = 1, 2, ..., N$), as discussed in Section 5.2.1. For the whole population, we assume that there are K_r canonical models $(f_{1,r}(t), f_{2,r}(t), ..., f_{K_r,r}(t))$ for rth preference dimension $\beta_r(t)$, representing K_r common changing patterns identified from all N individuals. Thus, formulation of the canonical models should be the same as the changing pattern of the preference dimension, which means the cubic polynomial equation, i.e., $f_{k,r}(t) = q_{kr,0} + q_{kr,1}t + q_{kr,2}t^2 + q_{kr,3}t^3$. Assigning each individual i with a membership vector $\boldsymbol{c}_{i,r} = [c_{1,ir}, c_{2,ir}, ..., c_{K_r,ir}]^T$, $\sum_k c_{k,ir} = 1$, representing the degree of resemblance of the individual's rth dimension of his preferences $\beta_{ir}(t)$, to the canonical models. Now, with the canonical models and the individual's membership vector, the individual i's preference $\beta_{ir}(t)$ could be expressed with a linear combination of the membership vector and the canonical models: $\beta_{ir}(t) = \sum_k c_{k,ir} f_{k,r}(t),\ k = 1, 2, ..., K$. Similarly, we could assume that other preference dimensions for individual i can also be expressed with the corresponding membership vectors and the canonical model sets. In total, there will be R different sets of canonical models for all R dimensions of the preferences, and each individual has R membership vectors correspondingly. With the canonical models and the membership vectors, all the individuals' preferences can be expressed.

Now we rewrite some of the notations such that we could use them in the formula expressions. The parameters of a canonical model $f_{k,r}(t) = q_{kr,0} + q_{kr,1}t + q_{kr,2}t^2 + q_{kr,3}t^3$ could be denoted by a R-dimensional parametric vector: $\boldsymbol{q}_{k,r} = [q_{kr,0}, q_{kr,1}, ..., q_{kr,R}]^T$. Then, $\beta_{ir}(t)$ could be written as a product of two vectors, i.e., $\beta_{ir}(t) = \sum_k c_{k,ir} \boldsymbol{q}_{k,r}{}^T \boldsymbol{v}_r(t)$, where $\boldsymbol{v}_r(t)$ is the time-dependent vector which gives the form of the canonical models. For example, if the canonical model is in a linear form $f_{k,r}(t) = q_0 + q_1 t$, it could be seen as $f_{k,r} = \boldsymbol{q}_{k,r}{}^T \boldsymbol{v}_r(t)$, where the parameter vector $\boldsymbol{q}_{k,r} = [q_0, q_1]^T$ and the time-dependent vector $\boldsymbol{v}_r(t) = [1, t]^T$.

For the binary logistic model as shown in Equation 5.1, we can have:

$$\begin{aligned} Pr(y_{it} = 1|\boldsymbol{x}_{it}) &= \frac{\exp(\boldsymbol{\beta}_i^T \boldsymbol{x}_{it})}{1 + \exp(\boldsymbol{\beta}_i^T \boldsymbol{x}_{it})} = \frac{\exp(\boldsymbol{x}_{it}^T (\boldsymbol{Q}\boldsymbol{C}_i)^T \boldsymbol{V}(t))}{1 + \exp(\boldsymbol{x}_{it}^T (\boldsymbol{Q}\boldsymbol{C}_i)^T \boldsymbol{V}(t))} \\ Pr(y_{it} = 0|\boldsymbol{x}_{it}) &= \frac{1}{1 + \exp(\boldsymbol{\beta}_i^T \boldsymbol{x}_{it})} = \frac{1}{1 + \exp(\boldsymbol{x}_{it}^T (\boldsymbol{Q}\boldsymbol{C}_i)^T \boldsymbol{V}(t))} \end{aligned} \tag{5.3}$$

The log-likelihood function of the binary logistic model then can be written as:

$$l(\beta) = -\log(1+\exp(\boldsymbol{\beta}_i^T \boldsymbol{x}_{it})) + y_{it}(\boldsymbol{\beta}_i^T \boldsymbol{x}_{it}) = -\log(1+\exp(\boldsymbol{x}_{it}^T (\boldsymbol{Q}\boldsymbol{C}_i)^T \boldsymbol{V}(t))) + y_{it}(\boldsymbol{x}_{it}^T (\boldsymbol{Q}\boldsymbol{C}_i)^T \boldsymbol{V}(t)) \tag{5.4}$$

To formulate the log-likelihood function in the collaborative learning framework for parameter estimation, we further write up the K_r canonical models as a matrix: $\boldsymbol{Q}_r \equiv [\boldsymbol{q}_{1,r}, \boldsymbol{q}_{2,r}, ..., \boldsymbol{q}_{K_r,r}] \in \mathbb{R}^{R \times K_r}$. Then, we could rewrite $\beta_{ir}(t) = (\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_r(t)$. Given that each individual's preferences have R dimensions, and each dimension has a set of canonical models, an individual may have R different membership vectors towards R dimensions of his preferences. Now the individual i's preference vector in Equation 5.2 can be further written as:

$$\boldsymbol{\beta}_i = [\beta_{i1}(t), \beta_{i2}(t), ..., \beta_{iR}(t)]^T = [(\boldsymbol{Q}_1\boldsymbol{c}_{i1})^T\boldsymbol{v}_1(t), (\boldsymbol{Q}_2\boldsymbol{c}_{i2})^T\boldsymbol{v}_2(t), ..., (\boldsymbol{Q}_R\boldsymbol{c}_{iR})^T\boldsymbol{v}_R(t)]^T$$

We also need to formulate $\boldsymbol{\beta}_i^T\boldsymbol{x}_{it}$. To do this, we combine the membership vectors for all dimensions of the preferences of individual i into a diagonal matrix $\boldsymbol{C}_i = diag(\boldsymbol{c}_{i1}, \boldsymbol{c}_{i2}, ..., \boldsymbol{c}_{iR})$. We further combine the canonical models for each dimension into a big matrix $\boldsymbol{Q} = diag(\boldsymbol{Q}_1, \boldsymbol{Q}_2, ..., \boldsymbol{Q}_R)$ and combine the time-dependent vectors $\boldsymbol{v}_r(t)$ for the R forms of the canonical models into $\boldsymbol{V}(t) = [\boldsymbol{v}_1(t), \boldsymbol{v}_2(t), ..., \boldsymbol{v}_R(t)]^T$. Then $\boldsymbol{\beta}_i^T\boldsymbol{x}_{it}$ could be written as $\boldsymbol{\beta}_i^T\boldsymbol{x}_{it} = \boldsymbol{x}_{it}^T(\boldsymbol{Q}\boldsymbol{C}_i)^T\boldsymbol{V}(t)$. At time step t, the result of $(\boldsymbol{Q}\boldsymbol{C}_i)^T\boldsymbol{V}(t)$ is a $R \times 1$ vector $[\beta_{i1}(t), \beta_{i2}(t), ..., \beta_{iR}(t)]^T$.

To learn all the parameters in $\boldsymbol{C}$ and $\boldsymbol{Q}$, we pool all the data from all individuals together, which leads to the following formulation:

$$\begin{aligned} \min_{\mathbf{C_1},\mathbf{C_2},...,\mathbf{C_N},\mathbf{Q}} \quad & \sum_{i=1}^{N}\frac{1}{n_i}\sum_{t=1}^{n_i}\left\{\log(1+\exp(\boldsymbol{x}_{it}^T\cdot(\boldsymbol{Q}\boldsymbol{C}_i)^T\boldsymbol{V}(t))) - y_{it}(\boldsymbol{x}_{it}^T\cdot(\boldsymbol{Q}\boldsymbol{C}_i)^T\boldsymbol{V}(t))\right\}, \\ \text{s.t.} \quad & \boldsymbol{c}_{ir} \geq 0,\ \boldsymbol{c}_{ir}^T\mathbf{1} = 1 \quad i = 1, ..., N; r = 1, ..., R. \end{aligned} \tag{5.5}$$

Here, n_i is the number of measurements/observations from individual i. We name the model of the optimization framework in Equation 5.5 as the Logistic Collaborative Model with Time-varying Preferences (LCM-T). In LCM-T, while t changes, the attributes $\boldsymbol{x}_{it}$ changes accordingly, representing the specific attributes for the scenario at time step t. The time-dependent vector $V(t)$ also changes, leading to the changes of each dimension of individual i's preferences. It could be noticed that now in our LCM-T, the sequences of the choice scenarios play a role in determining the decision variables. This is different from the previous optimization works done with the collaborative learning framework in [113, 112, 111].

As stated in the previous section, we assume that all the preference dimensions for an individual could be described with cubic polynomial models, i.e., $\boldsymbol{v}_1(t) = \boldsymbol{v}_2(t) = ... = \boldsymbol{v}_R(t) = [1, t, t^2, t^3]^T = \boldsymbol{v}_t$. Thus, the formation of all the canonical models ($k = 1, 2, ..., K_r$) for any dimension β_r could be written as $f_{k,r}(t) = {\boldsymbol{q}_{k,r}}^T\boldsymbol{v}_t$, where the parameter vector

$\boldsymbol{q}_{k,r} = [q_{1,kr}, q_{2,kr}, q_{3,kr}, q_{4,kr}]^T$, i.e., $R = 4, \forall r \in \{1, 2, ..., R\}$.

5.2.3 Online Logistic Collaborative Model with Time-varying Parameters (OLCM-T)

With the optimization problem shown in Equation 5.5, we can estimate the canonical parameters matrix $\boldsymbol{Q}$ and the membership matrix $\boldsymbol{C}$ iteratively with a parameter estimation algorithm similar to the one in [112]. However, the computation process may take time. Thus, this wholesome update of all parameters of all individuals based on only a few observations from one individual is not efficient. To update an individual's preferences when a new observation is available, we separate the process of canonical model updating and membership vector updating of the Logistic Collaborative Model with Time-varying Parameters (LCM-T) into two stages — "offline updating" and "online updating". The proposed Online Logistic Collaborative Model with Time-varying Parameters (OLCM-T) we are focusing on in this paper is the algorithm of the online stage.

In the online updating stage, an individual's membership vectors will be updated given his own data, while the canonical models are not updated. The online updating could also be used to learn the membership vectors when a new user comes to the system.

Assume that for individual i, there are n_i observations $y_{it}, t = 1, 2, ..., n_i$ available. In online updating process, the optimization problem in Equation 5.5 becomes:

$$\begin{aligned} \min_{\boldsymbol{C}_i} \quad & \sum_{t=1}^{n_i} \left\{ \log(1 + \exp(\boldsymbol{x}_{it}^T \cdot (\boldsymbol{Q}\boldsymbol{C}_i)^T \boldsymbol{V}_t)) - y_{it}(\boldsymbol{x}_{it}^T \cdot (\boldsymbol{Q}\boldsymbol{C}_i)^T \boldsymbol{V}_t) \right\}, \\ \text{s.t.} \quad & \boldsymbol{c}_{ir} \geq 0, \ \boldsymbol{c}_{ir}{}^T \mathbf{1} = 1 \quad r = 1, ..., R. \end{aligned} \tag{5.6}$$

Here, $\boldsymbol{V}_t = [\boldsymbol{v}_t, \boldsymbol{v}_t, ..., \boldsymbol{v}_t]^T$, where $\boldsymbol{v}_t = [1, t, t^2, t^3]^T$. $\boldsymbol{Q} = diag(\boldsymbol{Q}_1, \boldsymbol{Q}_2, ..., \boldsymbol{Q}_R)$ is the diagonal matrix of the canonical models of each dimension, where $\boldsymbol{Q}_r = [\boldsymbol{q}_{1,r}, \boldsymbol{q}_{2,r}, ..., \boldsymbol{q}_{K_r,r}]$, $r = 1, 2, ..., R$ is the parameter matrix of K_r canonical models for preference dimension r. $\boldsymbol{x}_{it}$ and y_{it}, $t = 1, 2, ..., n_i$ are the attributes and choices in the t's scenario for individual i, which are also known. The decision variables are individual i's R membership vectors towards each dimension of the preferences. For each dimension r, his membership vector $\boldsymbol{c}_{ir} \geq 0$, and the $\sum_{k=1}^{K_r} c_{k,ir} = 1$. Since each dimension of the individual's preferences is a linear combination of the corresponding canonical models, each preference dimension is also time-dependent and change in a form of the cubic polynomial.

5.2.4 Relationship between Online Updating and Offline Updating

In the work of [194], an algorithm based on Random Utility Maximization theory utilizing a particle filtering method is used to learn and update an individual's preferences, such that tracking preference changes is possible. The proposed LCM-T can capture the preferences changes by estimating $\boldsymbol{Q}$ and $\boldsymbol{C}$ iteratively. However, compared with the particle filter method, the updating process of the proposed Logistic Collaborative Model with Time-varying Parameters (LCM-T) would be cumbersome, especially when we are interested in updating only several individual's preferences (assume that the whole population is large). In other words, the whole computation process of the optimization problem needs to be taken, even when only a few individuals have new data available, and their preferences are to be updated. Because of this, the real-time preference updating for each individual may be inefficient to implement with the proposed Logistic Collaborative Model with Time-varying Parameters (LCM-T).

The two-stage updating structure of the proposed LCM-T could solve the problem: by having an online and an offline updating stage, we could apply the proposed online updating method to the new data obtained by an individual and only update his membership vectors with much shorter computation time. Since the canonical models are vital in the preference learning process but they are not updated in the online updating stage, the offline updating stage needs to be applied once after a time, during which both canonical models and individuals' membership vectors are updated using much more data in history. In other words, the online updating stage is taken more frequently than the offline updating stage.

Since the online and offline updating stages are taken iteratively in the LCM-T model, a linkage between the two stages should be built, and several questions need to be answered. For example, one question is related to the updating rule of the membership vectors. As the membership vectors are updated both in online and offline stages, the updates from the two stages may be different from each other. How to update the membership vectors in the iterative process such that little information is lost may worth exploration. One possible method is to add friction factors in the offline-updating process, such that the information from both online and offline stages could be maintained in the membership vectors. Another question could be related to the time interval between two offline updates. If offline updating is taken rarely, the model may have low accuracy, while if the offline updating is taken very frequently, the model may have low efficiency. This may need more explorations with real-world data, and the conclusions may differ significantly when having different applications. These questions are beyond the scope of the current paper, where we primarily focus on the online updating process when the canonical models are assumed known from the previous offline updating step. In the following sections, we are only considering the online updating stage.

5.3 Parameter Estimation Algorithm for Online Logistic Collaborative Model with Time-varying Parameters (OLCM-T)

In the online updating stage, an individual's membership vectors are to be updated. Based on Equation 5.6, given his new data $y_{i1}, ..., y_{it}$, the objective function for individual i could be written as:

$$\begin{aligned} \min_{\boldsymbol{c}_{ir}, r=1,2,...,R} \quad & \sum_{t=1}^{n_i} \Big\{ \log\Big(1 + \exp\big(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t\big)\Big) - y_{it}\big(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t\big)\Big\}, \\ \text{s.t.} \quad & \boldsymbol{c}_{ir} \geq 0, \ \boldsymbol{c}_{ir}{}^T \mathbf{1} = 1 \quad r = 1, ..., R. \end{aligned} \tag{5.7}$$

In Equation 5.7, the decision variables are the membership vectors of individual i. Given $\boldsymbol{Q}_1, \boldsymbol{Q}_2, ..., \boldsymbol{Q}_R$ and $\boldsymbol{v}_t$, the Lagrangian function of the formulation shown in Equation 5.7 could be written as:

$$L = \sum_{t=1}^{n_i} \Big\{ \log\Big(1 + \exp\big(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t\big)\Big) - y_{it}\big(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t\big)\Big\} + \sum_{r=1}^{R} \mu_r(\boldsymbol{c}_{\boldsymbol{ir}}^T \mathbf{1} - 1) \tag{5.8}$$

In equation 5.8, we introduce the Lagrangian multiplier μ_r for constraint $\boldsymbol{c}_{\boldsymbol{ir}}^T \mathbf{1} = 1$. To get the optimal $\boldsymbol{c}_{ir}, r = 1, 2, ..., R$, we could derive the gradient of the objective function regarding $\boldsymbol{c}_{ir}$:

$$\frac{\partial L}{\partial \boldsymbol{c}_{ir}} = \sum_{t=1}^{n_i} \left\{ \left(\frac{\exp(\sum_{r=1}^{R} x_{r,it} (\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t)}{1 + \exp(\sum_{r=1}^{R} x_{r,it} (\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t)} - y_{it} \right) \times x_{r,it} \boldsymbol{Q}_{\boldsymbol{r}}^T \boldsymbol{v}_t \right\} + \mu_r \mathbf{1} = 0 \tag{5.9}$$

According to Karush–Kuhn–Tucker conditions, we also have $\frac{\partial L}{\partial c_{k,ir}} c_{k,ir} = 0$ ($k = 1, 2, ..., K_r; r = 1, 2, ..., R$), which could lead to the following equation:

$$\frac{\partial L}{\partial c_{k,ir}} c_{k,ir} = \sum_{t=1}^{n_i} \left\{ \left(\frac{\exp(\sum_{r=1}^{R} x_{r,it} (\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t)}{1 + \exp(\sum_{r=1}^{R} x_{r,it} (\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t)} - y_{it} \right) \times x_{r,it} (\boldsymbol{Q}_r{}^T \boldsymbol{v}_t)_k c_{k,ir} \right\} + \mu_r c_{k,ir} = 0 \tag{5.10}$$

Given $\boldsymbol{c}_{ir} \mathbf{1} = 1$, i.e., $\sum_{k=1}^{K_r} c_{k,ir} = 1$, Equation 5.10 could further be modified as:

$$\sum_{t=1}^{n_i} \left\{ \left(\frac{\exp(\sum_{r=1}^{R} x_{r,it} (\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t)}{1 + \exp(\sum_{r=1}^{R} x_{r,it} (\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t)} - y_{it} \right) \times x_{r,it} (\boldsymbol{Q}_r{}^T \boldsymbol{v}_t)^T \boldsymbol{c}_{ir} \right\} + \mu_r = 0 \tag{5.11}$$

With Equation 5.11, we could write μ_r in the following way:

$$\mu_r = \sum_{t=1}^{n_i} y_{it} \times x_{r,it} (\boldsymbol{Q}_r{}^T \boldsymbol{v}_t)^T \boldsymbol{c}_{ir} - \sum_{t=1}^{n_i} \frac{\exp(\sum_{r=1}^{R} x_{r,it} (\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t)}{1 + \exp(\sum_{r=1}^{R} x_{r,it} (\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t)} \times x_{r,it} (\boldsymbol{Q}_r{}^T \boldsymbol{v}_t)^T \boldsymbol{c}_{ir} \tag{5.12}$$

Replace μ_r in Equation 5.10 with Equation 5.12:

$$\begin{aligned} & c_{k,ir} \times \left\{ \sum_{t=1}^{n_i} \left(\frac{\exp(\sum_{r=1}^{R} x_{r,it} (\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t)}{1 + \exp(\sum_{r=1}^{R} x_{r,it} (\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t)} - y_{it} \right) \times x_{r,it} (\boldsymbol{Q}_r{}^T v(t))_k \right. \\ & \left. + \sum_{t=1}^{n_i} \left(y_{it} - \frac{\exp(\sum_{r=1}^{R} x_{r,it} (\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t)}{1 + \exp(\sum_{r=1}^{R} x_{r,it} (\boldsymbol{Q}_r \boldsymbol{c}_{ir})^T \boldsymbol{v}_t)} \right) \times x_{r,it} (\boldsymbol{Q}_r{}^T \boldsymbol{v}_t)^T \boldsymbol{c}_{ir} \right\} = 0 \end{aligned} \tag{5.13}$$

Since $c_{k,ir}$ should be non-negative, we define $\delta_+(x) \equiv max(x, 0)$ and $\delta_-(x) \equiv min(x, 0)$, with which Equation 5.13 could be separated into a positive part and a negative part:

$$
\begin{aligned}
c_{k,ir} \times \Bigg\{ \sum_{t=1}^{n_i} \Bigg[&\delta_+ \Bigg(\frac{\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir})^T\boldsymbol{v}_t)}{1+\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir})^T\boldsymbol{v}_t)} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)_k \Bigg) - \delta_- \Bigg(y_{it} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)_k \Bigg) \\
&+\delta_+ \Bigg(y_{it} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)^T\boldsymbol{c}_{ir} \Bigg) - \delta_- \Bigg(\frac{\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir})^T\boldsymbol{v}_t)}{1+\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir})^T\boldsymbol{v}_t)} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)^T\boldsymbol{c}_{ir} \Bigg) \Bigg] \Bigg\} \\
-c_{k,ir} \times \Bigg\{ \sum_{t=1}^{n_i} \Bigg[&-\delta_- \Bigg(\frac{\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir})^T\boldsymbol{v}_t)}{1+\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir})^T\boldsymbol{v}_t)} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)_k \Bigg) + \delta_+ \Bigg(y_{it} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)_k \Bigg) \\
&-\delta_- \Bigg(y_{it} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)^T\boldsymbol{c}_{ir} \Bigg) + \delta_+ \Bigg(\frac{\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir})^T\boldsymbol{v}_t)}{1+\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir})^T\boldsymbol{v}_t)} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)^T\boldsymbol{c}_{ir} \Bigg) \Bigg] \Bigg\} = 0
\end{aligned}
\tag{5.14}
$$

With Equation 5.14, we could derive an iteratively updating rule for $\boldsymbol{c}_{ir}$ similarly as [111, 113, 112]:

$$
\begin{aligned}
c_{k,ir}^{m+1} = c_{k,ir}^{m} \times \Bigg\{ \sum_{t=1}^{n_i} \Bigg[&-\delta_- \Bigg(\frac{\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir}^m)^T\boldsymbol{v}_t)}{1+\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir}^m)^T\boldsymbol{v}_t)} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)_k \Bigg) + \delta_+ \Bigg(y_{it} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)_k \Bigg) \\
&-\delta_- \Bigg(y_{it} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)^T\boldsymbol{c}_{ir}^m \Bigg) + \delta_+ \Bigg(\frac{\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir}^m)^T\boldsymbol{v}_t)}{1+\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir}^m)^T\boldsymbol{v}_t)} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)^T\boldsymbol{c}_{ir}^m \Bigg) \Bigg] \Bigg\} \\
\Bigg/ \Bigg\{ \sum_{t=1}^{n_i} \Bigg[&\delta_+ \Bigg(\frac{\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir}^m)^T\boldsymbol{v}_t)}{1+\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir}^m)^T\boldsymbol{v}_t)} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)_k \Bigg) - \delta_- \Bigg(y_{it} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)_k \Bigg) \\
&+\delta_+ \Bigg(y_{it} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)^T\boldsymbol{c}_{ir}^m \Bigg) - \delta_- \Bigg(\frac{\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir}^m)^T\boldsymbol{v}_t)}{1+\exp(\sum_{r=1}^{R} x_{r,it}(\boldsymbol{Q}_r\boldsymbol{c}_{ir}^m)^T\boldsymbol{v}_t)} \times x_{r,it}(\boldsymbol{Q}_r{}^T\boldsymbol{v}_t)^T\boldsymbol{c}_{ir}^m \Bigg) \Bigg] \Bigg\}
\end{aligned}
\tag{5.15}
$$

In Equation 5.15, the superscript m refers to the order of iteration. By introducing δ-functions, we ensure that the numerator and the denominator are both non-negative. Therefore, given any positive initial membership vectors $\boldsymbol{c}_{ir}, r = 1, 2, ..., R$, the non-negativity requirement of the membership vectors is guaranteed.

In summary, the membership vectors could be learned and updated iteratively with the following steps:

1. Input: $\boldsymbol{Q}$, $\boldsymbol{v}_t = [1, t, t^2, t^3]^T$, $x_{r,it}$, y_{it}, initial value $\boldsymbol{c}_{ir}^0$, for all $r = 1, 2, ..., R$; $t = 1, 2, ..., n_i$;

2. For $m = 1, 2, ...$, iteratively update each dimension of each membership vector $c_{k,ir}^{m+1}$

with Equation 5.15, given $\boldsymbol{c}_{ir}^{m}$ calculated in the previous iteration step;

3. Give a pre-determined criteria ϵ. When $\gamma^m = \sum_{r=1}^{R} \|\boldsymbol{c}_{ir}^{m+1} - \boldsymbol{c}_{ir}^{m}\|_2^2 \leq \epsilon$, stop the iteration.

5.4 Simulations

The purpose of the simulation is to test the performance of the proposed model using online-updating strategy. Given the online-updating strategy is a real-time updating method, the model is expected to estimate and update an individual's preferences at each time step when new data points are available. Thus, we will (1) generate the true values of the preferences for each individual for several successive time steps, (2) generate attributes for the choice scenario at each time step, and (3) generate the responses to the scenario based on his the true preferences at each time step. Notice that when generating the true preferences for each individual at step (1), we need to generate the canonical models (polynomial models in the current study) and membership vectors for all individuals, such that we can obtain each individual's preference changing model with the method we propose in Section 2, and estimate his preferences at each time step.

The performance of the model is evaluated with two metrics, Average Absolute Error, and Prediction Accuracy.

Average Absolute Error measures the difference between learned coefficients and the true ones (Equation 5.16):

$$\text{Average Absolute Error} = \frac{1}{N}\sum_{i=1}^{N} |\beta_i - \hat{\beta}_i|. \tag{5.16}$$

Prediction Accuracy measures how the model performs in behavioral predictions, as shown in Equation 5.17:

$$\text{Prediction Accuracy} = \frac{N_{\text{Number of predictions that are consistent with the actual choices made}}}{N_{\text{Number of predictions made in total}}} \tag{5.17}$$

The performance of the proposed Online Logistic Collaborative Model with Time-varying Parameters (OLCM-T) with time-varying preferences is compared with several benchmark methods including (1) the independent logistic regression model (ILM) that learns the regression coefficients of each individual solely based on his own data; (2) the mixed-effect logistic regression model (logistic MEM) that considers the coefficients of individuals are extracted from a certain distribution; (3) the one-size-fits-all logistic regression model (LR) that treats all individuals homogeneously, combines all individuals' data and estimates one set of preferences; (4) the original Logistic Collaborative Learning (LCM) model, where individual preferences are constant values β_r rather than time-varying variables $\beta_r(t)$. In LCM, the canonical models represent different utility models in a format of $U = \sum_r \beta_r x_r$, and different canonical models have different β_r.

The simulations in this section test each model's performance when available data points increase over time. In other words, at the beginning of the updating period, the data points obtained from each individual may be limited (i.e., the so-called sparse sampling condition), and as time goes by, the sampling condition may become denser when more and more available data points could be collected. This happens commonly in various realistic circumstances, and we will show that the proposed method has advantages when dealing with it.

5.4.1 Data generation

Two coefficients are varying in the simulation, such that we could see the performance of the proposed algorithm and make comparisons: (1) the number of attributes in the utility model R, i.e., the number of factors in a choice scenario, or the number of the preference dimensions; (2) the number of canonical models for each preference dimension K (here in the simulation we are assuming that different preference dimension has the same number of canonical models, though the models are different). With R and K assigned, we could set the true preferences for the group of individuals with preference heterogeneity.

5.4.1.1 Generate suitable formulation for different canonical models

A given number of canonical models are generated by randomly setting the parameters of the cubic polynomial models. For each cubic polynomial model $q_0 + q_1 t + q_2 t^2 + q_3 t^2$, we need to set 4 parameters. Since the canonical models represent different preference changing patterns, they need to be different enough from each other. To guarantee this, for each preference dimension, we can let the canonical models be polynomial models of different degrees. For example, we may let one canonical model be a cubic polynomial, and other canonical models be a quadratic model, a linear model, and a constant model. We can also let different canonical models have opposite signs. For example, in canonical model $q_0^1 + q_1^1 t + q_2^1 t^2 + q_3^1 t^3$ and canonical model $q_0^2 + q_1^2 t + q_2^2 t^2 + q_3^2 t^3$, q_3^1 and q_3^2 could have opposite signs such that the two canonical models are significantly different from each other.

5.4.1.2 Generate appropriate parameters for canonical models

We obtain the magnitudes/ranges of the 4 parameters in the cubic polynomial model by running regressions with the learned preferences we obtained in our previous study [194]. As we stated in Section 4.3.2, we have individual preferences for 826 respondents in 13 binary choice scenarios. We run regressions for each individual, and obtain the magnitudes for q_0, q_1, q_2, and q_3 in the cubic polynomial model. Then, the parameters of each canonical model are selected randomly from corresponding ranges:

$$q_0 \in [-0.25, 0.25];\ q_1 \in [-0.05, 0.05];\ q_2 \in [-0.008, 0.008];\ q_3 \in [-0.0005, 0.0005]$$

Since we would like to guarantee that the canonical models are different from each other, some parameters are set to be 0 intentionally. For example, to obtain a quadratic polynomial model, we will let $q_3 = 0$ and $q_2 \neq 0$.

5.4.1.3 Generate membership vectors

With canonical models, each individual's membership vectors need to be decided such that his preferences $\boldsymbol{\beta}_i$ could be obtained via $\boldsymbol{\beta}_i = (\boldsymbol{Q}\boldsymbol{C}_i)^T \boldsymbol{V}_t$. Given the requirements that $\boldsymbol{c}_{ir}\mathbf{1} = 1$ and $c_{k,ir} \geq 0$, we use Dirichlet distribution to model the membership vectors (the commend is *rdirichlet* in R). In the simulation, we let each individual has his dominant preference changing model for each dimension of his preferences. This is achieved by setting one large value (e.g., 10) in the p-length vector $\boldsymbol{\alpha}$ and let other values of the vector be small (e.g., 1). With the given number of canonical models K, we equally split the total population into K sub-groups, and individuals in one sub-group will have the same "dominant model", which refers to the canonical model to which the preference changing patterns of these individuals will resemble more than others.

For example, assume that we are generating true preferences for a condition where the individual preferences are 3-dimensional and there are 2 canonical models for each dimension. The whole population is split into 2 sub-groups, each with $120/2 = 60$ individuals. Each individual has 3 membership vectors corresponding to the 3-dimensional preferences. Each membership vector has 2 values in response to the 2 canonical models. For the first sub-group, we may let the first canonical models be the "dominant models" for all preference dimensions. The corresponding value for the "dominant model" in each membership vector will be significantly larger than the other, e.g., $(0.9, 0.1)$. Meanwhile, for the second sub-group, we need to set the second canonical models be the "dominant models" for all preference dimensions. An example of the membership vector for one preference dimension of an individual in the second sub-group could be $(0.1, 0.9)$.

5.4.1.4 Generate attributes in choice scenarios

We then generate choice scenario attributes $\boldsymbol{x}_{it}$ for each individual i at each time step. With the generated attributes at each time step t, an individual's choice in the scenario can be predicted using binary logit model:

$$p_1 = Pr(y_{it}|\boldsymbol{x}_{it}) = \frac{\exp(\boldsymbol{x}_{it}^T\boldsymbol{\beta}_{it})}{1+\exp(\boldsymbol{x}_{it}^T\boldsymbol{\beta}_{it})} = \begin{cases} > 0.5 & y_{it} = 1 \\ \leq 0.5 & y_{it} = 0 \end{cases} \tag{5.18}$$

We randomly select the values of the attributes from a uniform distribution $U(0, 100)$, which is consistent with the magnitudes of the attributes in the online experiment in [194]. The consistency in the magnitudes of the attributes and the preferences helps ensure the reliability of the simulations in the current study. Given the number of attributes R, we first randomly select $R-1$ values from a uniform distribution $U(0, 100)$ for each individual at each time step. We then randomly select p_1 from a uniform distribution $U(0, 1)$. With p_1, we could calculate the last attribute required using the discrete choice-making model such that his probability to select the promoted choice is the pre-determined value p_1. By generating attributes in this way, we avoid the system from aborting due to large $\exp(\boldsymbol{x}_{it}^T \boldsymbol{\beta}_{it})$, and make it possible for us to control the balance of the responses (i.e., the percentages of acceptance and rejections in responses).

5.4.1.5 Levels of noise in responses

We also test the performance of the model given different levels of noise in responses, i.e., 10% (0.1), 20% (0.2), 30% (0.3), which represent the percentages of the incorrect responses for each individual in all his choices. To achieve this, a random number between 0 and 1 is generated after generating a scenario and the corresponding true response from the true preferences. If the number is smaller than the noise level (e.g., 0.1), we turn the response to the opposite (from acceptance to rejection, or from rejection to acceptance), making the response in this scenario an incorrect response.

5.4.1.6 simulation of the online-updating process

To mimic the online updating process over time, we assume that at each time point, one data point could be obtained from each individual. For a given time step, the data that could be used to learn the preferences of the individual is the set of the data points obtained from all previous time steps and the new data point obtained at the current time step. Another 10 data points are also generated at each time step for testing. In other words, at time step t, we generate $1 + 10 = 11$ data points with the performances of the time step t. Among all the 11 data points, one is the new observation at this time step, and the other 10 data points consist of the testing set.

In our simulation, we set the total number of individuals $N = 120$, set the number of preference parameters in utility function $R = 4$ (i.e., the individual preferences have four dimensions), and set the number of canonical models for each preference dimension $K = 2$. With this basic setting, each individual has four membership vectors towards the four dimensions of his preferences, and each dimension of his preferences could be described by the two canonical models of that preference dimension and his membership vector towards that dimension. Tests regarding the performances of the model with different numbers of canonical models and different numbers of dimensions (i.e., the number of variables in the utility function) are also conducted accordingly.

We present the results of the prediction accuracy, computation time, and the Average Absolute Error of an estimate for each model (Online Logistic Collaborative Model with Time-Varying preferences, Independent Logit Model using data from one individual, Logistic Regression with data from all individuals, Mixed Effect Model, and Logistic Collaborative Model with constant preference parameters) in online updating over time, with the OLCM-T (Online Logistic Collaborative Model with Time-Varying preferences) having four parameters in the utility model and two canonical models in each preference dimension. Notice that though the scenarios and responses are the same when testing all the five models, the division of training sets and testing sets are only used in simulations of OLCM-T model. We then show the differences in prediction accuracy rates when changing the number of canonical models in each preference dimension, and when changing the variable numbers for our proposed OLCM-T model.

5.4.2 Simulation results

The results of the prediction accuracy of each model at each time step are presented in Figure 5.2. In the experiment, at time $t = 1$, each individual has 1 data point available for estimation. As time goes from $t = 1$ to $t = 15$, the data points collected from each individual increase from 1 to 15. In general, the prediction accuracy of all the five models increases over time when more data are available in preference learning and updating process, among which the proposed OLCM-T (Online Logistic Collaborative Model with Time-Varying preferences) considering time-varying preferences has better performances in prediction accuracy comparing with all other models. Since the best number of canonical models of LCM (Logistic Collaborative Model with constant preference parameters) is not known in the simulation, cross-validation is applied to identify optimal K before evaluating the performance of LCM (Logistic Collaborative Model with constant preference parameters). The average absolute error of the estimates obtained by each model is presented in

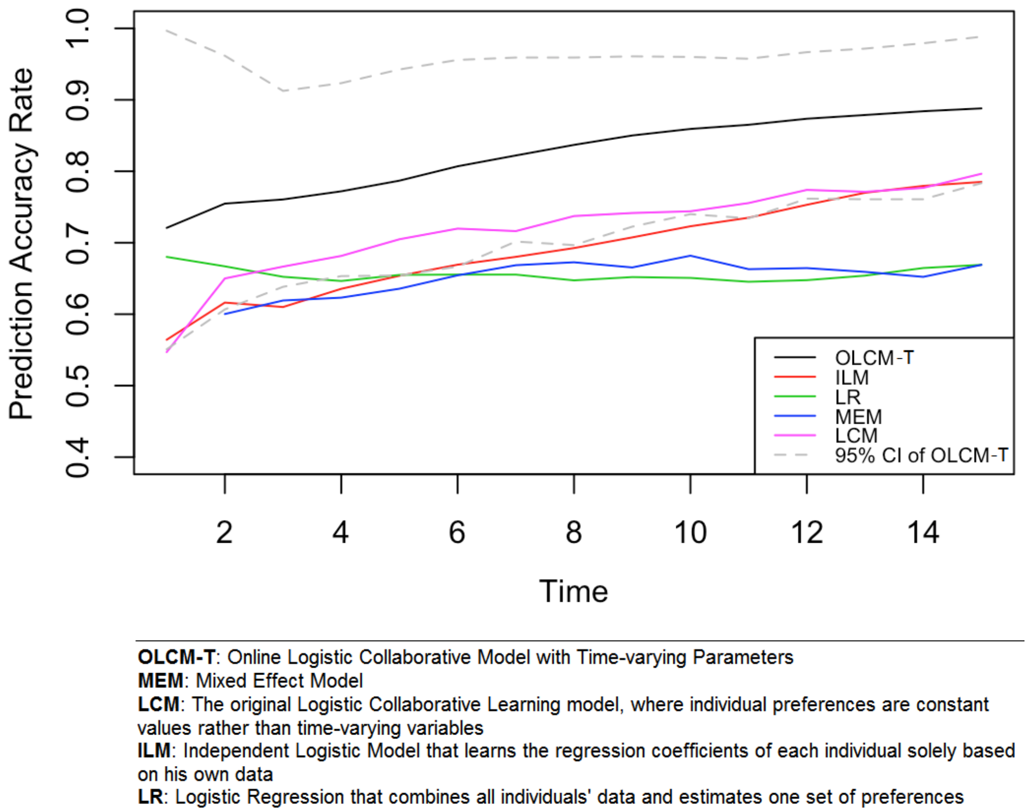


Figure 5.2: The prediction accuracy of each model over time in online updating process.

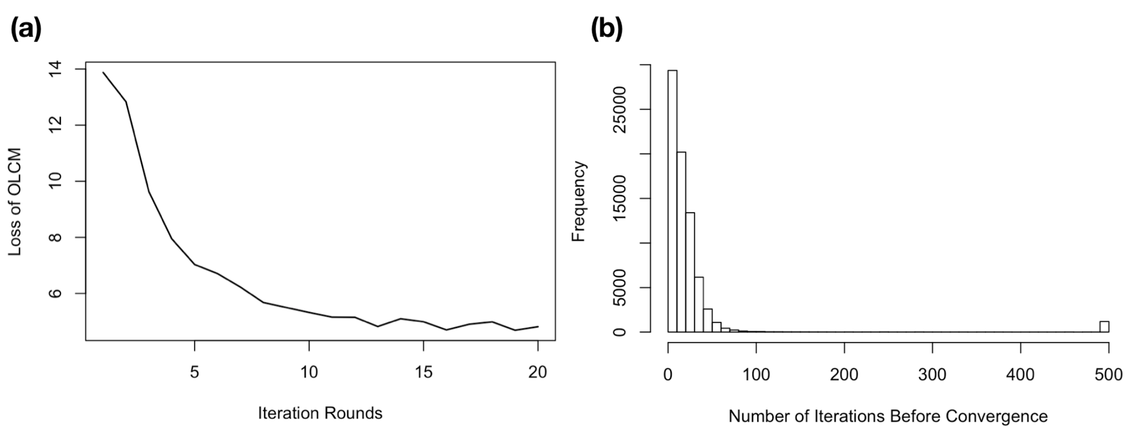


Figure 5.3: Convergence performance of the computational algorithm in the proposed Online Logistic Collaborative Model (OLCM-T)

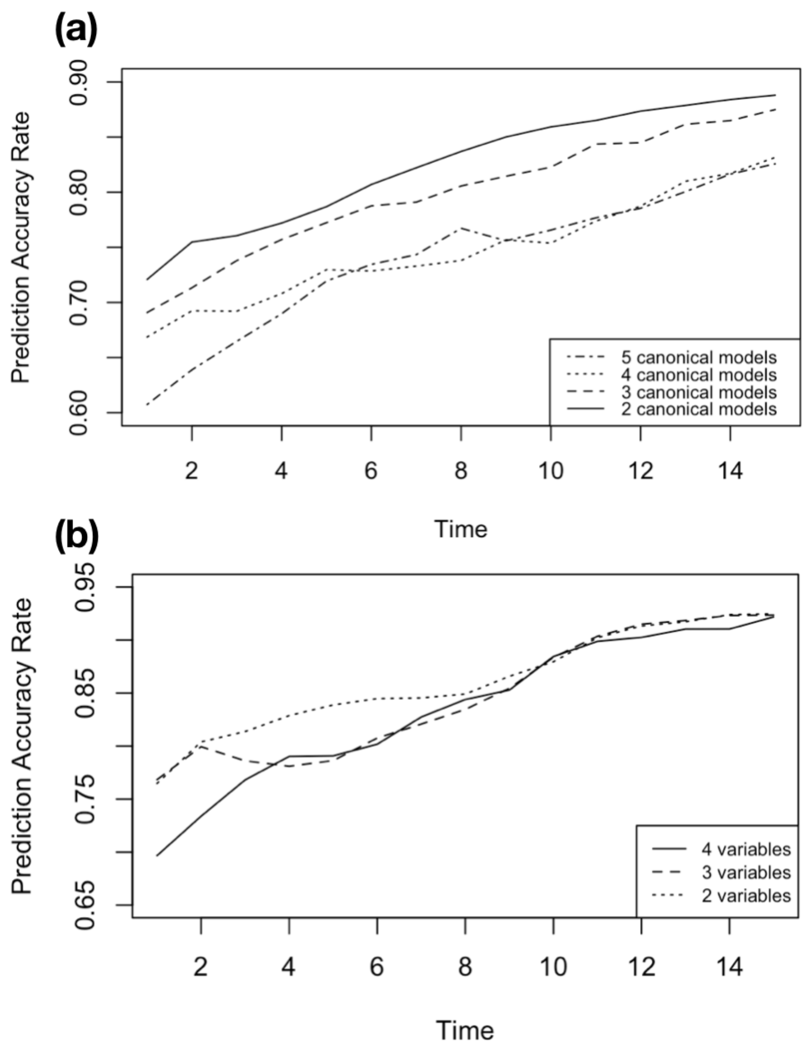


Figure 5.4: The prediction accuracy of (a) Online Logistic Collaborative Model with Time-Varying preferences (OLCM-T) with different number of canonical models, and (b) OLCM-T with different number of dimensions.

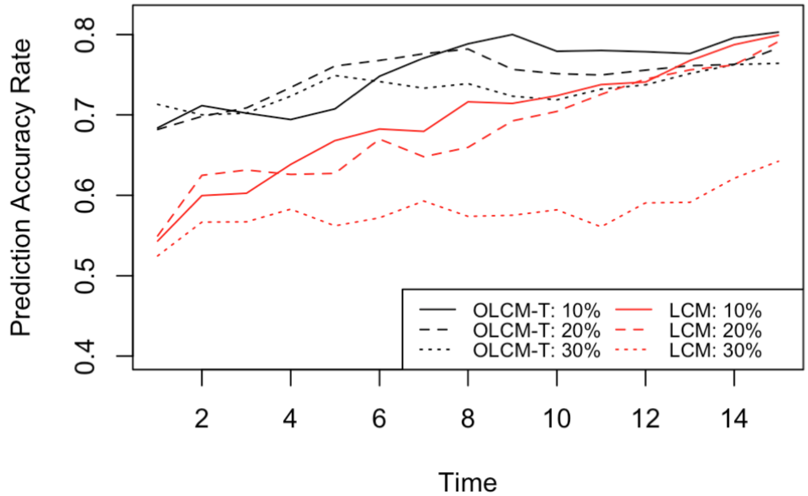


Figure 5.5: The prediction accuracy of Online Logistic Collaborative Model with Time-Varying preferences (OLCM-T) and original Logistic Collaborative Model (LCM) when having different level of noise in response.

Table 5.1. The proposed OLCM-T (Online Logistic Collaborative Model with Time-Varying preferences) has the lowest absolute error for each preference parameter when the available data is very limited.

The loss function of our model could be the sum of the minus log-likelihood of the logistic function. Figure 5.3a shows the average loss after each iteration of updating with the updating rule shown in Equation 5.15, and Figure 5.3b presents the histogram of the number of iterations before reaching a pre-specified convergence criterion, i.e., $|Loss(t) - Loss(t-1)| < 0.1$. Besides the criterion, we also set the maximum number of iterations to be 500. It could be seen that more than two-thirds of the total simulations converge within 20 iterations, and only about 1.5% of the simulations do not converge within 500 iterations.

The computation time for the proposed OLCM-T (Online Logistic Collaborative Model with Time-Varying preferences) is significantly longer (7788.792s) than the LCM (Logistic Collaborative Model with constant preference parameters) method (713.321s) where the preference changing is not considered. This could be reasonable because each individual now has multiple membership vectors (consistent with the number of preference dimensions),

Table 5.1: Results Model Comparison

Model	OLCM-T	LCM	ILM	LR	MEM
Average Computation Time (s)	7788.792	713.321	17.123	0.644	835.062
Average Absolute Error	0.101	0.945	5.175	1.100	0.445

such that more parameters in membership vectors need to be estimated. Because of this, the current method may require more iterations before convergence, which may significantly extend the computation time. Notice that the computation time counted here in Table 2 is the total computation time for 15 time steps and 120 individuals. As the proposed online algorithm OLCM-T is designed to be applied to each individual at each time step, the individual updating time at each time step would be much shorter (about 4.33 seconds).

We further test the performance of the proposed OLCM-T when the number of dimensions and the number of canonical models change. From Figure 5.4a we could see that when the number of canonical models increases, the performance of the model decreases. This is possibly because that when the number of canonical models increases, the number of parameters that need to be estimated also increases. Given that the available observations are the same, more parameters to be estimated may lead to lower accuracy in estimation. Meanwhile, the computation time does not show significant differences in the process. In Figure 5.4b we could see that for the first few time steps, the prediction accuracy would be higher when there are fewer numbers of dimensions (fewer variables in the model). As more data are obtained in the learning process, the performances would become similar to models with different numbers of dimensions. Figure 5.5 also shows that in general, the proposed OLCM-T has higher prediction accuracy compared with LCM when the responses fed to the model have some noises.

5.5 Real-world Case Study

To see the performance of OLCM-T with real-world data, we apply our model to the dataset we collected in an online experiment we conducted in [194]. Similar to the simulation section, in the real-world case study, we again compare the performance of the proposed OLCM-T with independent logit model with data from one individual (ILM), traditional logit model with data from all individuals (LM), mixed-effect model (MEM), and the original logistic collaborative model where the parameters are constant rather than time-varying (LCM). Moreover, we also compare the performance of OLCM-T with that of the latent class choice model (LCCM) and the particle filter method (PF) we used in [194].

5.5.1 Dataset of commuting departure time choices

The dataset includes responses from 826 individuals recruited from the AMT (Amazon Mechanical Turk) platform.

In the experiment, we randomly assign each participant into one of 9 experimental groups, each with a specific hypothetical background setting on his original commuting departure time, arrival time, typical commuting time, and the level of the flexibility in his working starting time. Then the participant is required to make binary choices on commuting departure time in 13 scenarios. In each scenario, he could select to depart at an earlier or later time point suggested in the experiment or to stay unchanged and depart at his original departure time. Information provided in each alternative includes departure time, arrival time, total commuting time, and possible reward points that are used to encourage the individual to accept the suggested alternative. Thus, we have four variables in the utility model: changes in arrival time (including both scheduled delay early SDE and scheduled delay late SDL), the minutes of travel time savings if the suggested departure time is accepted TTS and the reward points RP.

$$U = \beta_{SDE} SDE + \beta_{SDL} SDL + \beta_{TTS} TTS + \beta_{RP} RP \tag{5.19}$$

For each background setting of an experimental group, there are corresponding levels for the attributes of a proposed alternative in a scenario. Specifically, there are three levels for SDE and SDE (10 min, 25 min, and 60 min), two levels for TTS (10% for slight congestion, and 60 for severe congestion), and RP is no larger than 100. The experimental groups guarantee that there is preference heterogeneity in the population.

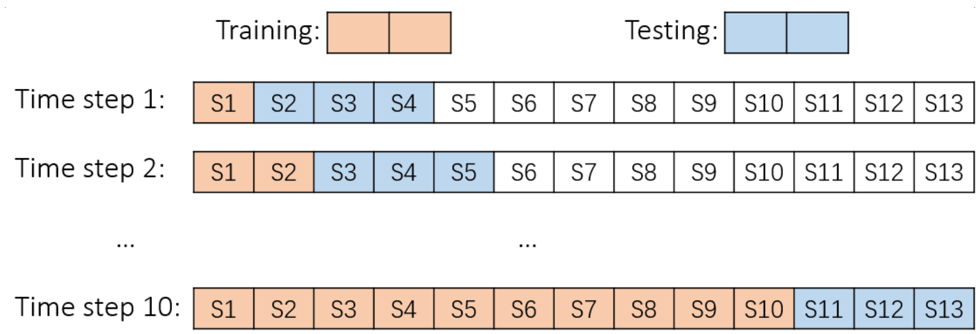


Figure 5.6: The data points used for model training and testing with real-world dataset at each time step.

The first two scenarios are the same for all respondents in the experiment. Start from the third scenario, the proposed alternative presented to a participant's in the next time step will be decided by the participant's choice at the current time step and his preferences learned by the individual preference learning algorithm proposed in [194]. With the updated preferences at each time step, the background algorithm further prepares a proposed alternative for the following scenario that is likely to be accepted by the participant. Thus, the amount of RP needed is calculated according to RUM such that the proposed alternative is attractive enough for the participant. If the required RP exceeds 100 points, the algorithm will change the alternative by adjusting the levels of other attributes.

5.5.2 Application procedures of different models

Similar to the simulation section, we assume that at each time point, one data point could be obtained from each individual. At each time step, the new data points obtained at the current time step, together with all the data points obtained in previous steps, could be used to learn the individual's preferences. Also, at each time step, the following three scenarios and choices are used as the test dataset to test the prediction accuracy given the preferences learned at the current time step (as shown in Figure 5.6).

For the independent logit model (ILM), we run logit regressions for each individual with his own data points at each time step. Since there are four parameters to be estimated, we could not obtain estimates from time step 1 to time step 3 with ILM. For the traditional logit model (LM), we run logit regression with data points from all individuals at each time step, and make predictions assuming that all the individuals have the same preferences. Since the scenarios at the first two time steps are the same for all respondents, we can only obtain estimates starting from the third time step. Similar problems also exist in the mixed-effect model (MEM) and the original logistic collaborative model (LCM).

For the particle filter method used in [194], we simply use the percentages of the acceptance at each time step in the dataset as the prediction accuracy. This is because that the preference learning algorithm utilizing particle filter approach itself is embedded in the experiment and each individual's preferences are updated at each time step, and the proposed alternative presented at each time step (starting from the third time step) to each respondent is expected to be accepted by the respondent. Thus, we are able to obtain the prediction accuracy of the particle filter algorithm proposed in [194] by calculating the acceptance ratio at each time step in the experiment.

For the proposed OLCM-T, we first decide the suitable degree of the polynomial models applied to this dataset using cross-validation. The results shown in Figure 5.7 suggest that the best model suitable for our dataset is the quadratic polynomial model. We then explore how many canonical models are needed for each dimension of the preferences. Similar to the conclusion we obtained from simulations (Figure 5.4a), the results also show that having two canonical models may lead to higher performance than having more canonical models. Thus in this real-world case study, we use two canonical modes for each preference dimension for OLCM-T. Cross-validation is also used to decide the number of canonical models for each preference dimension for the original LCM, and we let it be 5.

Since we are testing the performance of the online-updating strategy for the proposed LCM-T, at each time step, we will fix the canonical models and only update each individual's membership vector. Thus, we randomly select 80% of the 826 respondents (661 individuals) as a subset to learn the canonical models for this population, assuming that the canonical models are also applicable for the rest 20% respondents (165 individuals). Given the canonical models, the dataset of those 165 individuals is used to test the algorithm of OLCM-T (online collaborativemodel with time-varying preferences) presented in

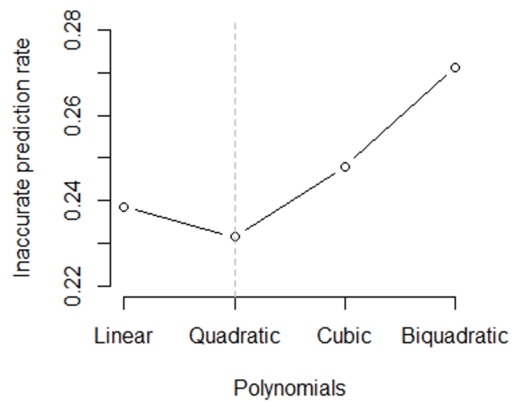
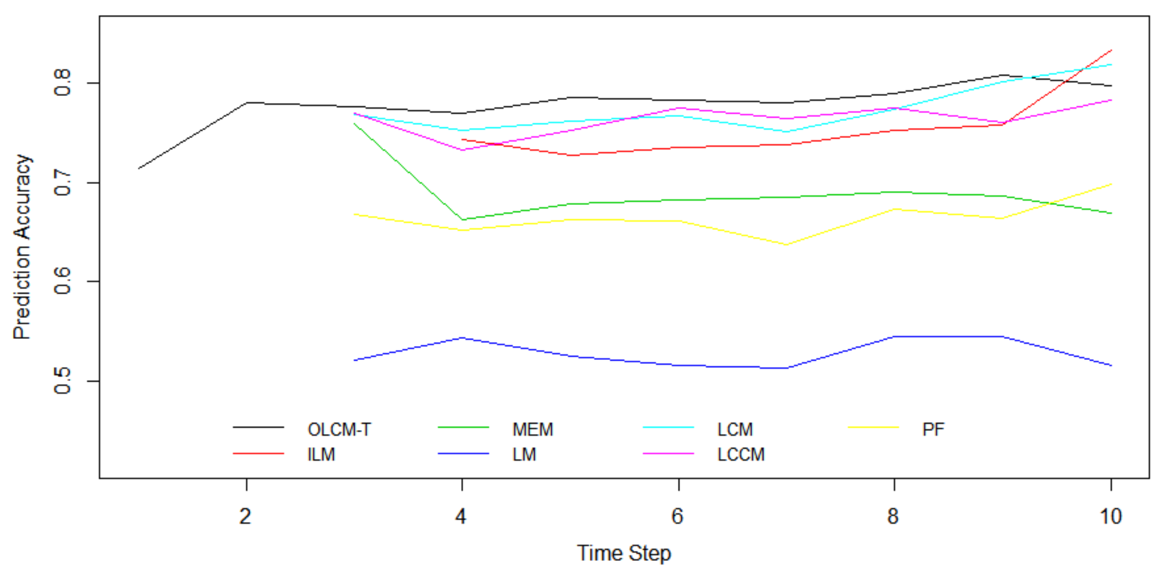


Figure 5.7: Inaccurate prediction rates with different degrees of polynomial models.



OLCM-T: Online Logistic Collaborative Model with Time-varying Parameters
MEM: Mixed Effect Model
LCM: The original Logistic Collaborative Learning model, where individual preferences are constant values rather than time-varying variables
ILM: Independent Logistic Model that learns the regression coefficients of each individual solely based on his own data
LM: Logit Model that combines all individuals' data and estimates one set of preferences
PF: Particle Filter, where the prediction accuracy at each time step is obtained in the experiment
LCCM: Latent Class Choice Model

Figure 5.8: The prediction accuracy of different models.

Section 5.2.2.2. As the 661 individuals in the canonical learning group are randomly selected from the population, we take independent random selection for 10 times.

5.5.3 Model application results

The performance of OLCM-T (online collaborative model with time-varying preferences) shown in Figure 5.8 is the average of the 10 trials. From Figure 5.8 we could see that the proposed OLCM-T has higher prediction accuracy than other models from the first time step until the 9th time step, at which the prediction accuracy of LCM exceeds that of OLCM-T (online collaborative model with time-varying preferences). This might because while the canonical models of OLCM-T are fixed in the whole process, the original LCM updates both canonical models and membership vectors at each time step, which may impact the performance of the model in later time steps. Another observation one may notice is that at the 10th time step (i.e., when 10 data points are available for each individual), the performance of the independent logit model (ILM) has a sudden increase and becomes higher than the performance of OLCM-T we propose. This is possibly due to some properties of the dataset we have in the real-world case study and the characteristics of the individuals in the experiment. Besides, our model is shown to be able to estimate a large number of unknown variables, where ILM may require many more data points to get estimates. From Figure 5.8, we could also notice that the performance of OCLM-T is just slightly better than other models in early time steps. One reason may be that the polynomial model used here to represent an individual's changing preference is just a tentative example. Better models that could capture the dynamic of the changing preferences may lead to better performance and would require further explorations in our future work.

We'd like to present some results showing the changing preferences we learned in the choice-making process. The results are obtained in one of the 10 times of independent algorithm running. Same as other 9 times, in the beginning, 661 individuals' data is randomly selected to learn canonical models for the population. For each preference dimension, two canonical models are identified. The dataset from the other 165 individuals is used to test OLCM-T (online collaborative model with time-varying preferences) and learn their

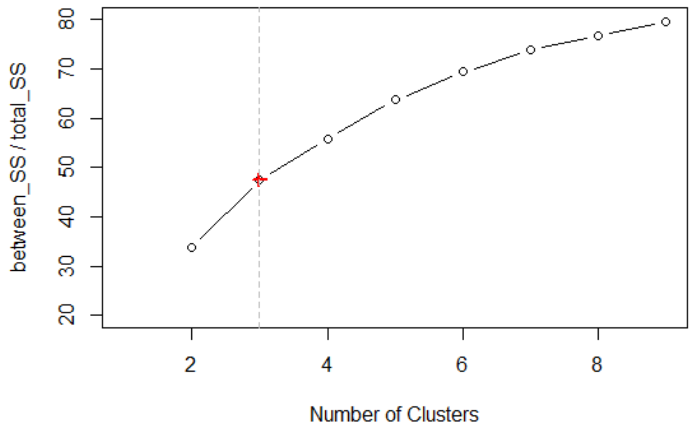


Figure 5.9: The ratio of between-cluster-sum-of-squares and total-sum-of-squares when dividing the individuals in to different number of clusters.

changing preferences.

To illustrate the changing preferences, we present the individual preferences learned by OLCM-T in an aggregated way: we cluster the 165 individuals into clusters using k-means according to their membership vectors (see Figure 5.9), and show each individual's preference changing curves in Figure 5.10. It can be seen that each individual has his own preferences learned by the model, and each individual's preferences are changing in his choice-making process. As we expected, most individuals have negative β_{SDE} and β_{SDL}, and positive β_{TTS} and β_{RP}, which is consistent with the common knowledge we have in transportation behaviors. The changing pattern of each individual is different from each other, possibly due to the different scenarios presented to him.

The results of the real-world case study show that our proposed model contributes to the prediction of the behaviors, which further proves that people's preferences may vary over time when making sequential choices.

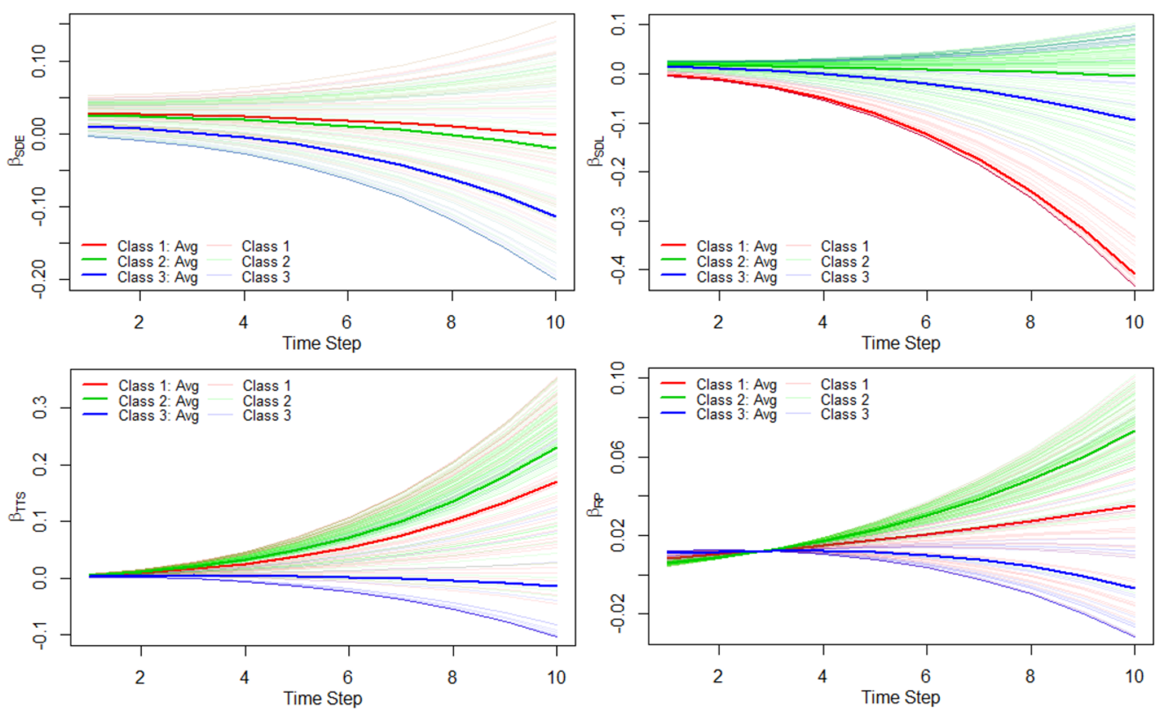


Figure 5.10: The curves of each individual's changing preferences (in light red, green, and blue), and the average preference changing curves in each clusters (in red, green, and blue).

Chapter 6

INTEGRATING THE COLLABORATIVE LEARNING MODEL WITH A TIME-INVARIANT MODEL

In Chapter 5, we elaborate on integrating a time-varying model with the collaborative learning model so that we could learn an individual's changing preferences. The time-varying model used in Chapter 5 is a polynomial model with one variable time step t. Though the polynomial model could capture the variations in an individual's preferences, it does not connect the preferences with other available information in each choice scenario. In this chapter, we will use another linear model as $\beta(t)$ could not only capture the preference changes but also describe how the related attributes impact the individual's preferences.

6.1 Linear model of preferences with more factors

From literature, it is known that an individual's preferences may vary according to specific choice scenarios and attributes, or evolve gradually with personal experiences [21, 97, 81, 104]. Given that, we formulate a model that takes both influences in the formulation.

The adoption of this model type into the preference learning process is based on three assumptions:

1. The attributes $\boldsymbol{x}_t$ of the alternatives presented to an individual may change his preferences when he makes a choice.

2. An individual's preferences $\boldsymbol{\beta}_t$ may also be related to his preferences $\boldsymbol{\beta}_{t-1}$ at the previous time step.

3. The impacts of the preferences at the previous time step and the attributes at the current time step to the preferences at the current time step may not evolve.

Denote the preference vector for individual i at time step t as $\boldsymbol{\beta}_{it}$. Also denote the attribute vector for individual i at time step t, i.e., the vector of the attribute differences between the two alternatives, as $\boldsymbol{x}_{it}$. Then we could write the preference evolution equation as:

$$\boldsymbol{\beta}_{it} = \boldsymbol{A}\boldsymbol{\beta}_{i(t-1)} + \boldsymbol{B}\boldsymbol{x}_{it} + \boldsymbol{v}_{it} \tag{6.1}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are two matrices converting vector $\boldsymbol{\beta}_{i(t-1)}$ and $\boldsymbol{x}_{it}$ into the new state $\boldsymbol{\beta}_{it}$, and $\boldsymbol{v}_{it}$ is a random noise vector. We could notice that different from the polynomial model in Chapter 5 where time t is directly in the equation, the time-varying model in this chapter is not a direct function of time. The preference $\boldsymbol{\beta}_{it}$ changes at each time step not directly because of t, but because that the attributes $\boldsymbol{x}_{it}$ and preferences $\boldsymbol{\beta}_{i(t-1)}$ change at time step t.

Further denote $\boldsymbol{\phi}_i = \text{vec}([\boldsymbol{A} \quad \boldsymbol{B}]^T)$. Given the notations, the predicted preferences of individual i at time step t given the preferences $\hat{\boldsymbol{\beta}}_{it}$ and the attributes of the alternatives $\boldsymbol{x}_{it}$ at time step t can be written as:

$$\hat{\boldsymbol{\beta}}_{it} = (\boldsymbol{I}_R \otimes [\boldsymbol{\beta}^T_{i(t-1)} \quad \boldsymbol{x}^T_{it}])\boldsymbol{\phi}_i = (\boldsymbol{I}_R \otimes \boldsymbol{Z}_{it})\boldsymbol{\phi}_i \tag{6.2}$$

where $\boldsymbol{\beta}$ is assumed to have R dimensions, $[\boldsymbol{\beta}^T_{i(t-1)} \quad \boldsymbol{x}^T_{it}]$ is denoted as $\boldsymbol{Z}_{it}$ (the input matrix as a whole), $\boldsymbol{I}_R$ is a identity matrix of size R, and $\otimes$ represents the Kronecker product of two matrices.

This equation we formulate shares similar formulation with Time-Invariant model, which is commonly used to describe dynamic systems where the state is also assumed to be impacted by the attributes at each time step and the state of the previous time step [30]. However, in the common time-invariant model, the state $\boldsymbol{\beta}_{it}$ is assumed measurable. In this kind of case, the observation can be directly compared with the predicted $\hat{\boldsymbol{\beta}}_{it}$, and a closed-form solution can be obtained [30]. However, in our study, the observation is individual i's choice behavior rather than the preferences (i.e., the state). Thus, we still have the observability issue here, since we could not directly observe the preferences but the choice behaviors. Again, we still use the discrete choice model to connect an individual's preferences and choice behaviors.

Given that the model shares some similarities with the time-invariant model, we also name the time-varying model of $\beta(t)$ shown in Equation 6.2. The specification of the integration of the time-varying model of $\beta(t)$ and the collaborative learning model is then illustrated in the following.

6.2 Time-Invariant Collaborative Model (TICM) specification

To be consistent with the notations of the canonical models in previous sections, hereafter we use $\boldsymbol{q}$ to replace $\boldsymbol{\phi}$s in the time-invariant model. Assume that there are K different parameter vectors $\boldsymbol{q}_1, \boldsymbol{q}_2, ..., \boldsymbol{q}_K$, representing K common types of mapping relationship from $\boldsymbol{\beta}_{t-1}$ and $\boldsymbol{x}_t$ (i.e. $\boldsymbol{Z}_{it}$) to $\boldsymbol{\beta}_t$ for all the N individuals in the population. We call the K common types of mapping relationship the "canonical models". Here $\boldsymbol{q}_k = \text{vec}([\boldsymbol{A}_k \quad \boldsymbol{B}_k]^T)$, and $\boldsymbol{\beta}_t = \boldsymbol{A}_k\boldsymbol{\beta}_{t-1} + \boldsymbol{B}_k\boldsymbol{x}_t + \boldsymbol{v}_t$. The canonical model matrix is denoted as $\boldsymbol{Q} = [\boldsymbol{q}_1, \boldsymbol{q}_2, ..., \boldsymbol{q}_K]$.

For each individual i, $i = 1, 2, ..., N$, there is a membership vector $\boldsymbol{c}_i = [c_{i1}, c_{i2}, \ldots, c_{iK}]^T$ assigned to him, representing the degree to which the mapping relationship of the individual resembles the corresponding K canonical models. For an individual i, $\boldsymbol{q}_i = \boldsymbol{Q}\boldsymbol{c}_i$

At time step t, the individual i's predicted preferences can be formulated as:

$$\boldsymbol{\beta}_{it} = \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i \tag{6.3}$$

And the probability of the promoted alternative being chosen by individual i at time step t is:

$$\pi_i(\boldsymbol{x}_{it}) = Pr(y_{it} = 1|\boldsymbol{x}_{it}) = \frac{\exp(\boldsymbol{x}_{it}^T\boldsymbol{I}_R \otimes \boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i)}{1 + \exp(\boldsymbol{x}_{it}^T\boldsymbol{I}_R \otimes \boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i)} \tag{6.4}$$

The log-likelihood function of the binary logit model given Equation 6.4 is:

$$l = -\log(1 + \exp\{\boldsymbol{x}_{it}^T\boldsymbol{I}_R \otimes \boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i\}) + y_{it}\{\boldsymbol{x}_{it}^T\boldsymbol{I}_R \otimes \boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i\} \tag{6.5}$$

To get the estimates, an objective function for the optimization problem integrating time-invariant model and Collaborative Learning Model can be formulated as:

$$\begin{aligned} \min_{\boldsymbol{Q},\boldsymbol{c}_i} \quad & \sum_{i=1}^{N}\sum_{t=1}^{T}\left(\log\left(1 + \exp(\boldsymbol{x}_{it}^T\boldsymbol{I}_R \otimes \boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i)\right) - y_{it}\big(\boldsymbol{x}_{it}^T\boldsymbol{I}_R \otimes \boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i\big)\right) \\ \text{s.t.} \quad & \boldsymbol{c}_i \geq 0,\ {\boldsymbol{c}_i}^T\boldsymbol{1} = 1 \quad i = 1, ..., N. \end{aligned} \tag{6.6}$$

The objective function of Equation 6.6 is to minimize $-l$, the Equation 6.5 (equals to maximizing the log-likelihood function of the binary choice model l).

6.3 Parameter estimation algorithm

To identify $\boldsymbol{Q}$ (canonical models) and $\boldsymbol{c}_i$ (individual membership vectors) separately with the formulation of the optimization problem shown in Equation 6.6, a two-step iteratively updating strategy is adopted, following existing works where a similar strategy is applied to a linear collaborative model [112, 111, 113]. Specifically, we iteratively optimize $\boldsymbol{Q}$ and $\boldsymbol{c}_i$ in the strategy: $\boldsymbol{c}_i$ is fixed when $\boldsymbol{Q}$ is optimized in "Q-step", and $\boldsymbol{Q}$ is fixed when $\boldsymbol{c}_i$ is optimized in "C-step". In the following, the details of the strategy are presented.

6.3.1 Parameter estimation for canonical models (Q-step)

At Q-step, the parameters of canonical models $\boldsymbol{Q}$ are updated, while the individual vectors $\boldsymbol{c}_i$ are fixed. This means that the $\boldsymbol{c}_i$s are known constants, i.e. the estimates obtained from the previous "C-step" iteration. Therefore, the objective function becomes:

$$\min_{\boldsymbol{Q}} \quad \sum_{i=1}^{N}\sum_{t=1}^{T}\Bigg(\log\big(1+\exp(\boldsymbol{x}_{it}^T\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i)\big)-y_{it}\big(\boldsymbol{x}_{it}^T\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i\big)\Bigg) \tag{6.7}$$

To optimize $\boldsymbol{Q}$, the optimization problem shown in Equation 6.7 needs to be reformulated:

1. The canonical model matrix $\boldsymbol{Q}$ (which has K columns and $2R^2$ rows) can be converted to a vector $\boldsymbol{q}$, $\boldsymbol{q} = \mathrm{vec}(\boldsymbol{Q})$. Assuming that an individual's preferences $\boldsymbol{\beta}_i$ has R dimensions, as well as the attribute vector $\boldsymbol{x}_{it}$, the length of the vector $\boldsymbol{q}$ is $2KR^2$.

2. The term $\boldsymbol{I}_R$ in Equation 6.7 is converted to $\boldsymbol{I}_{KR}$, which means that the size of the identity matrix is expanded to KR.

3. We further convert $\boldsymbol{x}_{it}^T$ to a matrix $\boldsymbol{X}_{it} = \boldsymbol{I}_K \otimes \boldsymbol{x}_{it}^T$, where $\boldsymbol{I}_K$ is a identity matrix of size K. Given the definition of the kronecker product, we have

$$\boldsymbol{X}_{it} = \begin{bmatrix} \boldsymbol{x}_{it}^T & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \boldsymbol{x}_{it}^T \end{bmatrix}_{K \times KR} \tag{6.8}$$

With the adjusted notations, the objective function of Equation 6.7 can be written as:

$$\min_{\boldsymbol{q}} \quad \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\log\left(1 + \exp(\boldsymbol{c}_i^T \boldsymbol{X}_{it} \boldsymbol{I}_{KR} \otimes \boldsymbol{Z}_{it} \boldsymbol{q})\right) - y_{it}\left(\boldsymbol{c}_i^T \boldsymbol{X}_{it} \boldsymbol{I}_{KR} \otimes \boldsymbol{Z}_{it} \boldsymbol{q}\right) \right) \tag{6.9}$$

Since $\boldsymbol{c}_i^T \boldsymbol{X}_{it} \boldsymbol{I}_{KR} \otimes \boldsymbol{Z}_{it}$ is known at each time step t, we may define $\boldsymbol{\chi}_{it} = \boldsymbol{c}_i^T \boldsymbol{X}_{it} \boldsymbol{I}_{KR} \otimes \boldsymbol{Z}_{it}$. Thus the objective function could be simplified as:

$$\min_{\boldsymbol{q}} \quad \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\log\left(1 + \exp(\boldsymbol{\chi}_{it} \boldsymbol{q})\right) - y_{it}\left(\boldsymbol{\chi}_{it} \boldsymbol{q}\right) \right) \tag{6.10}$$

This turns out to be the log-likelihood function of logistic regression, which could be estimated by many existing methods for optimization problem-solving. We use CVXR, an R package for specifying and solving convex programs [60] to solve the problem in Equation 6.10.

6.3.2 Parameter estimation for membership vectors (C-step)

At C-step, the parameters of canonical models $\boldsymbol{Q}$ are fixed to be the estimates obtained from the previous "Q-step", while the individual vectors $\boldsymbol{c}_i$ are to be estimated. Therefore, the objective function becomes:

$$\begin{aligned} \min_{\boldsymbol{c}_i} \quad & \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\log\left(1 + \exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)\right) - y_{it}\left(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i\right) \right) \\ \text{s.t.} \quad & \boldsymbol{c}_i \geq 0,\ {\boldsymbol{c}_i}^T \mathbf{1} = 1 \quad i = 1, ..., N. \end{aligned} \tag{6.11}$$

Given Q, the Lagrangian function of the objective shown in Equation 6.11 could be derived as:

$$\mathcal{L} = \sum_{i=1}^{N} \sum_{t=1}^{T} \Big(\log\big(1 + \exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)\big) - y_{it} \big(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i\big) \Big) + \sum_{i=1}^{N} \lambda_i (\boldsymbol{c}_i \mathbf{1} - 1) \tag{6.12}$$

The complementary condition gives that $\frac{\partial \mathcal{L}}{\partial c_{ik}} c_{ik} = 0$, where c_{ik} is the k's dimension of the membership vector of individual i. Thus, we :

$$\frac{\partial \mathcal{L}}{\partial c_{ik}} c_{ik} = \sum_{t=1}^{T} \left(\frac{\exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)}{1 + \exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)} - y_{it} \right) \big[(\boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q})^T \boldsymbol{x}_{it}\big]_k c_{ik} + \lambda_i c_{ik} = 0 \tag{6.13}$$

Adding $\boldsymbol{c}_i \mathbf{1} = 1$, i.e. $\sum_{k=1}^{K} c_{ik} = 1$, to equation 6.13, we could also have:

$$\sum_{t=1}^{T} \left(\frac{\exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)}{1 + \exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)} - y_{it} \right) \boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i + \lambda_i = 0 \tag{6.14}$$

With Equation 6.14, we could obtain the formulation for the Lagrangian multiplier λ_i:

$$\lambda_i = \sum_{t=1}^{T} \left(y_{it} - \frac{\exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)}{1 + \exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)} \right) \boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i \tag{6.15}$$

Plug Equation 6.15 into Equation 6.13, we have:

$$\begin{aligned} &\sum_{t=1}^{T} \left(y_{it} - \frac{\exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)}{1 + \exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)} \right) \big[(\boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q})^T \boldsymbol{x}_{it}\big]_k c_{ik} \\ &\quad = \sum_{t=1}^{T} \left(y_{it} - \frac{\exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)}{1 + \exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)} \right) \boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i c_{ik} \end{aligned} \tag{6.16}$$

I.e.,

$$\begin{aligned} &\sum_{t=1}^{T} \left(y_{it} \big[(\boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q})^T \boldsymbol{x}_{it}\big]_k - \frac{\exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)}{1 + \exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)} \big[(\boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q})^T \boldsymbol{x}_{it}\big]_k \right) c_{ik} \\ &\quad = \sum_{t=1}^{T} \left(y_{it} \boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i - \frac{\exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)}{1 + \exp(\boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i)} \boldsymbol{x}_{it}^T \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it} \boldsymbol{Q} \boldsymbol{c}_i \right) c_{ik} \end{aligned} \tag{6.17}$$

Since c_{ik} should be non-negative, we define $\delta_+(x) \equiv \max(x, 0)$ and $\delta_-(x) \equiv \min(x, 0)$, with which all terms in Equation 6.16 could be separated into a positive part and a negative part:

$$
\begin{aligned}
&c_{ik}\Bigg\{\sum_{t=1}^{T}\Bigg[\delta_+\Bigg(y_{it}\big[(\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q})^T\boldsymbol{x}_{it}\big]_k\Bigg)-\delta_+\Bigg(\frac{\exp(\boldsymbol{x}_{it}^T\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i)}{1+\exp(\boldsymbol{x}_{it}^T\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i)}\big[(\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q})^T\boldsymbol{x}_{it}\big]_k\Bigg)\\
&\quad+\delta_-\Bigg(y_{it}\big[(\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q})^T\boldsymbol{x}_{it}\big]_k\Bigg)-\delta_-\Bigg(\frac{\exp(\boldsymbol{x}_{it}^T\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i)}{1+\exp(\boldsymbol{x}_{it}^T\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i)}\big[(\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q})^T\boldsymbol{x}_{it}\big]_k\Bigg)\Bigg]\Bigg\}\\
&\quad=c_{ik}\Bigg\{\sum_{t=1}^{T}\Bigg[\delta_+\Bigg(y_{it}\boldsymbol{x}_{it}^T\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i\Bigg)-\delta_+\Bigg(\frac{\exp(\boldsymbol{x}_{it}^T\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i)}{1+\exp(\boldsymbol{x}_{it}^T\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i)}\boldsymbol{x}_{it}^T\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i\Bigg)\\
&\qquad+\delta_-\Bigg(y_{it}\boldsymbol{x}_{it}^T\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i\Bigg)-\delta_-\Bigg(\frac{\exp(\boldsymbol{x}_{it}^T\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i)}{1+\exp(\boldsymbol{x}_{it}^T\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i)}\boldsymbol{x}_{it}^T\boldsymbol{I}_R\otimes\boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i\Bigg)\Bigg]\Bigg\}
\end{aligned}
\tag{6.18}
$$

Let's further define $\boldsymbol{Q}^* = \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it}\boldsymbol{Q}$. With $\delta_+(\cdot)$ and $\delta_-(\cdot)$ in Equation 6.18, we are able to derive an updating rule for c_{ik} such that the updated c_{ik} is positive:

$$
\begin{aligned}
c_{ik}^{m+1} =& c_{ik}^{m}\times\\
&\Bigg\{\sum_{t=1}^{T}\Bigg[+\delta_+\Bigg(\frac{\exp(\boldsymbol{x}_{it}^T\boldsymbol{Q}^*\boldsymbol{c}_i)}{1+\exp(\boldsymbol{x}_{it}^T\boldsymbol{Q}^*\boldsymbol{c}_i)}\boldsymbol{x}_{it}^T\boldsymbol{Q}^*\boldsymbol{c}_i\Bigg)+\delta_+\Bigg(y_{it}\big[(\boldsymbol{Q}^*)^T\boldsymbol{x}_{it}\big]_k\Bigg)\\
&-\delta_-\Bigg(y_{it}\boldsymbol{x}_{it}^T\boldsymbol{Q}^*\boldsymbol{c}_i\Bigg)-\delta_-\Bigg(\frac{\exp(\boldsymbol{x}_{it}^T\boldsymbol{Q}^*\boldsymbol{c}_i)}{1+\exp(\boldsymbol{x}_{it}^T\boldsymbol{Q}^*\boldsymbol{c}_i)}\big[(\boldsymbol{Q}^*)^T\boldsymbol{x}_{it}\big]_k\Bigg)\Bigg]\Bigg\}\\
&/\\
&\Bigg\{\sum_{t=1}^{T}\Bigg[\delta_+\Bigg(\frac{\exp(\boldsymbol{x}_{it}^T\boldsymbol{Q}^*\boldsymbol{c}_i)}{1+\exp(\boldsymbol{x}_{it}^T\boldsymbol{Q}^*\boldsymbol{c}_i)}\big[(\boldsymbol{Q}^*)^T\boldsymbol{x}_{it}\big]_k\Bigg)+\delta_+\Bigg(y_{it}\boldsymbol{x}_{it}^T\boldsymbol{Q}^*\boldsymbol{c}_i\Bigg)\\
&-\delta_-\Bigg(y_{it}\big[(\boldsymbol{Q}^*)^T\boldsymbol{x}_{it}\big]_k\Bigg)-\delta_-\Bigg(\frac{\exp(\boldsymbol{x}_{it}^T\boldsymbol{Q}^*\boldsymbol{c}_i)}{1+\exp(\boldsymbol{x}_{it}^T\boldsymbol{Q}^*\boldsymbol{c}_i)}\boldsymbol{x}_{it}^T\boldsymbol{Q}^*\boldsymbol{c}_i\Bigg)\Bigg]\Bigg\}
\end{aligned}
\tag{6.19}
$$

In summary, the learning algorithm for the proposed Time-Invariant Collaborative Model (TICM) is:

TICM Learning Algorithm

Input

Data $\boldsymbol{x}_{it}$ for individual $i(i = 1, 2, \cdots, N)$ at time step $t(t = 1, 2, \cdots, T)$;

Initial value $\boldsymbol{Q}^{(0)}$ and $\boldsymbol{c}_i^{(0)}$;

Maximum number of rounds of iteration $M_{Iteration}$;

Output

$\boldsymbol{Q}^{M_{Iteration}+1}$, $\boldsymbol{c}_i^{M_{Iteration}+1}$ for $i = 1, 2, \cdots, N$.

1. **for** each m in $[0, M_{Iteration}]$:
2. Convert $\boldsymbol{Q}^m$ to $\boldsymbol{q}^m$, $\boldsymbol{q}^m = \text{vec}(\boldsymbol{Q}^m)$; Convert $\boldsymbol{x}_{it}^T$ to $\boldsymbol{X}_{it}$, $\boldsymbol{X}_{it} = \boldsymbol{I}_K \otimes \boldsymbol{x}_{it}^T$;
3. Calculate $\chi_{it} = \boldsymbol{c}_i^T \boldsymbol{X}_{it} \boldsymbol{I}_{KR} \otimes \boldsymbol{Z}_{it}$ with $\beta_{i(t-1)}$ and current $\boldsymbol{c}_i^m$;
4. Solve Equation 6.10 and get $\boldsymbol{q}^{m+1}$;
5. Transform $\boldsymbol{q}^{m+1}$ to $\boldsymbol{Q}^{m+1}$ by partitioning $\boldsymbol{q}^{m+1}$ to the $2R^2 \times K$ matrix;
6. Calculate $\boldsymbol{c}_i^{m+1}$ by applying Equation 6.19.
7. **end for**

6.4 Simplified Time-Invariant Collaborative Model (TICM)

While converting preferences $\boldsymbol{\beta}_{t-1}$ and $\boldsymbol{x}_t$ to the current preferences $\boldsymbol{\beta}_t$ with matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ (evaluation equation 6.1), we are assuming that each dimension of the current preferences may be impacted by all the dimensions of the preferences and attributes. This is clearly reflected in the equation: Let the rth dimension of the preference vector $\boldsymbol{\beta}_{it}$ for individual i at time step t be denoted as β_{it}^r. According to the evolution equation, β_{it}^r generated by adding the product of the rth row of matrix $\boldsymbol{A}$ and $\boldsymbol{\beta}_{i(t-1)}$, and the product of the rth row of $\boldsymbol{B}$ and $\boldsymbol{x}_t$. This shows that the value of β_{it}^r not only depends on $\beta_{i(t-1)}^r$ or x_t^r, but is also impacted by other dimensions of $\boldsymbol{\beta}_{it}$ and $\boldsymbol{x}_t$. We may interpret this assumption in the following way: the existence of other dimensions of $\boldsymbol{\beta}_{it}$ and the present of other dimensions of $\boldsymbol{x}_t$ might possibly influence how the individual feels about x_t^r, i.e. β_t^r.

While this assumption can be reasonable, another issue arises. Assume that we have

K canonical models with R-dimensional preferences, and the total number of parameters to be estimated for canonical models is $2KR^2$. When the number of dimensions increases, the number of parameters to be estimated increases from $2KR^2$ to $2K(R+1)^2$. It can be noticed that not only the number grows when R is getting bigger, but also the growth itself increases. This may bring a heavy burden to the computation of the model. For example, when $R = 4$, i.e., there are five attributes in each scenario that may be influential in the choice-making behavior (thus, the individual preferences also have 4 dimensions), for each canonical model, the number of parameters to be estimated is $2 \times 4^2 = 32$. Each time when the number of canonical models K increases by one, there are 32 more parameters to be estimated. While the dimension increases to 5, the number of parameters for one canonical model increases to $2 \times 5^2 = 50$. If the number of canonical models is comparatively large, say, 7 or 8, the total number of parameters would be as many as 400. This may slow down the computation speed, while at the same time reduce the learning accuracy of the algorithm. Based on this consideration, we simplified the proposed model by assuming that each dimension of the preferences is only impacted by the corresponding dimension of preferences at the previous time step and the corresponding attribute. In other words, we may let both $\boldsymbol{A}$ and $\boldsymbol{B}$ be diagonal matrices. This significantly reduces the number of parameters to be learned for each canonical model.

In the following sections, we test this simplified Time-Invariant Collaborative Model (TICM) by running simulation and applying the model on a real-world dataset.

6.5 Simulations

In the simulations, each individual's preferences are described by time-dependent functions, which may be impacted by the preferences of the previous time step and the attributes of the current step, and his choice made at each time step is simulated based on his predicted preferences at the current time step according to his estimated model and the given attributes.

Similar with the previous chapter, the performance of the proposed TICL is compared with several benchmark methods including (1) the independent logistic regression model (ILM) that learns the regression coefficients of each individual solely based on her own data;

(2) the mixed-effect logistic regression model (logistic MEM) that considers the coefficients of individuals are extracted from a certain distribution; (3) the one-size-fits-all logistic regression model (LR) that treats all individuals homogeneously, combines all individuals' data together and estimates one set of preferences; (4) the original LCM model, where individual preferences are constant values rather than time-varying variables (the canonical models represent different types of utility models). The simulations in this section test each model's performance with all data available.

Given that the true coefficients are known in the simulations, Average Absolute Error is used as a metric to evaluate the performance of the models, which measures the difference between learned coefficients and the true ones (Equation 6.20):

$$\text{Average Absolute Error} = \frac{1}{N}\sum_{i=1}^{N}|\beta_i - \hat{\beta}_i|. \tag{6.20}$$

We also evaluate the performance of the models using Prediction Accuracy, as shown in Equation 6.21:

$$\text{Prediction Accuracy} = \frac{N_{\text{Number of predictions that are consistent with the actual choices made}}}{N_{\text{Number of predictions made in total}}} \tag{6.21}$$

6.5.1 Design of the simulation experiment and data generation

In simulations of the proposed Time-Invariant Collaborative Model (TICM), a given number of canonical models are generated by manually setting the parameters. Since the canonical models represent different preference changing patterns, they need to be different enough from each other. For example, when the number of canonical models is set to be 2, we may let the matrix $\boldsymbol{A}$ of one canonical model to be a matrix of zero, and the matrix $\boldsymbol{B}$ of the other canonical model be zero. With this setting, the preferences generated by the first canonical model are all impacted by the preferences of the previous time step, and the preferences generated by the second canonical model are all decided by the attributes of the current time step. Similarly, when we set the number of the canonical models to be 3 or more, we could let the different dimensions of the preferences be impacted only by either $\boldsymbol{A}$

or $\boldsymbol{B}$. For instance, assuming that we are generating 3 canonical models for preferences with two dimensions. Two of the canonical models could still be obtained by setting $\boldsymbol{A}$ or $\boldsymbol{B}$ as a matrix of zero. For the third canonical model, we may let the first preference dimension be impacted by matrix $\boldsymbol{A} = \begin{bmatrix} a_{11} & 0 \\ 0 & 0 \end{bmatrix}$, and the second preference dimension impacted by matrix $\boldsymbol{B} = \begin{bmatrix} 0 & 0 \\ 0 & b_{22} \end{bmatrix}$. Notice that since we use simplified Time-Invariant Collaborative Model (TICM), both $\boldsymbol{A}$ and $\boldsymbol{B}$ are diagonal matrices. Then the three canonical models are significantly different from each other.

With canonical models, each individual's membership vectors need to be decided such that her preferences $\boldsymbol{\beta}_i$ could be obtained via $\boldsymbol{\beta}_i = \boldsymbol{I}_R \otimes \boldsymbol{Z}_{it}\boldsymbol{Q}\boldsymbol{c}_i$. Given the requirements that $\boldsymbol{c}_i\boldsymbol{1} = 1$ and $c_{ki} \geq 0$, we use Dirichlet distribution to model the membership vectors (the commend is *rdirichlet* in R). In the simulation, we let each individual has her dominant preference changing model. This is achieved by setting one big value (e.g., 10) in the p-length vector $\boldsymbol{\alpha}$ and let other values of the vector be small (e.g., 1). Also, to guarantee that the $\boldsymbol{\beta}_i$ are suitable in simulation, the $\boldsymbol{\beta}_i$ are mandatorily constrained to be between -1 and 1. This is just to make sure that the calculated utilities will not be too large and cause errors in simulations.

We then generate choice scenario attribute vector $\boldsymbol{x}_{it}$ for each individual i at each time step t. Considering the utility function $U_{it} = \boldsymbol{x}_{it}\boldsymbol{\beta}_{it}$ used in the simulation, we randomly select the value of each dimension of the attribute vector from a uniform distribution.

We also test the performance of the model given different levels of noise in responses, i.e., 5% (0.05), 10% (0.1), 15% (0.15), which represent the percentages of the incorrect responses for each individual in all his choices. To achieve this, a random number between 0 and 1 is generated after generating a scenario and the corresponding true response from the true preferences. If the number is smaller than the noise level (e.g., 0.1), we turn the response to the opposite (from acceptance to rejection, or from rejection to acceptance), making the response in this scenario an incorrect response.

The proposed TICM learns the rule for how an individual's preferences change over time. With the rule, the individual's preferences and his response given a new scenario with new attributes could be predicted. In this chapter, we first fed all the data into the algorithm in the learning process. Then we tried to update the preferences at each time step by setting a moving window to decide the datasets for training and testing shown in Figure 6.1. In the figure, each cell represents one data point from one individual, and each row represents all the data points from one individual. The moving window allows us to use the latest data to train the model and make predictions. Similar to the Online Collaborative Model with Time-Varying parameters (OLCM-T) discussed in Chapter 5, the evolution of the Time-Invariant Collaborative Model (TICM) parameters over time can also be updated with online and offline stages when new observations are available. Since this chapter aims to preliminarily present a model that could be used to capture preference changes when given different scenarios over time, how to update the parameters such that the evolution of the model could be tracked is beyond the scope of this chapter, thus related simulations and discussions are not provided here. To update the parameters and capture the evolution of the models can be future studies of this dissertation.

In our simulation, we set the total number of individuals $N = 240$, the number of preference parameters in utility function $R = 2$ (i.e., the individual preferences have 2 dimensions), and the number of canonical models for each preference dimension $K = 2$. With this basic setting, each individual a 2-dimensional membership vector towards the 2 dimensions of his preferences, and his preferences could be described by the two canonical models and his membership vector. Tests regarding the performances of the model with different numbers of canonical models and different numbers of dimensions (i.e., the number of variables in the utility function) are also conducted accordingly.

We present the results of the prediction accuracy, computation time, and the Average Absolute Error of an estimate for each model, with the Time-Invariant Collaborative Model (TICM) having same numbers of preference dimension and canonical models.

6.5.2 Simulation results

Figure 6.2 shows the prediction accuracy of different models when in simulation settings there 2, 3, 4, and 5 canonical models respectively. Notice that the number of canonical

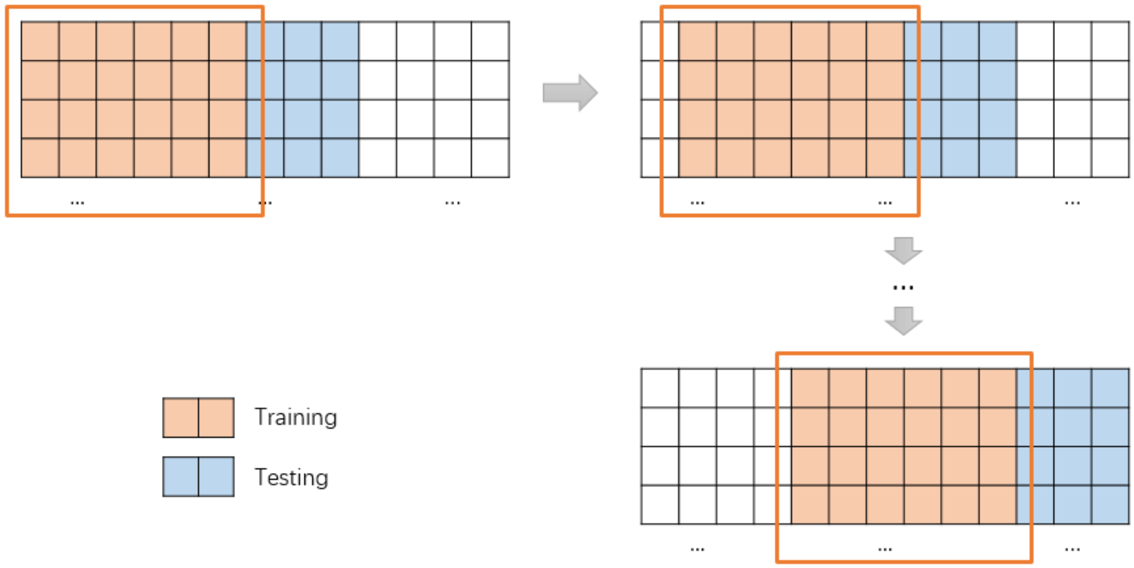


Figure 6.1: Illustration of the moving window in model training and testing.

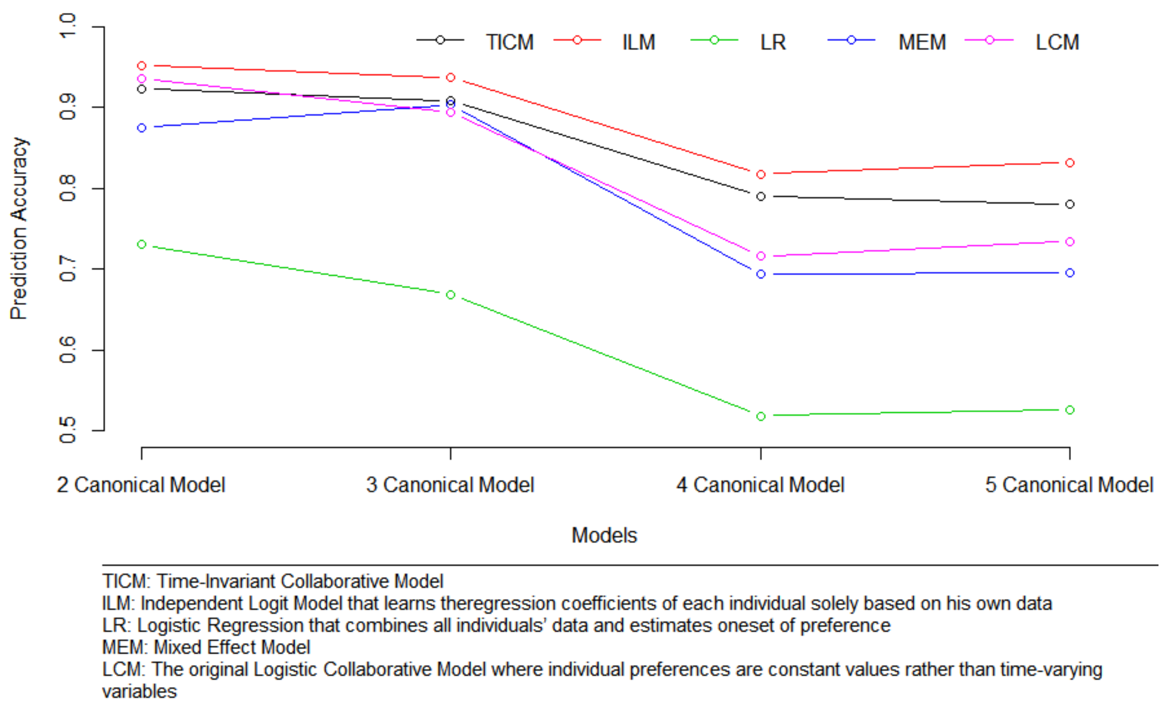


Figure 6.2: Prediction accuracy when having different number of canonical models (with 2-dimensional preferences).

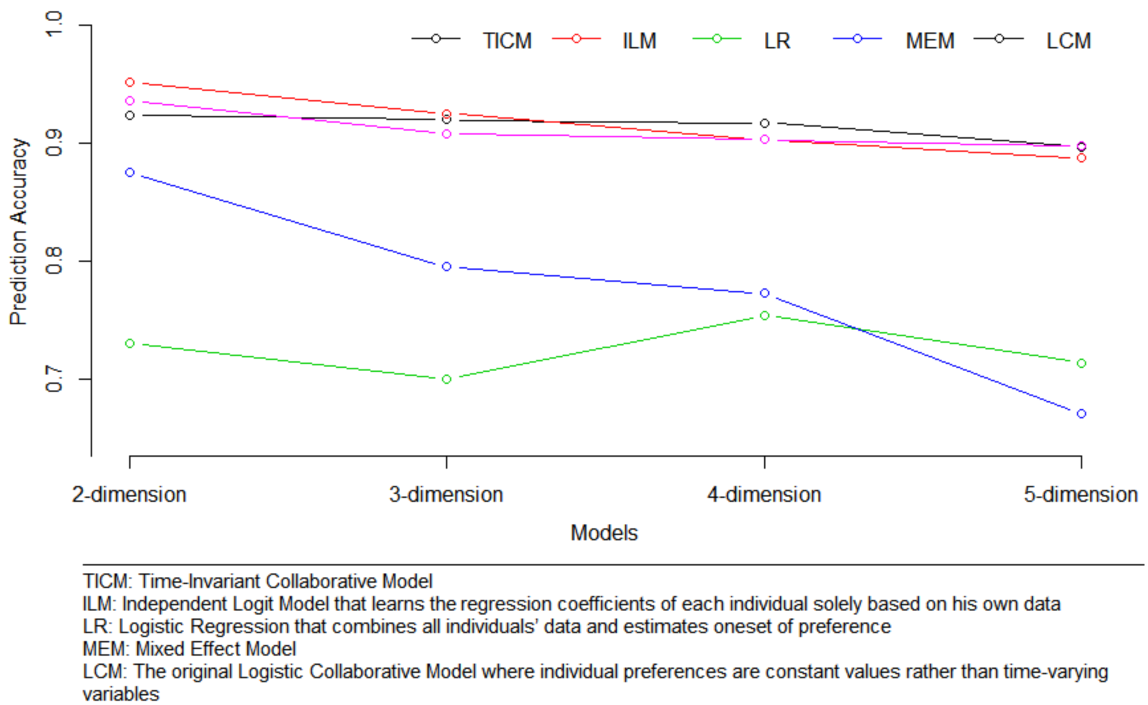


Figure 6.3: Prediction accuracy when having different number of dimensions in preferences (with 2 canonical models).

Table 6.1: Results Model Comparison

Model	TICM	LCM	ILM	LR	MEM
Average Computation Time (s)	589.79	50.81	0.67	0.01	278.81
Average Absolute Error	0.721	0.605	0.375	253.9	0.911

TICM: Time-Invariant Collaborative Model; LCM: Original Logistic Collaborative Model; ILM: Independent logit Model; LR: Logistic Regression; MEM: Mixed Effect Model

Table 6.2: Prediction accuracy when data has different levels of noise.

Noise Level	Prediction Accuracy
0%	0.92
5%	0.82
10%	0.76
15%	0.69

models is the setting for the true preferences of the population in the simulation. Thus the Prediction Accuracy (PA) of each model is the prediction accuracy if the model is applied to the data. It can be noticed that the PA (prediction accuracy) of LR (logistic regression), MEM (mixed effect model), and LCM (original logistic collaborative model) significantly decreases when the number of canonical models increases, which means the heterogeneity in the population increases. LR (logistic regression) has the lowest PA (prediction accuracy) is reasonable because the heterogeneity could not be captured by the logistic regression (LR), which learns one set of preferences for all individuals. Though in different ways, LCM and MEM (mixed effect model) could handle some taste variations. However, since the mechanism of the preference changes is different from what is assumed in MEM and LCM (original logistic collaborative model), their learning performances are not satisfactory. ILM shows better prediction accuracy than the proposed TICM. This may be due to the settings we have in the simulation, and this issue is further discussed in the last chapter.

Table 6.1 shows the average computation time and the average absolute error for different models. Notice that the computation time and absolute error shown in the column of TICM (time-invariant collaborative model) are for TICM with 2 canonical models and 3 dimensions, and the criteria for the iteration process is $|Loss_t - Loss_{t-1}| < 0.05$. The results in Table 6.2 come from 10 independent simulation runs of TICM for 2-dimensional preferences with 2 canonical models, at time step 10 when 10 data points are available.

Figure 6.4 shows the prediction accuracy at each time step when we use the moving

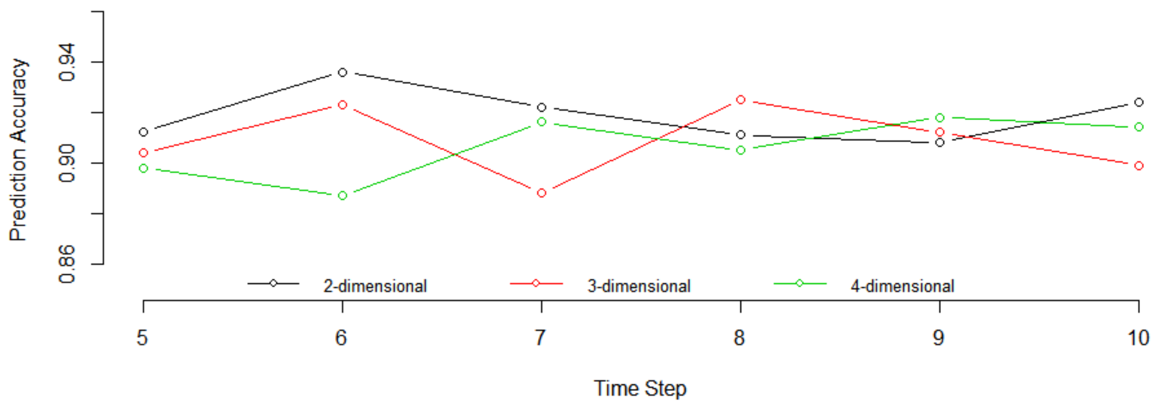


Figure 6.4: Prediction accuracy when the parameters are updated at each time step with a moving window for training dataset (the number of canonical models is 2).

window in the updating process. From time step 5, the previous 5 data points from each individual are aggregated to be the learning dataset, and the following 5 data points are used to test. Different numbers of preference dimensions are tested in the process, while the number of canonical models is fixed at 2. No obvious trend could be found from the figure, showing that while the size of the training dataset is unchanged, the prediction accuracy remains at a stable level. Similar to Figure 6.3, the increasing number of dimensions has mild impacts on the performance of the model.

6.6 Application to a real-world dataset

The dataset we use in this section is the same subset of the whole dataset collected in an online experiment we conducted in our previous study, which includes 200 individuals and their choice-making data of 13 sequential scenarios. For each individual, the first 10 data points are used to train the model, and the last 3 data points are used for testing.

Similarly, Prediction Accuracy (PA) is used here as a metric to evaluate the performance of the model when applied to the dataset.

$$PA = \frac{N_{\text{Times when the prediction is inconsistent with the true choice}}}{N_{\text{Times of the predictions made in total}}} \tag{6.22}$$

Before applying the proposed TICM to the dataset, we need to decide the number of

Table 6.3: Model Comparison

Model	TICM	LCM	ILM	LR	MEM
Prediction Accuracy	**0.847**	0.827	0.818	0.468	0.695

TICM: Time-Invariant Collaborative Model; LCM: Original Logistic Collaborative Model; ILM: Independent logit Model; LR: Logistic Regression; MEM: Mixed Effect Model

the canonical models. Figure 6.5 suggest that the best number of the canonical model for the dataset is 5, with which the percentage of prediction error is 0.153. We also apply ILM, LR, MEM, and LCM to the dataset. Similar to the previous chapter, here, the number of canonical models for LCM is still chosen as 5. From the results shown in Table 6.3 we could see that TICM has the highest prediction accuracy $PA = 0.847$. LCM, which integrates the collaborative learning structure with a utility function, has $PA = 0.827$, slightly lower than the one of TICM. Since the dataset used here in this section is the same as the one in Section 5.5, here the PAs of LCM (original Logistic Collaborative Model), ILM (Independent Logit Model), LR (Logistic Regression), and MEM (Mixed Effect Model) are the same as the results of these models at the 10th time step.

We further randomly sampled 10 other dataset with 200 individuals from the original dataset obtained with the online experiment with 1839 individuals, and applied TICM to estimate the average PA for each model. Table 6.4 shows that the proposed TICM does return results with a higher accuracy rate than other models. With the proposed model, an individual's preferences at each time step are different, corresponding to the time step and the attributes of the scenario. For example, the changes in the learned preference of dimension SDE for an individual in our dataset are shown in Figure 6.6(a). We could see that the higher the scheduled delay early (SDE) is, the lower his preferences towards the attribute is, which means that he/she dislikes it more.

We also test the performance of TICM if we update the parameters at each time step, as shown in Figure 6.1. At each time step, the previous 5 data points are used in training

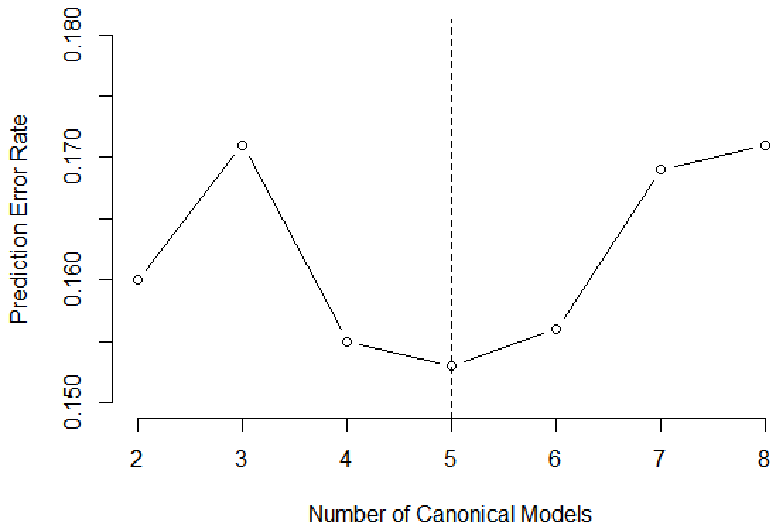
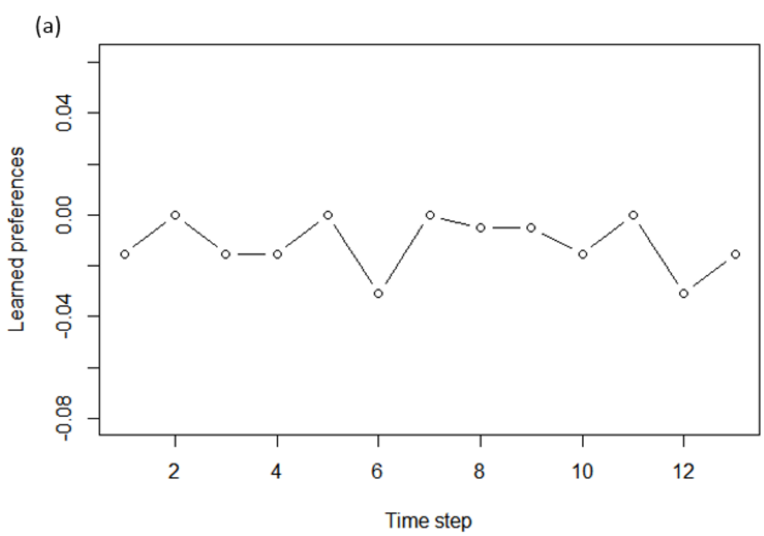


Figure 6.5: Prediction error rates when having different number of canonical models in Time-Invariant Collaborative Model (TICM).



(b)

sde	sdl	tts	reward_pts	y
30	0	5	20	1
0	30	5	20	0
30	0	10	20	1
30	0	5	30	0
0	10	5	20	0
60	0	10	20	0
0	30	10	20	1
10	0	5	20	1
10	0	5	30	1
30	0	10	30	1
0	10	10	20	1
60	0	10	30	0
30	0	10	20	1

Figure 6.6: (a) Changes of the learned preference to attribute “Scheduled Delay Early” (SDE) of an individual in his 13 sequential scenarios; (b) The scenario sequence displayed to him (column “sde”, “sdl”, “tts”, and “reward pts”) and his responses (column “y”).

Table 6.4: Prediction Accuracy - 10 Random selected Datasets.

Dataset No.	TICM	LCM	ILM	LR	MEM
1	**0.874**	0.832	0.850	0.443	0.641
2	**0.861**	0.805	0.805	0.447	0.625
3	**0.866**	0.818	0.817	0.460	0.632
4	**0.855**	0.803	0.795	0.468	0.637
5	**0.859**	0.840	0.793	0.478	0.697
6	**0.856**	0.817	0.828	0.488	0.642
7	**0.851**	0.815	0.833	0.485	0.708
8	**0.879**	0.812	0.807	0.500	0.680
9	**0.862**	0.840	0.813	0.518	0.568
10	**0.881**	0.841	0.855	0.397	0.673
Average	**0.864**	0.822	0.820	0.469	0.640

TICM: Time-Invariant Collaborative Model; LCM: Original Logistic Collaborative Model; ILM: Independent logit Model; LR: Logistic Regression; MEM: Mixed Effect Model

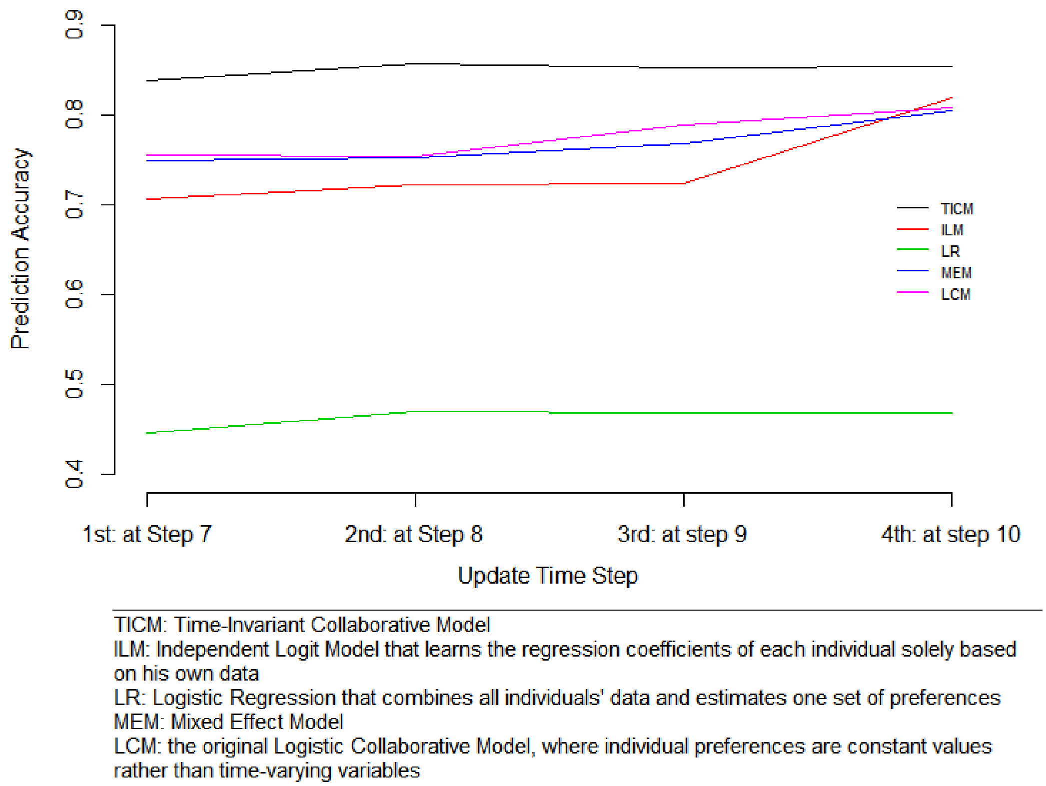


Figure 6.7: Prediction accuracy if parameters in Time-Invariant Collaborative Model (TICM) is updated at each time step with a moving window.

dataset, and the following 3 data points are used for testing. Figure 6.7 shows the results: The proposed TICM has the best performance, and the prediction accuracy at each time step is significantly higher than other models.

Chapter 7

CONCLUSION AND DISCUSSION

7.1 Conclusion

Learning an individual's preferences accurately and knowing how much the individual values each influential factor is essential when providing personalized services or incentives. The significant difficulties in addressing this problem are: (1) an individual's data is always limited, which may not be enough to learn and update his preferences with some commonly used methods such as regression; (2) the observability of the data points is limited, as an individual's preferences could not be directly observed, and we could only observe the individual's choice behavior. In this dissertation, we propose a personalized control system that could interact with individuals by presenting travel alternatives to an individual and collecting his response (the choice data). The system may further learn the individual's preferences towards the influential factors with an individual preference learning algorithm and calculate the appropriate amount of incentives to provide to the individual along with a promoted alternative to increase the probability for the individual to accept the alternative.

An online experiment is designed and conducted in the dissertation to collect individual behavioral data. The respondents of the experiment are recruited from Amazon Mechanical Turk platform, and are asked to make sequential binary choices for commuting departure time. Considering the validity and reliability of the data collected from online experiment, the experiment is carefully designed in (1) a full factorial experiment with departure time choice scenarios, and (2) the design of data quality control strategies, i.e., methods that help reduce and identify low-quality data. Several data quality control checks are proposed in the dissertation, such as understanding check, response consistency check, responding time record, and social desirability scale. The experiment and the data analysis process show that these checks can successfully reduce the low-quality data in the experiment, and are very practical to be applied to other online experiments and behavioral studies. With the online experiment, a choice behavior dataset is also collected from the real world.

Concerning the difficulties in individual preference learning, the dissertation proposes a methodology of integrating a collaborative learning framework with a time-dependent model. The collaborative learning framework captures the common preference patterns underlying the heterogeneous population with data from all individuals while also identifying a unique membership vector for each individual with his observations. The collaborative learning framework helps deal with the problem of insufficient observations in individual preference learning. Also, this dissertation proposes two different time-dependent models in the integration. The first one is a single variate time-dependent polynomial model, which allows the existence of the preference changes over time. The second one is a time-invariant model that captures both the evolution of the preferences and the impacts of the attributes presented in each scenario. In other words, the first model makes the learned preferences more accurately by allowing the fluctuations in preferences when people make choices. The predicted preferences at each time step are the learned preferences of the previous step, assuming that the preferences will not change dramatically quickly. In this model, this dissertation further proposes a two-stage updating algorithm to update an individual's preferences when a new observation is available for him. The second model learns the evolution model of the preferences. The predicted preferences at each time step are calculated by the evolution model learned at the previous time step and the attributes of the given scenario at the current time step. The parameter estimation algorithm uses all the data, which means that the updating method is not real-time due to its computation time.

The dissertation first uses simulations to explore the properties of the integration of the collaborative learning framework and the two time-dependent models. Results show that the proposed models can learn individual preferences, capture the preference changes, and even provide more accurate behavior predictions than other models such as traditional logistic models and mixed-effect models. To further test the performance of the proposed models in the real world, this dissertation further designs and conducts a randomized experiment to collect data from people. In the experiment, respondents are randomly assigned to different groups and presented with different hypothetical background settings. After that, the respondents are asked to respond to a sequence of different departure time alternatives are presented to the individuals with different amounts of incentives displayed. The collected dataset is used to test the preference learning algorithms. Results reveal that the proposed methods can predict the individuals' choices in a more accurate way than other models.

7.2 Discussion

7.2.1 The simulation and the real-world case study of Time-Invariant Collaborative Model (TICM)

Time-Invariant Collaborative Model (TICM) integrates the canonical learning structure and the time-invariant model, such that the evolution function for an individual's preferences can be learned, and his preferences can be predicted given a scenario with a set of attributes. When testing the performance of TICM, we use both simulations and application to a real-world dataset. The prediction accuracy for different models in Figure 6.2 and Figure 6.3 shows that the independent logit model has better performance than the proposed TICM. However, the performance comparison when applying the models to the real-world dataset indicates that TICM has significantly higher Prediction Accuracy (PA) than other models. A possible reason for this inconsistency may be the different number of data points fed to the model in simulation and real-world application. When exploring the impact of preference dimensions and the number of canonical models in simulations (Figure 6.3 and Figure 6.2), we use 15 data points for each individual. Given the number of parameters to be estimated in independent logit model (ILM) is limited (i.e., the number of unknown parameters are just the number of dimensions, which is set to be 2, 3, 4, and 5 in the simulation tests), the data points can be enough to learn an individual's preferences. However, in real-world case study, the data points that are available are much fewer than that in the simulation – 10 in Table 6.4 and 7 in Figure 6.7. We could see that when there are only a few available data points, the advantage of the TICM may come to light: that the proposed Time-Invariant Collaborative Model can learn individual preference better than the Independent Logit Model when the available observations are limited. Another possible reason is that the data generation process in simulation is not appropriate, such that the generated data could not significantly distinguish the proposed Time-Invariant Collaborative Model (TICM) and Independent Logit Model (ILM) as the real-world dataset does. Further explorations are needed with synthetic dataset to understand the properties of TICM and its underlying mechanism.

7.2.2 Problems and Limitations of the Integrated Models

Individuals' choice-making behaviors in transportation have been well-studied for decades. Countless research has shown and verified findings in topics such as influential factors, model types, and parameter magnitudes in travel-related choices. This information is known as the domain knowledge in the transportation area, further building up the models of OLCM-T (Online Logistic Collaborative Model with Time-Varying parameters) and TICM (Time-Invariant Collaborative Model) models in this dissertation. For instance, we know well that individuals' valuation of time may change correspondingly when the duration changes or when at a different time of day [7, 175]. Thus, in TICM, we assume that an attribute's value can impact a certain dimension of the preferences. We also know that the major influential factors of the departure time decision are scheduled delays and travel time savings, and the utility function for commuting departure time choices is typically formulated as a linear function [64, 162, 119, 132, 49, 91]. There is also evidence showing that different tastes towards a certain attribute exist among the population. For example, while in general commuters would like to shorten their commuting time, some individuals who choose to take public transit may gain positive emotional experiences as they behave in a social desirable way and contribute to reduce traffic congestion and pollution [179]. With this knowledge, we can embed the logistic model with the utility function and the format of the evolution equation in the collaborative learning framework to capture the heterogeneity of the population, and apply the models to a real-world choice data on commuting departure time.

With the domain knowledge taking part in the model formulation, the integrated models can capture the heterogeneity in the population and learn individual preferences in an interpretable way. However, there would be at least two issues related to the reliance on the domain knowledge. For the first, those behaviors that are not reflected in the domain knowledge may not be captured by the proposed models, e.g., the choice-making behaviors that can not be described by the discrete choice model. For the second, in areas where domain knowledge is limited, the proposed model framework may be hard to apply. The application and generalization of the proposed model framework still needs more explorations.

Moreover, there are some other problems that remain unclear in the logistic canonical model (LCM). For the first, comparing with linear canonical model [111, 112, 113], the parameter estimation algorithm for LCM shows some instability in the mathematical property of convergence. In simulations, this is reflected not only by the longer computation time and more runs of the iteration but also by some jumps of the loss (obtained by the loss function). Figure 7.1 shows an example of the loss curve for one simulation round with 3 canonical models and 4 dimensions in preferences. The loss reaches the lowest level after around 50 to 60 iterations, but then jumps back a bit and fluctuate there till the end. More explorations may be needed to study the convergence of the model.

For the second, the uniqueness of the solution may also be in question. In the simulations, though the predicted preferences are close to the pre-set true preferences (and the predictions are accurate), the identified canonical models and the learned membership vectors can be significantly different from the true ones. Given the iteratively updating parameter estimation algorithm to the objective function, it is unknown whether the estimates would be local optimal or global optimal. Besides, another possible reason is that there exist multiple solutions to the problem (also described in [194]). Given that the utility model here is linear, it might be possible that multiple sets of solutions can satisfy the constraints, especially when there are only a few observations available for each individual.

For the third, it can be noticed that at the initialization of the optimization problem at the Q-Step in each iteration may impact the final estimates. In this dissertation, three different settings of the initial values are tested: (1) fixed values, e.g., a vector of 0; (2) the estimates obtained at the previous iteration; (3) a vector with random values. In most cases, the three settings may return the same results, while the estimates can occasionally be different. This may support that the solution returned by the estimation algorithm is the

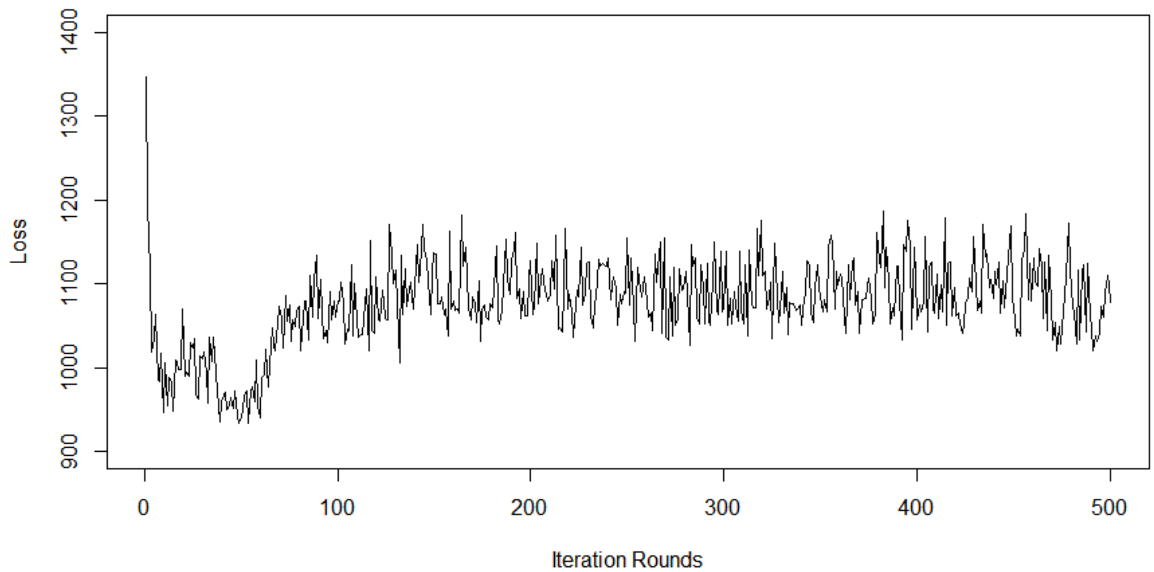


Figure 7.1: An example of a loss curve of one simulation.

local optimal. One possible way to deal with this can be to solve the optimization problem multiple times with different initialization settings. However, this may be time-consuming. In the simulation, we finally choose the second setting, i.e., the initialization is set to be the estimates of the previous iteration. If the objective function with the estimates of the previous iteration would return infinity or other incalculable values at a certain calculating step, we will give a vector of 0 for the initialization at the step.

7.2.3 Comparison of the Online Logistic Collaborative Model (OLCM-T) and the Time-Invariant Collaborative Model (TICM)

The proposed TICM differs from OLCM-T mainly in the formulation of the time-varying model $\beta(t)$. With the two different formulations, the way the model predicts the next step's preferences also differs. We will illustrate this in this section.

In the polynomial formulation of one preference dimension in OLCM-T $\beta_r(t) = q_0 + q_1 t + q_2 t^2 + q_3 t^3$, the time t is directly impacting the preference value. In the learning process, the polynomial model of t is used to capture the changes in the consecutive time steps. An example is shown in Figure 7.2: in the figure, the square points represents an individual's true preferences at each time step, while the curves with different colors represent different polynomial models that aim to learn and capture the changes of the preferences. We could see that the polynomial models enable the individual preferences to fluctuate in the sequential choices so that the estimated preferences at each time step can be closer to the true value. In other words, the polynomial model contributes to more accurate estimates at each given time step. When making predictions for the next time step, the OLCM-T model assumes that the preferences would stay unchanged, since the model could not tell us how the preference will change even we know the future choice scenarios.

The formulation of TICM, as we elaborated earlier in Chapter 6, learn the preferences in another way. It learns how the preferences react to the choice scenario, i.e., given the attributes, how the preferences would be. In other words, TICM captures the underlying mechanism of an individual's preference changes over time. Because of this, when predicting future behaviors, the learned preferences will change. The assumption in TICM is that the way the individual's preferences respond to the attributes in the choice scenario will be unchanged (i.e., the $\boldsymbol{A}$ and $\boldsymbol{B}$ are unchanged over time in Equation 6.1). Notice that in the dissertation, we also update the model at each time step with a moving window, which enables $\boldsymbol{A}$ and $\boldsymbol{B}$ to change over time. Considering this, the model together with the updating strategy makes the time-invariant model of $\beta(t)$ in TICM be a time-variate model, in which $\boldsymbol{A}$ and $\boldsymbol{B}$ become $\boldsymbol{A}_t$ and $\boldsymbol{B}_t$.

7.3 Future works

This dissertation proposes a personalized control system to learn an individual's preferences from his choice data, such that possible personalized incentives could be provided to trigger behavioral changes. In this dissertation, two models are proposed as the methodology in the personalized control system's UPDATE module. A two-stage preference updating method is proposed for model OLCM-T, with test results showing optimistic performance in both simulation and real-world dataset applications. However, it is still unclear whether the two-stage updating method can be a universal updating strategy for this kind of preference learning and updating problem where new observations are coming all the time. At least,

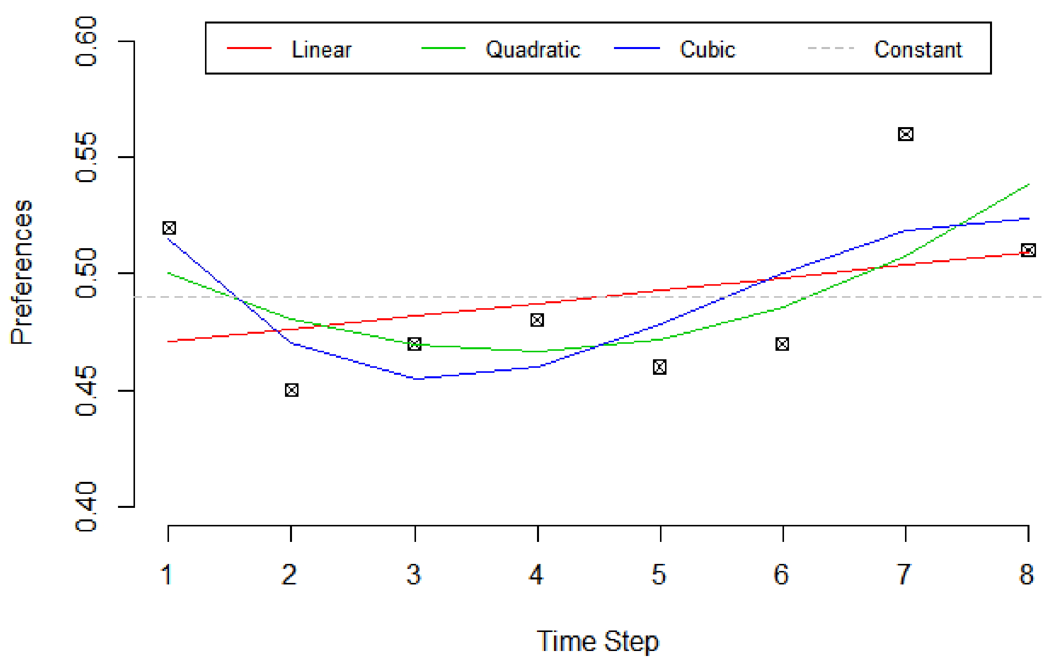


Figure 7.2: An example of fitting a polynomial curve to capture preference changes in consecutive time steps.

the two-stage updating strategy has not been adopted by TICM in this dissertation, and whether it is good to use it here is still a question. Possible designs and corresponding simulations are expected to be taken in the future to test whether TICM could also utilize this updating strategy and update each individual's preferences in a real-time way.

Another very related topic that should be included in the dissertation is the interpretation of the results from the behavior aspect. In the online experiment, the respondents are randomly assigned to different groups with different hypothetical background settings. This is to guarantee that there is underlying heterogeneity among the respondents. Since both models can identify the common patterns in the preference changes or preference evolution with canonical models, further data analysis, and interpretations from the choice-making behavioral aspect can be expected.

The personalized control system with individual preference learning algorithm aims to provide personalized incentives to trigger individuals' behavioral changes. Thus, a randomized experiment could be designed, built, and conducted to test whether personalized incentives can make more people change their behaviors. The procedure of the experiment may refer to the flowchart of the personalized control system in Chapter 3 and the design of the online experiment in Chapter 4: In the experiment, while an individual makes a selection in a scenario among alternatives, the data is immediately fed into the individual preference learning algorithm running at the backstage. The updated preferences obtained from the learning algorithm are used to calculate the appropriate amount of incentives needed by the individual in the next scenario. A control group is needed in the experiment, where the respondents are provided with a random amount of incentives.

Also, at the current stage, the personalized control system only focuses on preference learning and incentive provision at the individual level. For possible implementation in the real world, some more works are needed. For instance, one question that needs to be solved is what alternative should be proposed to each individual. To answer this question, some other systems need to be developed and integrated besides the individual preference learning system proposed in the dissertation, such as a road network with real-time traffic information, an algorithm that could identify a travel alternative (e.g., a departure time, a route) to be promoted to an individual, and a simulation platform that could predict future traffic conditions when some individuals change their behaviors. How to provide personalized incentives, on the other hand, should also consider the performance of the whole road network. For example, we also need a system or a rule to decide to whom the personalized incentives should be provided (e.g., to those with whose changes the traffic congestion could be relieved the most).

BIBLIOGRAPHY

[1] James Agarwal, Wayne S DeSarbo, Naresh K Malhotra, and Vithala R Rao. An interdisciplinary review of research in conjoint analysis: recent developments and directions for future research. *Customer Needs and Solutions*, 2(1):19–40, 2015.

[2] Erin L Allwein, Robert E Schapire, and Yoram Singer. Reducing multiclass to binary: A unifying approach for margin classifiers. *Journal of machine learning research*, 1(Dec):113–141, 2000.

[3] Foued Aloulou. The application of discrete choice models in transport. *Statistics: Growing Data Sets and Growing Demand for Statistics*, page 85, 2018.

[4] Theo A Arentze, Harmen Oppewal, and Harry JP Timmermans. A multipurpose shopping trip model to assess retail agglomeration effects. *Journal of Marketing Research*, 42(1):109–115, 2005.

[5] Kenneth J Arrow. Utilities, attitudes, choices: A review note. *Econometrica: Journal of the Econometric Society*, pages 1–23, 1958.

[6] Kenneth J Arrow. *Social choice and individual values*, volume 12. Yale university press, 2012.

[7] Javier Asensio and Anna Matas. Commuters' valuation of travel time variability. *Transportation Research Part E: Logistics and Transportation Review*, 44(6):1074–1085, 2008.

[8] Karl Johan Åström and Richard M Murray. *Feedback systems: an introduction for scientists and engineers*. Princeton university press, 2010.

[9] Bilge Atasoy, Takuro Ikeda, Xiang Song, and Moshe E Ben-Akiva. The concept and impact analysis of a flexible mobility on demand system. *Transportation Research Part C: Emerging Technologies*, 56:373–392, 2015.

[10] Carlos Lima Azevedo, Ravi Seshadri, Song Gao, Bilge Atasoy, Arun Prakash Akkinepally, Eleni Christofa, Fang Zhao, Jessika Trancik, and Moshe Ben-Akiva. Tripod: sustainable travel incentives with prediction, optimization, and personalization. In *the 97th Annual Meeting of Transportation Research Board*, 2018.

[11] Prateek Bansal, Rico Krueger, Michel Bierlaire, Ricardo A Daziano, and Taha H Rashidi. Bayesian estimation of mixed multinomial logit models: Advances and simulation-based evaluations. *Transportation Research Part B: Methodological*, 131:124–142, 2020.

[12] Richard G Baraniuk and Douglas L Jones. Signal-dependent time-frequency analysis using a radially gaussian kernel. *Signal processing*, 32(3):263–284, 1993.

[13] Felix Becker, Mazen Danaf, Xiang Song, Bilge Atasoy, and Moshe Ben-Akiva. Bayesian estimator for logit mixtures with inter-and intra-consumer heterogeneity. *Transportation Research Part B: Methodological*, 117:1–17, 2018.

[14] Gary S Becker. Irrational behavior and economic theory. *Journal of political economy*, 70(1):1–13, 1962.

[15] Prawira Fajarindra Belgiawan, Ilka Dubernet, Basil Schmid, and Kay Axhausen. Context-dependent models (crrm, murrm, prrm, ram) versus a context-free model (mnl) in transportation studies: a comprehensive comparisons for swiss and german sp and rp data sets. *Transportmetrica A: Transport Science*, 15(2):1487–1521, 2019.

[16] David A Belsley, Edwin Kuh, and Roy E Welsch. The condition number. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, 100:104, 1980.

[17] Moshe Ben-Akiva, Daniel McFadden, Kenneth Train, et al. Foundations of stated preference elicitation: Consumer behavior and choice-based conjoint analysis. *Foundations and Trends® in Econometrics*, 10(1-2):1–144, 2019.

[18] Moshe Ben-Akiva, Daniel McFadden, Kenneth Train, Joan Walker, Chandra Bhat, Michel Bierlaire, Denis Bolduc, Axel Boersch-Supan, David Brownstone, David S

Bunch, et al. Hybrid choice models: progress and challenges. *Marketing Letters*, 13(3):163–175, 2002.

[19] Moshe E Ben-Akiva, Steven R Lerman, and Steven R Lerman. *Discrete choice analysis: theory and application to travel demand*, volume 9. MIT press, 1985.

[20] Adam J Berinsky, Michele F Margolis, and Michael W Sances. Separating the shirkers from the workers? making sure respondents pay attention on self-administered surveys. *American Journal of Political Science*, 58(3):739–753, 2014.

[21] James R Bettman, Mary Frances Luce, and John W Payne. Constructive consumer choice processes. *Journal of consumer research*, 25(3):187–217, 1998.

[22] Aditya Bhargava and Grzegorz Kondrak. Language identification of names with svms. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 693–696. Association for Computational Linguistics, 2010.

[23] Chandra R Bhat, Naveen Eluru, and Rachel B Copperman. Flexible model structures for discrete choice analysis. In *Handbook of Transport Modelling: 2nd Edition*, pages 75–104. Emerald Group Publishing Limited, 2007.

[24] Chandra R Bhat and Rupali Sardesai. The impact of stop-making and travel time reliability on commute mode choice. *Transportation Research Part B: Methodological*, 40(9):709–730, 2006.

[25] Orme BK. Which conjoint method should i use? sawtooth software 2009. http://www.sawtoothsoftware.com/education/techpap.shtml. Last accessed: 2020-04-25.

[26] William S Breffle, Edward R Morey, and Jennifer A Thacher. A joint latent-class model: combining likert-scale preference statements with choice data to harvest preference heterogeneity. *Environmental and Resource Economics*, 50(1):83–110, 2011.

[27] Klaus Brinker, Johannes Fürnkranz, and Eyke Hüllermeier. Label ranking by learning pairwise preferences. Technical report, Citeseer, 2007.

[28] Michael D Buhrmester, Sanaz Talaifar, and Samuel D Gosling. An evaluation of amazon's mechanical turk, its rapid rise, and its effective use. *Perspectives on Psychological Science*, 13(2):149–154, 2018.

[29] Danny Campbell, David A Hensher, and Riccardo Scarpa. Non-attendance to attributes in environmental choice analysis: a latent class specification. *Journal of environmental planning and management*, 54(8):1061–1076, 2011.

[30] Fredrik Bagge Carlson. Machine learning and system identification for estimation in physical systems. *arXiv preprint arXiv:1906.02003*, 2019.

[31] J Douglas Carroll and Paul E Green. Psychometric methods in marketing research: Part i, conjoint analysis. *Journal of marketing Research*, 32(4):385–391, 1995.

[32] Cynthia Chen, Jingtao Ma, Yusak Susilo, Yu Liu, and Menglin Wang. The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transportation research part C: emerging technologies*, 68:285–299, 2016.

[33] Dapeng Chen, Zejian Yuan, Gang Hua, Nanning Zheng, and Jingdong Wang. Similarity learning on an explicit polynomial kernel feature map for person re-identification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1565–1573, 2015.

[34] Hao Chen, Hesham A Rakha, and Shereef Sadek. Real-time freeway traffic state prediction: A particle filter approach. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 626–631. IEEE, 2011.

[35] Siddhartha Chib and Edward Greenberg. Understanding the metropolis-hastings algorithm. *The american statistician*, 49(4):327–335, 1995.

[36] A Chiuso and G Pillonetto. System identification: A machine learning perspective. *Annual Review of Control, Robotics, and Autonomous Systems*, 2:281–304, 2019.

[37] Caspar Chorus, Sander van Cranenburgh, and Thijs Dekker. Random regret minimization for consumer choice modeling: Assessment of empirical evidence. *Journal of Business Research*, 67(11):2428–2436, 2014.

[38] Caspar G Chorus. A generalized random regret minimization model. *Transportation research part B: Methodological*, 68:224–238, 2014.

[39] Caspar G Chorus, Theo A Arentze, and Harry JP Timmermans. A random regret-minimization model of travel choice. *Transportation Research Part B: Methodological*, 42(1):1–18, 2008.

[40] Keith Chrzan and Jefferson Forkner. The random regret minimization choice modeling paradigm: An introduction with empirical tests. *Sawtooth Software, RESEARCH PAPER SERIES*, 2014.

[41] Keith Chrzan and Bryan Orme. An overview and comparison of design strategies for choice-based conjoint analysis. *Sawtooth software research paper series*, 98382, 2000.

[42] Yi-Shih Chung. Seemingly irrational driving behavior model: The effect of habit strength and anticipated affective reactions. *Accident Analysis & Prevention*, 82:79–89, 2015.

[43] Courtney Cobbs. Seattle decreased solo driving commutes, and chicago can too [february 20th 2020]. `https://chi.streetsblog.org/2020/02/20/seattle-decreased-solo-driving-commutes-and-chicago-can-too/`. Last accessed: 2020-09-01.

[44] Matthew JC Crump, John V McDonnell, and Todd M Gureckis. Evaluating amazon's mechanical turk as a tool for experimental behavioral research. *PloS one*, 8(3), 2013.

[45] Carolina Feher da Silva and Marcus Vinícius Chrysóstomo Baldo. A simple artificial life model explains irrational behavior in human decision-making. *PloS one*, 7(5):e34371, 2012.

[46] Mazen Danaf, Felix Becker, Xiang Song, Bilge Atasoy, and Moshe Ben-Akiva. Online discrete choice models: Applications in personalized recommendations. *Decision Support Systems*, 119:35–45, 2019.

[47] Peter J Danaher. Using conjoint analysis to determine the relative importance of service attributes measured in customer satisfaction surveys. *Journal of retailing*, 73(2):235–260, 1997.

[48] Samson Yusuf Dauda and Jongsu Lee. Quality of service and customer satisfaction: a conjoint analysis for the nigerian bank customers. *International Journal of Bank Marketing*, 2016.

[49] André De Palma, Asad J Khattak, and Deepak Gupta. Commuters' departure time decisions in brussels, belgium. *Transportation Research Record*, 1607(1):139–146, 1997.

[50] Rameswar Debnath and Haruhisa Takahashi. Kernel selection for the support vector machine. *IEICE transactions on information and systems*, 87(12):2903–2904, 2004.

[51] Nicolas Depraetere and Martina Vandebroek. A comparison of variational approximations for fast inference in mixed logit models. *Computational Statistics*, 32(1):93–125, 2017.

[52] Inderjit S Dhillon, Yuqiang Guan, and Brian Kulis. Kernel k-means: spectral clustering and normalized cuts. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 551–556, 2004.

[53] Reginald T Dillingham II 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.

[54] Arnaud Doucet and Adam M Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of nonlinear filtering*, 12(656-704):3, 2009.

[55] Benjamin Edelman, Michael Luca, and Dan Svirsky. Racial discrimination in the

sharing economy: Evidence from a field experiment. *American Economic Journal: Applied Economics*, 9(2):1–22, 2017.

[56] Editor. Hyphenation exception log. *TUGboat*, 7(3):145, 1986.

[57] Donald G Fischer and Carol Fick. Measuring social desirability: Short forms of the marlowe-crowne social desirability scale. *Educational and Psychological Measurement*, 53(2):417–424, 1993.

[58] Peter Frazier, David Kempe, Jon Kleinberg, and Robert Kleinberg. Incentivizing exploration. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 5–22, 2014.

[59] Christine Fricker and Nicolas Gast. Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity. *Euro journal on transportation and logistics*, 5(3):261–291, 2016.

[60] Anqi Fu, Balasubramanian Narasimhan, and Stephen Boyd. Cvxr: An r package for disciplined convex optimization. *arXiv preprint arXiv:1711.07582*, 2017.

[61] Johannes Fürnkranz and Eyke Hüllermeier. Pairwise preference learning and ranking. In *European conference on machine learning*, pages 145–156. Springer, 2003.

[62] Johannes Fürnkranz and Eyke Hüllermeier. Preference learning and ranking by pairwise comparison. In *Preference learning*, pages 65–82. Springer, 2010.

[63] Tommy Gärling and Kay W Axhausen. Introduction: Habitual travel choice. *Transportation*, 30(1):1–11, 2003.

[64] Donald P Gaver Jr. Headstart strategies for combating congestion. *Transportation Science*, 2(2):172–181, 1968.

[65] Andrew Gelman et al. Objections to bayesian statistics. *Bayesian Analysis*, 3(3):445–449, 2008.

[66] David K George and Cathy H Xia. Fleet-sizing and service availability for a vehicle rental system via closed queueing networks. *European journal of operational research*, 211(1):198–207, 2011.

[67] Genevieve Giuliano. Transportation demand management: promise or panacea? *Journal of the American Planning Association*, 58(3):327–335, 1992.

[68] Michel Goossens, Frank Mittelbach, and Alexander Samarin. *The LaTeX Companion*. Addison-Wesley, 1994.

[69] Dinesh Ambat Gopinath. Modeling heterogeneity in discrete choice processes: Application to travel demand. 1996.

[70] Samuel D Gosling, Simine Vazire, Sanjay Srivastava, and Oliver P John. Should we trust web-based studies? a comparative analysis of six preconceptions about internet questionnaires. *American psychologist*, 59(2):93, 2004.

[71] Paul E Green and V Srinivasan. Conjoint analysis in marketing research: New developments and directions. *Journal of Marketing*, 54(4):3, 1990.

[72] Paul E Green and Venkatachary Srinivasan. Conjoint analysis in consumer research: issues and outlook. *Journal of consumer research*, 5(2):103–123, 1978.

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

[74] Arthur Gretton. Introduction to rkhs, and some simple kernel algorithms. *Adv. Top. Mach. Learn. Lecture Conducted from University College London*, 16, 2013.

[75] Joseph F Hair, William C Black, Barry J Babin, Rolph E Anderson, Ronald L Tatham, et al. *Multivariate data analysis*, volume 5. Prentice hall Upper Saddle River, NJ, 1998.

[76] Juho Hamari, Mimmi Sjöklint, and Antti Ukkonen. The sharing economy: Why people participate in collaborative consumption. *Journal of the association for information science and technology*, 67(9):2047–2059, 2016.

[77] David J Hauser and Norbert Schwarz. Attentive turkers: Mturk participants perform better on online attention checks than do subject pool participants. *Behavior research methods*, 48(1):400–407, 2016.

[78] Jerry A Hausman and Daniel McFadden. Specification tests for the multinomial logit model. 1981.

[79] David A Hensher. Stated preference analysis of travel choices: the state of practice. *Transportation*, 21(2):107–133, 1994.

[80] Stephane Hess, Andrew Daly, and Richard Batley. Revisiting consistency with random utility maximisation: theory and implications for practical work. *Theory and Decision*, 84(2):181–204, 2018.

[81] Stephane Hess and Marek Giergiczny. Intra-respondent heterogeneity in a stated choice survey on wetland conservation in belarus: first steps towards creating a link with uncertainty in contingent valuation. *Environmental and Resource Economics*, 60(3):327–347, 2015.

[82] Stephane Hess and David Palma. Apollo: a flexible, powerful and customisable freeware package for choice model estimation and application. *Journal of Choice Modelling*, page 100170, 2019.

[83] Stephane Hess and Kenneth E Train. Recovery of inter-and intra-personal heterogeneity using mixed logit models. *Transportation Research Part B: Methodological*, 45(7):973–990, 2011.

[84] Christoph Hirnschall, Adish Singla, Sebastian Tschiatschek, and Andreas Krause. Learning user preferences to incentivize exploration in the sharing economy. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[85] Tom Howley and Michael G Madden. The genetic kernel support vector machine: Description and evaluation. *Artificial intelligence review*, 24(3-4):379–395, 2005.

[86] Tom Howley and Michael G Madden. An evolutionary approach to automatic kernel construction. In *International Conference on Artificial Neural Networks*, pages 417–426. Springer, 2006.

[87] Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, et al. A practical guide to support vector classification, 2003.

[88] Xianbiao Hu, Yi-Chang Chiu, and Lei Zhu. Behavior insights for an incentive-based active demand management platform. *International Journal of Transportation Science and Technology*, 4(2):119–133, 2015.

[89] Eyke Hüllermeier, Johannes Fürnkranz, Weiwei Cheng, and Klaus Brinker. Label ranking by learning pairwise preferences. *Artificial Intelligence*, 172(16-17):1897–1916, 2008.

[90] Giuseppe Iarossi. *The power of survey design: A user's guide for managing surveys, interpreting results, and influencing respondents.* The World Bank, 2006.

[91] Rong-Chang Jou, Ryuichi Kitamura, Mei-Chuan Weng, and Chih-Cheng Chen. Dynamic commuter departure time choice under uncertainty. *Transportation Research Part A: Policy and Practice*, 42(5):774–783, 2008.

[92] Richard Katzev. Car sharing: A new approach to urban transportation problems. *Analyses of Social Issues and Public Policy*, 3(1):65–86, 2003.

[93] S Sathiya Keerthi and Chih-Jen Lin. Asymptotic behaviors of support vector machines with gaussian kernel. *Neural computation*, 15(7):1667–1689, 2003.

[94] Jeremy Kees, Christopher Berry, Scot Burton, and Kim Sheehan. An analysis of data quality: Professional panels, student subject pools, and amazon's mechanical turk. *Journal of Advertising*, 46(1):141–155, 2017.

[95] Kwang In Kim, Keechul Jung, and Hang Joon Kim. Face recognition using kernel principal component analysis. *IEEE signal processing letters*, 9(2):40–42, 2002.

[96] Sun Kim, Haibin Liu, Lana Yeganova, and W John Wilbur. Extracting drug–drug interactions from literature using a rich feature-based linear kernel approach. *Journal of biomedical informatics*, 55:23–30, 2015.

[97] Ran Kivetz, Oded Netzer, and Rom Schrift. The synthesis of preference: Bridging behavioral decision research and marketing science. *Journal of Consumer Psychology*, 18(3):179–186, 2008.

[98] Ran Kivetz, Oded Netzer, and V Srinivasan. Alternative models for capturing the compromise effect. *Journal of marketing research*, 41(3):237–257, 2004.

[99] Donald E. Knuth. *The TEX book*. Addison-Wesley, 1984.

[100] Donald E. Knuth. *TEX: The Program*. Addison-Wesley, 1986.

[101] Donald E. Knuth. *Computer Modern Typefaces*. Addison-Wesley, 1986.

[102] Donald E. Knuth. *The Metafont book*. Addison-Wesley, 1986.

[103] Ilan Kremer, Yishay Mansour, and Motty Perry. Implementing the "wisdom of the crowd". *Journal of Political Economy*, 122(5):988–1012, 2014.

[104] Rico Krueger, Prateek Bansal, Michel Bierlaire, Ricardo A Daziano, and Taha H Rashidi. Variational bayesian inference for mixed logit models with unobserved inter- and intra-individual heterogeneity. *arXiv preprint arXiv:1905.00419*, 2019.

[105] Leslie Lamport. *LATEX: A Document Preparation System*. Addison-Wesley, 2nd edition, 1994.

[106] K Lancaster. Consumer demand: A new approach. 1971.

[107] Paul Felix Lazarsfeld and Neil W Henry. *Latent structure analysis*. Houghton Mifflin Co., 1968.

[108] Waiyan Leong and David A Hensher. Contrasts of relative advantage maximisation with random utility maximisation and regret minimisation. *Journal of Transport Economics and Policy (JTEP)*, 49(1):167–186, 2015.

[109] Dewei Li, Yufang Gao, Ruoyi Li, and Weiteng Zhou. Hybrid random regret minimization and random utility maximization in the context of schedule-based urban rail transit assignment. *Journal of Advanced Transportation*, 2018, 2018.

[110] Hsuan-Tien Lin and Chih-Jen Lin. A study on sigmoid kernels for svm and the training of non-psd kernels by smo-type methods. *submitted to Neural Computation*, 3:1–32, 2003.

[111] Ying Lin, Kaibo Liu, Eunshin Byon, Xiaoning Qian, and Shuai Huang. Domain-knowledge driven cognitive degradation modeling for alzheimer's disease. In *Proceedings of the 2015 SIAM International Conference on Data Mining*, pages 721–729. SIAM, 2015.

[112] Ying Lin, Kaibo Liu, Eunshin Byon, Xiaoning Qian, Shan Liu, and Shuai Huang. A collaborative learning framework for estimating many individualized regression models in a heterogeneous population. *IEEE Transactions on Reliability*, 67(1):328–341, 2018.

[113] Ying Lin, Shan Liu, and Shuai Huang. Selective sensing of a heterogeneous population of units with dynamic health conditions. *IISE Transactions*, 50(12):1076–1088, 2018.

[114] Xiaofang Liu and Chun Yang. A kernel spectral angle mapper algorithm for remote sensing image classification. In *2013 6th International Congress on Image and Signal Processing (CISP)*, volume 2, pages 814–818. IEEE, 2013.

[115] L Ljung. System identification-theory for the user 2nd edition ptr prentice-hall. *Upper Saddle River, NJ*, 1999.

[116] Pasquale Lops, Marco De Gemmis, and Giovanni Semeraro. Content-based recommender systems: State of the art and trends. In *Recommender systems handbook*, pages 73–105. Springer, 2011.

[117] Jordan Louviere and Harry Timmermans. Stated preference and choice models applied to recreation research: a review. *Leisure Sciences*, 12(1):9–32, 1990.

[118] Jordan J Louviere. Conjoint analysis modelling of stated preferences: a review of theory, methods, recent developments and external validity. *Journal of transport economics and policy*, pages 93–119, 1988.

[119] Hani S Mahmassani and Gang-Len Chang. Experiments with departure time choice dynamics of urban commuters. *Transportation Research Part B: Methodological*, 20(4):297–320, 1986.

[120] Michael R Maniaci and Ronald D Rogge. Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality*, 48:61–83, 2014.

[121] Daniel McFadden. The revealed preferences of a government bureaucracy: Theory. *The Bell Journal of Economics*, pages 401–416, 1975.

[122] Daniel McFadden. Econometric models of probabilistic choice. *Structural analysis of discrete data with econometric applications*, 198272, 1981.

[123] Daniel McFadden and Anders Karlqvist. Spatial interaction theory and planning models. *Modeling the choice of residential location*, pages 75–96, 1978.

[124] Daniel McFadden and Kenneth Train. Mixed mnl models for discrete response. *Journal of applied Econometrics*, 15(5):447–470, 2000.

[125] Craig RM McKenzie, John T Wixted, and David C Noelle. Explaining purportedly irrational behavior by modeling skepticism in task parameters: an example examining confidence in forced-choice tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(5):947, 2004.

[126] Adam W Meade and S Bartholomew Craig. Identifying careless responses in survey data. *Psychological methods*, 17(3):437, 2012.

[127] Vahid Hooshmand Moghaddam and Javad Hamidzadeh. New hermite orthogonal polynomial kernel and combined kernels in support vector machine classifier. *Pattern Recognition*, 60:921–935, 2016.

[128] Joshua L Moore, Shuo Chen, Douglas Turnbull, and Thorsten Joachims. Taste over time: The temporal dynamics of user preferences. In *ISMIR*, pages 401–406. Citeseer, 2013.

[129] Kevin P Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.

[130] Elizabeth A Necka, Stephanie Cacioppo, Greg J Norman, and John T Cacioppo. Measuring the prevalence of problematic respondent behaviors among mturk, campus, and community participants. *PloS one*, 11(6):e0157732, 2016.

[131] Anton J Nederhof. Methods of coping with social desirability bias: A review. *European journal of social psychology*, 15(3):263–280, 1985.

[132] Robert Noland and Kenneth A Small. Travel-time uncertainty, departure time choice, and the cost of morning commutes. *Transportation research record*, (1493):150–158, 1995.

[133] J Novakovic and A Veljovic. C-support vector classification: Selection of kernel and parameters in medical diagnosis. In *2011 IEEE 9th International Symposium on Intelligent Systems and Informatics*, pages 465–470. IEEE, 2011.

[134] Nadezhda Novotorova. *A conjoint analysis of consumer preferences for product attributes: the case of Illinois apples*. PhD thesis, University of Illinois at Urbana-Champaign, 2007.

[135] American Association of State Highway and Transportation Officials. Commuting in america 2013: The national report on commuting patterns and trends, 2013.

[136] Michael P O'Mahony, Pádraig Cunningham, and Barry Smyth. An assessment of machine learning techniques for review recommendation. In *Irish Conference on Artificial Intelligence and Cognitive Science*, pages 241–250. Springer, 2009.

[137] Deuk Hee Park, Hyea Kyeong Kim, Il Young Choi, and Jae Kyeong Kim. A literature review and classification of recommender systems research. *Expert systems with applications*, 39(11):10059–10072, 2012.

[138] Roger Parker. Human behavior modeling: The necessity of narrative. In *Computer Architecture in Industrial, Biomechanical and Biomedical Engineering*. IntechOpen, 2019.

[139] Julius Pfrommer, Joseph Warrington, Georg Schildbach, and Manfred Morari. Dynamic vehicle redistribution and online price incentives in shared mobility systems. *IEEE Transactions on Intelligent Transportation Systems*, 15(4):1567–1578, 2014.

[140] Kathryn A Phillips, Tara Maddala, and F Reed Johnson. Measuring preferences for health care interventions using conjoint analysis: an application to hiv testing. *Health services research*, 37(6):1681–1705, 2002.

[141] Gianluigi Pillonetto and Giuseppe De Nicolao. A new kernel-based approach for linear system identification. *Automatica*, 46(1):81–93, 2010.

[142] Gianluigi Pillonetto and Giuseppe De Nicolao. Kernel selection in linear system identification part i: A gaussian process perspective. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 4318–4325. IEEE, 2011.

[143] Gianluigi Pillonetto, Francesco Dinuzzo, Tianshi Chen, Giuseppe De Nicolao, and Lennart Ljung. Kernel methods in system identification, machine learning and function estimation: A survey. *Automatica*, 50(3):657–682, 2014.

[144] Sofia Ranchordás. Does sharing mean caring: Regulating innovation in the sharing economy. *Minn. JL Sci. & Tech.*, 16:413, 2015.

[145] Thomas A Ranney. Models of driving behavior: a review of their evolution. *Accident Analysis & Prevention*, 26(6):733–750, 1994.

[146] Michael Reed and Barry Simon. *Methods of modern mathematical physics. Volume I: Functional analysis*. Academic press, 1980.

[147] David Revelt and Kenneth Train. Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of economics and statistics*, 80(4):647–657, 1998.

[148] Ganit Richter, Daphne R Raban, and Sheizaf Rafaeli. Studying gamification: The effect of rewards and incentives on motivation. In *Gamification in education and business*, pages 21–46. Springer, 2015.

[149] Mandy Ryan and Shelley Farrar. Using conjoint analysis to elicit preferences for health care. *Bmj*, 320(7248):1530–1533, 2000.

[150] Mandy Ryan, Diane Alison Scott, C Reeves, A Bate, EM Russell, M Napper, CM Robb, et al. Eliciting public preferences for healthcare: a systematic review of techniques., 2001.

[151] Becca Savransky. Seattle ranks 2nd worst commute time in us, report finds [december 1st 2019]. https://komonews.com/news/local/spend-a-lot-of-time-commuting-seattle-ranks-2nd-worst-in-us-report-finds. Last accessed: 2020-08-10.

[152] Inc. Sawtooth Software. Aca technical paper (2007). https://www.sawtoothsoftware.com/support/technical-papers/aca-related-papers/aca-technical-paper-2007. Last accessed: 2020-04-25.

[153] Inc. Sawtooth Software. Cbc technical paper (2017). https://www.sawtoothsoftware.com/support/technical-papers/cbc-related-papers/cbc-technical-paper-2013. Last accessed: 2020-04-25.

[154] Inc. Sawtooth Software. The cbc/hb system for hierarchical bayes estimation version 5.0 technical paper. https://www.sawtoothsoftware.com/support/technical-papers/hierarchical-bayes-estimation/cbc-hb-technical-paper-2009. Last accessed: 2020-04-25.

[155] Angie Schmitt. Seattle cut car commuting downtown while adding 60,000 jobs [february 15th 2018]. https://usa.streetsblog.org/2018/02/15/seattle-cut-car-commuting-downtown-while-adding-60000-jobs/. Last accessed: 2020-09-01.

[156] Bernhard Schölkopf, Alexander J Smola, Francis Bach, et al. *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2002.

[157] Juliet B Schor and Connor J Fitzmaurice. Collaborating and connecting: the emergence of the sharing economy. In *Handbook of research on sustainable consumption*. Edward Elgar Publishing, 2015.

[158] David Schrank, Bill Eisele, Tim Lomax, and Jim Bak. 2015 urban mobility scorecard. 2015.

[159] Daniela Selloni. New forms of economies: sharing economy, collaborative consumption, peer-to-peer economy. In *CoDesign for Public-Interest Services*, pages 15–26. Springer, 2017.

[160] William Shakespeare. *Hamlet*. F.S. Crofts & Co., Inc., NY, 1946. Act I, Scene 3, Lines 70-72, are apropos.

[161] Adish Singla, Marco Santoni, Gábor Bartók, Pratik Mukerji, Moritz Meenen, and Andreas Krause. Incentivizing users for balancing bike sharing systems. In *Twenty-Ninth AAAI conference on artificial intelligence*, 2015.

[162] Kenneth A Small. The scheduling of consumer activities: work trips. *The American Economic Review*, 72(3):467–479, 1982.

[163] Kenneth A Small. A discrete choice model for ordered alternatives. *Econometrica: Journal of the Econometric Society*, pages 409–424, 1987.

[164] X Song, B Atasoy, and M Ben-Akiva. Sequential personalized menu optimization through bandit learning. Technical report, Working paper, MIT, 2018.

[165] Xiang Song, Mazen Danaf, Bilge Atasoy, and Moshe Ben-Akiva. Personalized menu optimization with preference updater: a boston case study. *Transportation Research Record*, 2672(8):599–607, 2018.

[166] Geoffrey N Soutar and Julia P Turner. Students' preferences for university: A conjoint analysis. *International journal of educational management*, 2002.

[167] César R Souza. Kernel functions for machine learning applications. *Creative Commons Attribution-Noncommercial-Share Alike*, 3:29, 2010.

[168] Spivak, M.D., Ph.D. *PCTEX Manual.* Personal TEX, Inc., CA, 1985.

[169] Spivak, M.D., Ph.D. *The Joy of TEX.* American Mathematical Society, RI, 1986.

[170] Peter R Stopher. Reducing road congestion: a reality check. *Transport Policy*, 11(2):117–131, 2004.

[171] Robert Strahan and Kathleen Carrese Gerbasi. Short, homogeneous versions of the marlow-crowne social desirability scale. *Journal of clinical psychology*, 28(2):191–193, 1972.

[172] Liang Tang, Chenfeng Xiong, and Lei Zhang. Decision tree method for modeling travel mode switching in a dynamic behavioral process. *Transportation Planning and Technology*, 38(8):833–850, 2015.

[173] Kenneth Train. *Qualitative choice analysis: Theory, econometrics, and an application to automobile demand*, volume 10. MIT press, 1986.

[174] Kenneth E Train. *Discrete choice methods with simulation.* Cambridge university press, 2009.

[175] Yin-Yen Tseng and Erik T Verhoef. Value of time by time of day: A stated-preference study. *Transportation Research Part B: Methodological*, 42(7-8):607–618, 2008.

[176] Graduate School University of Washington. Format guidelines for theses and dissertations, 2012.

[177] Sander van Cranenburgh. Advanced random regret minimization models. https://www.advancedrrmmodels.com/classical-rrm. Last accessed: 2020-01-13.

[178] Sander van Cranenburgh, Cristian Angelo Guevara, and Caspar G Chorus. New insights on random regret minimization models. *Transportation Research Part A: Policy and Practice*, 74:91–109, 2015.

[179] Mark Van Vugt, Paul AM Van Lange, and Ree M Meertens. Commuting by car or public transportation? a social dilemma analysis of travel mode judgements. *European journal of social psychology*, 26(3):373–395, 1996.

[180] Marco Vriens, Joel Huber, and Wittink Dick. The commercial use of conjoint in north america and europe: Preferences choices and self-explicated data. *Unpublished working paper*, 1997.

[181] Milena Vukic, Marija Kuzmanovic, and Milica Kostic Stankovic. Understanding the heterogeneity of generation y's preferences for travelling: A conjoint analysis approach. *International Journal of Tourism Research*, 17(5):482–491, 2015.

[182] Shenghua Wan, Tengfei Ye, Maoqing Li, Hongchao Zhang, and Xin Li. Efficient spherical parametrization using progressive optimization. In *International Conference on Computational Visual Media*, pages 170–177. Springer, 2012.

[183] Wenjian Wang, Zongben Xu, Weizhen Lu, and Xiaoyun Zhang. Determination of the spread parameter in the gaussian kernel for classification and regression. *Neurocomputing*, 55(3-4):643–663, 2003.

[184] Yanmei Wang and Yanzhu Zhang. Facial recognition based on kernel pca. In *2010 Third International Conference on Intelligent Networks and Intelligent Systems*, pages 88–91. IEEE, 2010.

[185] Ariel Waserhole, Vincent Jost, and Nadia Brauner. Pricing techniques for self regulation in vehicle sharing systems. *Electronic Notes in Discrete Mathematics*, 41:149–156, 2013.

[186] Huw CWL Williams. On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and planning A*, 9(3):285–344, 1977.

[187] La'Marcus Wingate et al. Evaluating consumer preferences for medicare part d using conjoint analysis. 2011.

[188] Chenfeng Xiong and Lei Zhang. Dynamic travel mode searching and switching analysis considering hidden model preference and behavioral decision processes. *Transportation*, 44(3):511–532, 2017.

[189] M-H Yang, Narendra Ahuja, and David Kriegman. Face recognition using kernel eigenfaces. In *Proceedings 2000 International Conference on Image Processing (Cat. No. 00CH37101)*, volume 1, pages 37–40. IEEE, 2000.

[190] Tian Youwen, Li Tianlai, and Niu Yan. The recognition of cucumber disease based on image processing and support vector machine. In *2008 congress on image and signal processing*, volume 2, pages 262–267. IEEE, 2008.

[191] Junyi Zhang. A generalized relative utility based choice model with multiple context dependencies. In *Proceedings of the International Choice Modelling Conference*, pages 3–5, 2013.

[192] Haotian Zhou and Ayelet Fishbach. The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of personality and social psychology*, 111(4):493, 2016.

[193] Bin Zhu, Lu Jiang, Yaguang Luo, and Yang Tao. Gabor feature-based apple quality inspection using kernel principal component analysis. *Journal of Food Engineering*, 81(4):741–749, 2007.

[194] Xi Zhu, Feilong Wang, Cynthia Chen, and Derek D Reed. Personalized incentives for promoting sustainable travel behaviors. *Transportation Research Part C: Emerging Technologies*, 113:314–331, 2020.

[195] Zheng Zhu, Xiqun Chen, Chenfeng Xiong, and Lei Zhang. A mixed bayesian network for two-dimensional decision modeling of departure time and mode choice. *Transportation*, 45(5):1499–1522, 2018.