Transportation Research Part C 108 (2019) 245–254

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Modeling effects of travel time reliability on mode choice using cumulative prospect theory

Sepehr Ghader, Aref Darzi, Lei Zhang*

Maryland Transportation Institute, Department of Civil and Environmental Engineering, 1173 Glenn Martin Hall, University of Maryland, College Park, MD 20742, United States

ARTICLE INFO

Keywords:
Cumulative prospect theory
Mode choice
Travel time reliability
Observational data

ABSTRACT

This paper utilizes the well-known Cumulative Prospect Theory (CPT) to study how travel mode choice is affected by travel time reliability. Prospect Theory (PT) was originally developed to model decision making under risk, and later extended to CPT, which is also applicable to decision making under uncertainty. Travel time reliability is related to uncertainties in travel time, and its effect on travel behavior can be properly modeled using CPT. PT and CPT applications in transportation are mainly limited to experimental studies. This paper introduces a CPT-based framework for mode choice modeling using observational data. The framework can help travel demand modelers utilize CPT for modeling reliability in their trip-based or activity-based models, which are typically estimated from observational data (household travel surveys). To showcase model applications, the model parameters are estimated using a combination of revealed preference household travel survey data and empirically observed travel time data for a mode choice scenario in the Washington, D.C. area. We estimated the parameters of the utility function, the parameters of the CPT value function, and the parameters of the CPT weighting function and discussed how they imply diminishing sensitivity, risk-aversion in small probability losses, and risk-seeking in high-probability losses. The paper is focused on mode choice, but its extension to other choice dimensions is discussed.

1. Introduction

Travel behavior is a very important aspect of transportation studies. This behavior mainly consists of four choice dimensions: destination choice, departure-time choice, mode choice, and route choice. For each trip, a traveler chooses where to go between all possible destinations, when to depart between different time period alternatives, which mode of transportation to select between available modes, and which route to take between their origin and destination. A comprehensive transportation demand model should be able to consider each of these choices to predict demand accurately.

Travel choice depends on the attributes of available alternatives. Variations in certain characteristics of an alternative are among the important attributes considered by travelers. Travel time reliability is related to the variability of travel time. Although travel time reliability can be correlated with travel time, it is a different phenomenon. A congested road is considered reliable if travelers experience the same amount of congestion every day. According to Bhat and Sardesai (2006), travelers consider reliability for two main reasons. First, commuters may face timing requirements and consequences associated with an early or late arrival. Second, travelers inherently feel uncomfortable with unreliability because it brings worry and pressure. This behavioral consideration has

* Corresponding author.
E-mail addresses: sghader@umd.edu (S. Ghader), adarzi@umd.edu (A. Darzi), lei@umd.edu (L. Zhang).

https://doi.org/10.1016/j.trc.2019.09.014
Received 13 April 2018; Received in revised form 23 August 2019; Accepted 20 September 2019

been noted in many studies where they observed that some travelers accept longer travel times in order to make their trip more reliable (Jackson and Jucker, 1982). A reliability-related term has become a significant part of travel models since early studies (Gaver Jr, 1968, Prashker, 1979). Many theoretical and experimental studies have considered reliability in their departure-time choice, route choice, or mode choice models, using stated preference (SP) or revealed preference (RP) surveys. While SP surveys describe a hypothetical situation for respondents, RP surveys ask about their actual experience and do not contain the usual perceptual errors found in SP surveys.

Travel choice models usually follow utility maximization theory. This theory assumes that decision makers have complete information about their alternatives and their corresponding attributes and make a rational decision to maximize their utility. In order to account for unobserved factors and measurement errors, an error term is usually added to the utility function. The assumption on the distribution of the error term identifies the probability of choosing each alternative and leads to models such as Logit and Probit (Ben-Akiva and Lerman, 1985). In the transportation literature, there are two main approaches to considering reliability within the utility maximization theory. The first is called the mean-variance approach, where a dispersion measure of travel time distribution, such as variance or standard deviation, is added to the utility function by linear summation. The second is called scheduling approach, where expected earliness penalty and lateness penalty are added to the utility function by linear summation (Carrion and Levinson, 2012).

Travel choice with alternatives that have unreliable attributes is a type of decision making under risk or uncertainty. In decision making under risk, the decision maker faces alternatives that each may lead to several outcomes. The probability of each outcome is objective and known to the decision maker. An example of decision making under risk is deciding between two gambles; in one, the gambler wins $100 with the probability of 0.5 and wins nothing with the probability of 0.5, in the other, the gambler wins $50 with the probability of 0.7 and wins nothing with the probability of 0.3. In this example, the probabilities corresponding to each gamble are objective and known to the decision maker. In decision making under uncertainty, the decision maker again faces alternatives that each may lead to several outcomes; however, the probability of each outcome is not known to the decision maker. An example of decision making under uncertainty is deciding between two new gambles; in one, the gambler wins $100 if the Dow Jones index gains at least 10 points and wins nothing otherwise, in the other, the gambler wins $50 if the Dow Jones index gains at least 5 points and wins nothing otherwise. Ellsberg (1961) showed that people behave differently under risk and uncertainty, even under different sources of uncertainty. This concept is generally known as source dependence, where decision making under risk and uncertainty should be treated differently. In the case of travel choice with alternatives that have unreliable attributes, such as a mode choice with unreliable travel times, the choice situation can be modeled as a case of decision making under risk if we assume that the traveler is aware of the travel time distribution; otherwise, the choice situation should be modeled as a case of decision making under uncertainty. We believe a traveler choosing between travel options is not aware of all possible outcomes and their corresponding probabilities; therefore, in this paper, we frame the travel choices as cases of decision making under uncertainty.

The utility maximization theory, typically used for modeling travel choices, is a normative theory, meaning that it is concerned with what a rational decision maker should do. The decision making literature also includes descriptive theories, mainly dealing with decision makers' actual behavior, beliefs, and preferences (Kahneman and Tversky, 2013). The literature shows various cases in which the normative theories fall short in explaining some aspects of decision making under risk and uncertainty (Kahneman and Tversky, 1979). The Prospect Theory (PT; see Kahneman and Tversky (1979)) is a widely used descriptive theory in psychology and behavioral economics for decision making under risk. In PT, it is assumed that in a risky condition, the decision maker chooses the alternative that has the highest prospect value. The prospect value is evaluated relative to a reference point, using a probability weighting function and a value function. The theory states that the gain or loss of each prospect is evaluated by comparison with a reference point. The role of the value function is to convert objective values to subjective ones, and the role of the weighting function is to convert probabilities to decision weights. The PT accounts for risk-aversion and diminishing sensitivity. The PT had some shortcomings, as it could not incorporate stochastic dominance and could not be extended to a large number of outcomes. Quiggin (1982) utilized rank-dependence and cumulative functions to address these limitations for decision making under risk. Schmeidler (1989) also utilized rank-dependence to address the limitations for decision-making under uncertainty. Tversky and Kahneman (1992) utilized the rank dependence to extend PT to uncertain prospects as well as risky prospects with a large number of outcomes, and introduced Cumulative Prospect Theory (CPT). In this paper, we utilize CPT as it is applicable to decision making under uncertainty.

PT and CPT have been utilized in some previous transportation studies to examine the effect of uncertainty or reliability on travel behavior. The previous reliability studies using PT and CPT often involve data from controlled experiments, which tend to be costly to collect and suffer from relatively smaller sample sizes compared to observational data. They were mostly departure-time or route choice experiments focusing on a choice between multiple given routes or departure-time alternatives. The reference points and value and weighting function parameters were mainly set to fixed values. Nobody can question the importance of the studies using experimental data in the literature, but the lack of a study focusing on PT or CPT with observational data is evident in the travel time reliability literature. Most of the travel demand models, whether trip-based or activity-based, are estimated from observational data such as revealed-preference household travel surveys. The focus of this paper is to introduce a framework to model reliability based on CPT using observational data. This paper proposes a general framework that uses revealed-preference survey data and travel time data to model the effect of perceived reliability on mode choice, based on CPT. It also estimates value and weighting function parameters. The proposed framework's data requirement is similar to the requirement of the mean-variance model, which is widely used for studying travel time reliability.

To showcase the applications of the proposed model, we describe and study a mode choice scenario between driving and rail in the Washington, D.C. area. The CPT-based model parameters are estimated using INRIX travel time data and 2007–2008

Transportation Planning Board (TPB)-Baltimore Metropolitan Council (BMC) household travel survey data. The estimated model is presented and the meaning of the estimated parameters is also discussed.

This paper continues with a review of transportation literature on applications of PT and CPT in the travel time reliability literature. The modeling framework that utilizes CPT to model the effect of perceived travel time reliability on mode choice behavior is presented afterward. The modeling framework section is followed by a section describing a model application example in which the estimation procedure and the results are described. The last sections present some further discussion and concluding remarks.

2. Literature review

The transportation literature has identified reliability as one of the key factors in travel behavior. Many studies have focused on the impacts of travel time reliability on travel choices (Bates et al., 2001, Carrion and Levinson, 2013, Small, 1999). Travel time reliability is usually related to the day-to-day variation of travel time for a certain link, route, or network. The less day-to-day variability in travel time is related to higher reliability. Noland and Polak (2002), Carrion and Levinson (2012), and Li et al. (2010) have reviewed the travel time reliability literature that is mainly based on Random Utility Maximization (RUM) framework. Within this framework, mean-variance models and scheduling models are the main approaches to consider reliability. Li et al. (2016) developed a framework by integrating the mean-variance and scheduling approaches in order to investigate the schedule delay and trip time variability, with the main focus on a more in-depth representation of uncertainty. In their study, they used two stated preference datasets from China and Australia.

Kahneman and Tversky (1979) introduced PT to model decision making under risk. While expected utility theory (EU) assumes that decision makers evaluate alternatives by their utility, which is the sum of the product of probabilities and outcome values, PT assumes outcomes are converted to subjective values using a value function, and probabilities are converted to decision weights using a weighting function. PT assumes that decision makers evaluate alternatives (prospects) by comparing the sum of products of converted values and decision weights and choose the alternative that has the highest prospect value. PT models have the following differences relative to the normative EU models:

- They assume that decision makers reframe or edit the prospects as gains and losses relative to a reference point, and evaluate the reframed prospects.
- They assume different value functions for gains and losses.
- They assume a concave function for gains and a convex function for losses which results in diminishing sensitivity.
- They assume disutility of a loss is greater than the utility of the same amount of gain, which results in loss aversion.
- They assume that decisions are affected by the decision weights, not actual probabilities. Decision weights and actual probabilities are related through a probability weighting function.

In 1992, Tversky and Kahneman (1992) used the rank-dependent utility theory proposed by Quiggin (1982) and extended their model to CPT. In CPT, probability weights are influenced by the preference rank of an outcome. CPT is also capable of modeling decision making under uncertainty.

In transportation literature, a number of studies have used PT to explain travel behavior (Li and Hensher, 2011). De Palma et al. (2008) described the theory of integrating risk and uncertainty in discrete choice models and utilizing non-expected-utility-based models within the discrete choice framework. Avineri and Prashker (2005) conducted one of the earliest studies that used CPT to find the market share between two hypothetical unlabeled routes. They drew travel times and probabilities from hypothetical distributions, assumed the reference point to be 31.5 min, collected SP data from the respondents, and used parameters estimated by Tversky and Kahneman (1992) to estimate prospect value and market share of each route. In a numerical example, Avineri (2004) used PT to study the choice between two bus lines with different headway distributions. He also used parameters estimated by Tversky and Kahneman (1992) and assumed different reference points to study the sensitivity of prospect values to the reference point. Gao et al. (2010) used simulated choice data and parameters estimated by Tversky and Kahneman (1992) to compare expected utility theory (EUT) and PT for route choice in a risky network. They estimated their models within the Multinomial Logit framework and suggested that CPT performs better than EUT. Michea and Polak (2006) studied train travelers' behavior using different models such as CPT. They used SP data and showed that a CPT model gives a better fit compared to an EUT model.

Senbil and Kitamura (2004) conducted a survey asking respondents about three consecutive commuting days, using PT in the departure-time choice context. They showed that this travel behavior dimension is consistent with PT assumptions. Connors and Sumalee (2009) modeled network equilibrium on a hypothetical network using PT. They assumed the parameters of weighting and value functions are fixed to certain values and studied the sensitivity of their model to the reference point. Xu et al. (2011) utilized CPT to introduce a decision-making rule for route choice behavior. They explained how the reference point value can be set in the route choice setting, and used designed experiment and survey questionnaires to estimate some of the value and probability weighting function parameters. Jou and Chen (2013) applied CPT to study drivers' attitude toward risk in the route choice setting using a computer-aided survey in Taiwan, designed to capture drivers risk attitude. Yang and Jiang (2014) is another example of route choice modeling based on CPT. They improved the assumption on the reference travel times by considering traveler's reserved travel time restrictions. They showed that compared to EUT, CPT can better explain route choice behavior under risk. Zhou et al. (2014) showed another application of PT in route choice by studying travel behavior under dynamic message signs. Chen et al. (2017) investigated the impacts of travel time variability specific to train station choice for park-and-ride users, using a framework based on discrete choice theory, CPT, and mean-variance approach. The dataset used in this study was a combination of a stated preference experiment and revealed preference data for obtaining the travel time of travelers to the station. Their result showed that the respondents had risk-averse attitudes toward the travel time variability. Huang et al. (2017) utilized CPT to study the degree of risk-averseness in choosing managed lanes using data collected from a designed internet-based SP survey, and compared the CPT model with a utility-based model. Zhang et al. (2018) is a recent route choice study focused on CPT. They considered the effects of social interaction among friends on route choice behavior. They used a laboratory experiment to collect data.

We see decision making under unreliable travel time as a case of decision making under uncertainty, for which CPT offers a descriptive decision-making theory. The review of transportation literature revealed some valuable studies incorporating PT (or CPT) in travel behavior, but most of the studies were in an experimental setting. They had to collect new surveys for the experiment, which tends to be costly and suffers from relatively smaller sample sizes, and used limited assumptions on the reference point, in addition to the value and weighting function parameters. In addition, focus on mode choice was very limited. This paper proposes a CPT-based mode choice framework that utilizes household travel survey data and empirically observed INRIX travel time data to study the effects of unreliability on travel behavior. This is valuable, as the framework can be applied in many of the existing travel demand models to incorporate reliability without requiring any additional experiments or surveys. The modeling framework is general, and can be extended to other types of uncertainty (cost uncertainty) and other choice dimensions. The possible extensions of the model are outlined in the discussion section.

3. Methodology

As discussed in the introduction, we consider the effect of reliability on travel choices as a case of decision making under uncertainty and apply CPT. CPT argues that decision makers perceive values (travel time in the case of this paper) differently from the actual values. CPT also argues that probabilities (probability of experiencing a certain travel time for a given trip) are converted to decision weights. Assume a decision maker wants to choose between M alternatives (choice between driving and rail). Choosing each alternative can lead to k different outcomes, quantified by $x_1, x_2, \cdots, x_k$ (driving travel time can be $t_1, t_2, \cdots, t_k$). The outcome i can happen with probability p_i. Each alternative can be seen as a prospect (x_i, p_i). In a typical expected utility maximization model, it is assumed that the decision maker chooses the alternative that has the highest expected value: $\sum_i x_i * p_i$. The PT argues that the outcome values (x_i) are first reframed into gains and losses with respect to a reference point and then converted to perceived values by a value function $(v(x_i))$. The probabilities are also converted to decision weights by a probability weighting function $(w(p_i))$. PT states that the decision maker chooses the prospect of the highest value based on $V = \sum_i v(x_i) * w(p_i)$.

The main difference between PT and CPT is in the decision weights $(w(p_i))$. PT works with marginal probabilities, while CPT uses cumulative probabilities. In CPT, it is assumed that the outcomes of each prospect are first arranged in increasing order. Therefore, if $i > j$, then $x_i > x_j$. A prospect can be positive (all outcomes are gains), negative (all outcomes are losses), or mixed (outcomes are a combination of gains and losses). Assume that outcomes are sorted from the biggest loss to the biggest gain with positive indices for gains $(1, 2, \cdots, k^+)$ and negative indices for loss $(-k^-, \cdots, -2, -1)$. In CPT, the value of a prospect f is calculated by:

$$V(f) = V(f^+) + V(f^-)$$

$$V(f^+) = \sum_{i=1}^{k^+} \pi_i^+ * v(x_i);\ V(f^-) = \sum_{i=-1}^{-k^-} \pi_i^- * v(x_i); \tag{1}$$

The π's are calculated from the cumulative probabilities and the probability weighting function by:

$$\pi_{k^+}^+ = w(p_{k^+})\pi_{k^-}^- = w(p_{-k^-}),$$

$$\pi_i^+ = w(p_i + \cdots + p_{k^+}) - w(p_{i+1} + \cdots + p_{k^+}),\ 0 < i < k^+$$

$$\pi_i^- = w(p_{-k^-} + \cdots + p_{-i}) - w(p_{-k^-} + \cdots + p_{-i-1}),\ -k^- < i < 0 \tag{2}$$

Tversky and Kahneman (1992) provided parametric formulas for the value and weighting functions to formulate the prospect theory. The following value function suggested by Tversky and Kahneman is used in this paper:

$$v(x) = x^\alpha \quad x \geq 0 \tag{3}$$

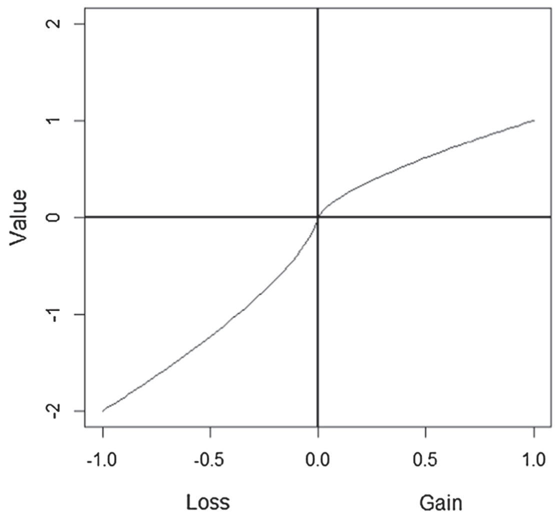
$$v(x) = (-\lambda)(-x)^\beta \quad x < 0 \tag{4}$$

In these equations, α is the exponent of the value function over the gain region, and β is the exponent of the value function over the loss region. λ indicates the degree of loss aversion.

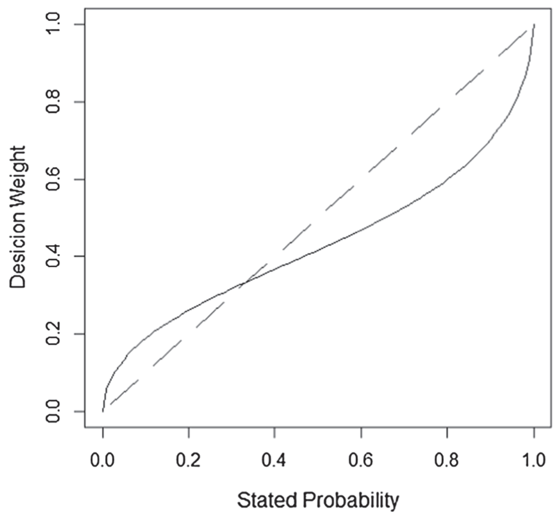
The following probability weighting function suggested by Tversky and Kahneman (1992) is used in this paper:

$$w(p) = \frac{(p)^\gamma}{[(p)^\gamma + (1-p)^\gamma]^{\frac{1}{\gamma}}} \tag{5}$$

where $w(p)$ is the decision weight corresponding to probability p, and γ is the probability weighting parameter. Fig. 1a and b show examples of a value function and a weighting function, respectively. These figures follow the examples provided by Tversky and Kahneman (1992). The combination of the value function and the weighting function leads to risk-seeking behavior for low-probability gains and high-probability losses, and risk-averse behavior for high-probability gains and low-probability losses.



a. Example of a value function with $\alpha = 0.7$, $\beta = 0.7, \lambda = 2$



b. Example of a weighting function with $\gamma = 0.6$

Fig. 1. Examples of the value and the weighting function.

In order to apply the CPT theory for mode choice, assume a traveler wants to decide between M modes $m_1,\ m_2,\ \cdots, m_M$. Travel time reliability can be modeled by considering variations in the value of travel time using CPT. For mode alternative m, the travel time can be $t_{m,1},\ t_{m,2},\ \cdots, t_{m,k}$ with probabilities $p_{m,1},\ p_{m,2},\ \cdots, p_{m,k}$, respectively. There may be other certain factors, such as travel cost, that affect the decision-making process. For the sake of simplicity, this study only considers cost. Other factors such as comfort can be also considered. The corresponding prospect for mode m is:

$$f_m = (\beta_m + \beta_{cost} * Cost_m + \beta_{time} * t_{m,i},\ p_i),\ i = 1,\ \cdots, k \tag{6}$$

In which, β_m is the mode-specific constant added to capture the effect of unforeseen variables (β_m should be set to zero for one of the modes for identification), β_{cost} is the coefficient of cost, and β_{time} is the coefficient of time. The prospect can be formed for each mode and CPT can be applied to calculate the value of each mode prospect ($V(f_m)$) as well as model decision making between the M prospects.

The following items must be assumed in order to apply CPT:

- The reference point
- The form of the value function

- The form of the probability weighting function

As both cost and travel time are disutility, we model all the outcomes as losses and set the reference point at zero. Eq. (4), with λ set to 1 (since we only model in the loss domain) is assumed to be the value function. Eq. (5) is assumed to be the form of the probability weighting function. One major advantage of the PT and CPT is that they can account for differences between gains and losses and model loss aversion. Defining all outcomes in the loss domain is a limitation of the model introduced in this section. In the discussion section, we discuss the capability of the framework to be applied in the gain domain by adding other attributes with a positive effect such as comfort.

The model parameters that must be estimated are β_m for $m = 2, \cdots, M$, β_{cost}, β_{time}, β (parameter of the value function), and γ (parameter of the probability weighting function). The widely used method in the literature for formulating the probability of choosing mode m is to use the logit formulation (Huang et al., 2017; Jou and Chen, 2013; Zhou et al., 2014). Once $V(f_m)$ is calculated for all the modes using CPT, the probability of choosing mode m can be written by the logit formulation:

$$p(m) = \frac{\exp(V(f_m))}{\sum_i \exp(V(f_i))} \tag{7}$$

The model estimation can be completed using the Maximum Likelihood Estimation, or MLE (Cramer, 1986). The proposed CPT-based model requires the same data as a typical mean-variance model. The mean-variance model requires data on travelers' choice, such as travel survey data and travel time variation of each mode to calculate means and variances. The proposed model does not require anything further. Data from household travel surveys that are publicly available can be combined with historical travel time observations for the estimation of the parameters in the proposed model. Historical travel time observations from probe vehicles, GPS, and other sources are available in many areas.

The next section of the paper uses the 2007–2008 Transportation Planning Board (TPB) and Baltimore Metropolitan Council (BMC) Household Travel Survey and INRIX travel time data to showcase how the proposed model can be applied in a real-world situation, and how the estimated parameters can be interpreted.

4. Empirical application

Drivers tend to dislike high travel time variations resulting from accidents, bad weather, roadwork, fluctuation in demand, etc. Compared to driving, rail usually has more reliable travel times since it follows a fixed schedule of operation. It would be interesting to explore how this difference in travel time reliability would affect travelers' choice between these two modes. The proposed CPT-based model is applied to this mode choice scenario, the parameters are estimated, and their implications are discussed.

4.1. Data

Two datasets were used in this application: travel time observations from INRIX and the 2007-2008 TPB-BMC Household Travel Survey. INRIX travel times were used to obtain origin-destination (OD)-level travel time observations, and 2007–2008 TPB-BMC Household Travel Survey data was used to provide information on observed trips and their chosen mode. As the travel survey does not include information on travel cost, skim matrices from the Washington, D.C. travel demand model were used to obtain cost between the origin-destination pairs.

INRIX Travel Time Data: INRIX collects road speed information from millions of mobile phones, connected cars, trucks, delivery vans, and other fleet vehicles with GPS devices, and provides these speed and travel time data to users. Traffic Message Channels (TMC) location codes are used in INRIX data as location indices. Each road segment in the INRIX network has one unique TMC location code. INRIX can provide minute-by-minute travel time and speed information throughout the day for selected road segments, which can be used to calculate travel time and travel time reliability for any time-of-day for a specific road by using variation among different days. Kim and Coifman (2014) validated the quality of INRIX data by comparing it with loop detector data. In this study, INRIX travel time data was used to obtain travel time observations between selected OD pairs.

2007-2008 TPB-BMC Household Travel Survey: This survey was conducted by the Transportation Planning Board (TPB) from February 2007 to April 2008 to gather information about demographics and socioeconomic and trip-making characteristics of residents in the Washington and Baltimore metropolitan areas. 14,000 households (about 31,000 persons) participated in this survey, and the data was geocoded at the traffic analysis zone (TAZ) level. The data contains four major components: household data, person data, vehicle data, and trip data.

In this application example, OD pairs in the Washington, D.C. area that have both rail and driving trips in the travel survey were selected and studied. In these OD pairs, both travel modes are available and are competing with each other. In total, there were 160 OD pairs that contained both rail and driving trip records and that had available INRIX data. The reported travels between these 160 OD pairs formed a major component of this application example. In these 160 OD pairs, 179 rail trips, 193 driving trips, and 37 trips using other travel modes were observed, as shown in Table 1.

Due to the method of OD pair selection, rail and driving are the major travel options in the selected OD pairs. Because of the few observations for trips made with other modes and for the sake of simplicity, it was assumed that rail and driving are the only available alternatives for travelers. Therefore, only driving and rail travel modes were considered in the mode choice model. 179 rail trips and 193 driving trips (372 in total) formed the observations in the mode choice problem. We acknowledge that considering other modes

Table 1
Trip records in the 160 OD pairs from the travel survey.

Travel Mode	Rail	Driving	Other	Sum
Number of Trips	179	193	37	409
Percent	43.8%	47.2%	9.0%	100.0%

in our model can lead to improvements; however, the purpose of this application example is not to build the best mode choice model for the study area. The purpose of presenting the empirical application example is to illustrate how the modeling framework can be applied to real world observational data.

The INRIX travel time data between the origin and the destination at the time of departure were collected for the 372 survey observations. One year of INRIX data corresponding to the year 2012 was obtained, and day-to-day travel time variation (after removing outliers) at the time of departure was used to get travel time distribution for each observation. It should be noted that 2012 was the year closest to 2008 with the proper coverage on the studied OD pairs. The inconsistency between the survey year and the travel time year is one of the limitations of this application example. TMC level INRIX data had to be processed to obtain OD-level travel time observations. The process for obtaining OD-level travel time observations from TMC-based data is explained in previous papers (Mishra et al., 2017, Tang et al., 2015). Fig. 2 summarizes the obtained INIRX travel time data for observations with a travel time standard deviation higher than five minutes. The x-axis corresponds to different observations, and the Y-axis shows the INRIX travel time variation. The dots show the mean travel times, and the intervals correspond to the mean ± standard deviation.

INRIX data was used for the travel time distribution of the driving mode. INRIX does not have information on rail travel time. Due to the method used to select OD pairs, they all had rail and driving trip records in the survey. Therefore, the travel time of the rail mode for each OD pair was obtained from the survey data. The mean travel time for the rail mode was calculated by averaging the reported travel time of all rail trips between the origin and the destination. Rail was assumed to be very reliable, meaning that rail travel time between each OD pair was assumed to be constant and equal to the average reported rail travel time between the OD pair. Although rail trips are usually reliable and follow their schedule, assuming them to be perfectly reliable is another limitation of this application example. Data on rail reliability can certainly improve this example.

4.2. Mode choice model estimation results

For the CPT-based model, we should first define the travel time outcomes. These outcomes were obtained from the travel time distribution (INRIX data) by dividing the observed travel times into one-minute intervals. Each one-minute interval can be seen as an outcome, and its corresponding probability can be obtained from the travel time distribution. An example can be seen in Fig. 3.

Once outcomes are defined, the prospects for driving and rail can be written as:

$$f_{driving} = (\beta_{cost} * Cost_{driving} + \beta_{time} * t_{driving,i}, p_i), \ i \in outcomes \tag{8}$$

$$f_{rail} = (\beta_{rail} + \beta_{cost} * Cost_{rail} + \beta_{time} * t_{rail}, 1) \tag{9}$$

The notations for these prospects were previously described in the methodology section. It should be noted again that the rail travel time is assumed to be fixed, therefore only one outcome with probability of one is assumed. The CPT formulation can now be applied to the two prospects as described in the methodology section.

MLE was used to estimate model parameters. The MLE was coded in R (RCore, 2012). Table 2 summarizes the estimation results of the CPT-based model.

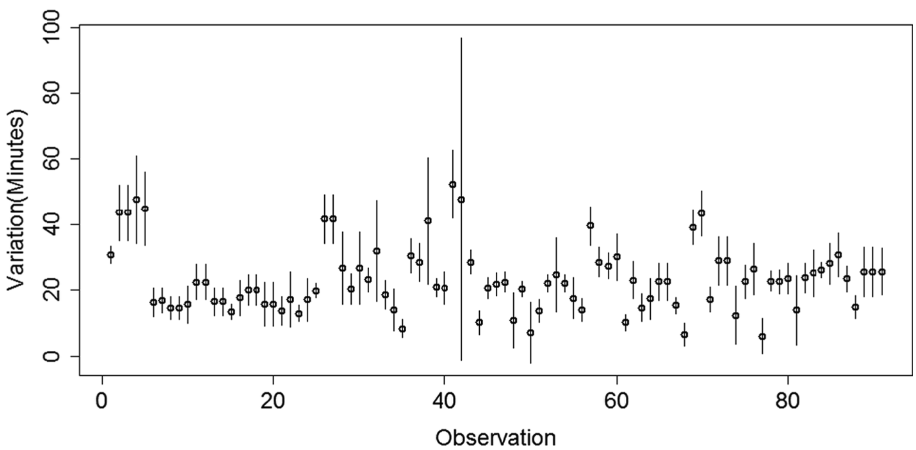


Fig. 2. INRIX travel time variation between selected OD pairs.

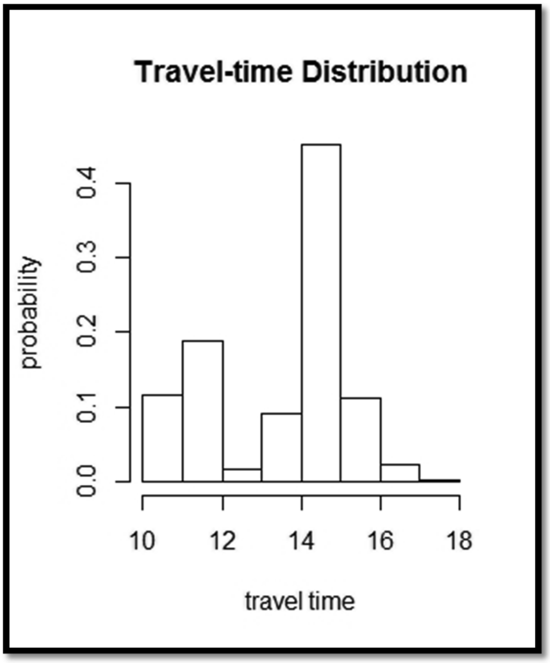


Fig. 3. Example of travel time distribution divided into one-minute outcomes.

Table 2
Estimation results for the CPT-based model.

Parameter	Coefficient	Std. Err.	t stat
β_{cost}	0.022	0.009	2.44
β_{time}	0.008	0.000	280
ASC_{rail}	−0.522	0.14	3.72
β	0.863	0.008	17.12
γ	0.643	0.113	3.15
Log Likelihood	−257.850		

The following items should be discussed about the estimation results:

- The estimated coefficients for cost and travel time were both positive. This means their effects are negative, as they are used with a negative sign in the value function over the loss domain (please refer to Eq. (4)). They are also statistically significant with $\alpha = 0.05$. The results are consistent with the expectation, because higher cost and higher travel time are both undesired. This shows that travelers prefer to experience lower travel times and lower travel costs.
- The value function adds diminishing sensitivity to the model. It also contributes to attitude toward risk in combination with the probability weighting function. The value of β must be between 0 and 1 for the value function to comply with the diminishing sensitivity property. A lower β value corresponds to more extreme diminishing sensitivity. The median value for β in Tversky and Kahneman's original CPT paper was 0.88. Later studies also showed a value of β between 0 and 1. Here, the estimated β is 0.86, consistent with the literature, and in accordance with diminishing sensitivity.
- The value of γ should be positive by definition. A negative γ will result in a decreasing function, while the probability weighting function needs to increasing. A γ smaller than 1 will result in a weighting function that increases low probabilities and decreases high probabilities. A γ value bigger than 1 has the reverse effect. The median value for γ in Tversky and Kahneman's original CPT paper was 0.61. Later studies also showed a value of γ between 0 and 1. Here, the estimated γ is 0.64, consistent with the literature. In combination with the estimated convex value function in the loss domain, this value of γ leads to risk-seeking behavior in high-probability losses and risk-averse behavior in low-probability losses. This value also contributes to diminishing sensitivity in probability weighting function, meaning that the effect of a given change in probability gets smaller as we move further away from probability bounds, zero and one (convex weighting function near one and concave weighting function near zero).
- The value of time obtained from the estimated results is about \$21.00 per hour, reasonable for the Washington, D.C. area.

5. Discussion

5.1. Extension to other choice dimensions

This paper was focused on mode choice, but can be extended to other choice dimensions. In the case of mode choice studied in this paper, we defined the prospect of each mode considering the uncertainties in travel time. Similarly, one can define the prospect of each time-of-day alternative considering uncertainties in travel time at each time-of-day for time-of-day modeling, or define the prospect of each destination considering the uncertainties in travel time to each destination for destination choice modeling. The described time-of-day choice modeling would require the distribution of travel time at each time-of-day and the described destination-choice modeling would require the distribution of travel time between the origin and each destination. The same data source used in this paper, INRIX travel time data, or any other sort of travel time data can be used to obtain the required travel time distributions. The extension to other choice dimensions should be further studied.

5.2. Extension to other sorts of travel uncertainties

This paper was focused on travel time uncertainty, but the methodology can be extended to other sorts of uncertainty, such as travel cost uncertainty. We defined the prospect of each mode based on uncertainties in travel time. If the travel cost is also uncertain, and the data on cost distribution is available, one can also consider cost uncertainty in defining the prospect of each mode. This paper was focused on travel time uncertainty because travel time uncertainty is the most widely-used type of uncertainty in transportation reliability studies.

5.3. Gain domain

While the prospect examples used in the paper were only defined in the loss domain, the general framework is not limited to the loss domain. As travel time and cost are both disutility, we would only have loss domain in a typical mode choice model, unless attributes such as comfort and mode incentives are added to the model. In Tversky and Kahneman (1992), they talk about positive, negative, and mixed prospects, and discuss how the theory can be applied in each case. The application of our framework to mixed or positive prospect is possible, but should be further studied. This issue has been previously discussed for network user equilibrium models (Avineri, 2006, Wang et al., 2019).

5.4. Sensitivity to reference point

One important assumption of PT and CPT is the reference point, which divides the outcome spectrum into loss and gain domains. The reference point can be different for each individual based on their information, experience, etc. In transportation literature, Avineri (2004) have previously shown that CPT predictions are sensitive to the reference point assumption. Avineri and Bovy (2008) suggested three approaches for setting the reference point. We assumed that the reference point is zero for all individuals, which makes the prospect values closely related to the utilities in a utility-based model. Another approach is to use dynamic reference points, similar to the research by Yang et al. (2017) that used dynamic reference points for a dynamic mode choice study. An interesting next step would be studying the reference point assumption and evaluating other approaches for setting the reference points.

5.5. Dividing travel time observations into a discrete set of outcomes

In order to define outcomes from the travel time observations, we used one-minute intervals, as seen in Fig. 3. This discretization may affect the model estimates and performance because of non-linearity in the value and weighting functions. Another interesting next step would be studying the sensitivity of the model to changes in the discretization and approximation scheme used in defining outcomes.

6. Summary and conclusion

This paper introduced a methodology to consider the effect of travel time reliability on mode choice behavior using the cumulative prospect theory (CPT). We considered mode choice with an unreliable travel time attribute as a case of decision making under uncertainty. The literature offers many examples of situations where normative models such as the utility maximization fail to explain some of the decision-making aspects under risk and uncertainty. CPT was introduced in the fields of psychology and behavioral economics to model decision making under risk and uncertainty. CPT is capable of explaining phenomena such as risk-aversion, diminishing sensitivity, and certainty effect. The applications of PT and CPT in transportation are mainly limited to experimental studies. These studies involve designing experiments and collecting experimental data, which is usually costly. We introduced a reliability modeling framework based on CPT that can be applied to observational data, which means researchers and practitioners can apply the framework to their existing data and incorporate reliability into their models with a behaviorally realistic descriptive method. As an example of model application, the proposed model was estimated for a mode choice problem between unreliable driving and reliable rail using a combination of revealed preference survey data and empirically observed travel time data in the Washington, D.C. area. The estimated parameters for the CPT-based model were consistent with the previous findings in the literature and showed diminishing sensitivity, risk-aversion in small probability losses, and risk-seeking in high-probability losses. Many of the previous PT-related papers in transportation assumed PT parameters (parameters of the value and the weighting functions) as fixed, equal to the values originally estimated by Tversky and Kahneman (1992). One could question the validity of this assumption, as the context in which Tversky and Kahneman estimated their values was very different from the context of transportation studies. However, the estimation results of this paper, which are based on empirically observed data for mode choice modeling, show that the estimates of the PT parameters were not significantly different from the original estimates of Tversky and Kahneman. This should be further studied for other choice dimensions, as this might not be the case for the other dimensions.

Acknowledgments

This research was partly supported by the Maryland Transportation Institute. Any opinions, findings, conclusions, or recommendations expressed are those of the authors and do not necessarily reflect the views of the Institute. We also thank Dr. James Grand for his guidance on the paper and Dr. Liang Tang for his help in the data processing.

References

Avineri, E., 2004. A cumulative prospect theory approach to passengers behavior modeling: waiting time paradox revisited. Intelligent Transportation Systems. Taylor & Francis.
Avineri, E., 2006. The effect of reference point on stochastic network equilibrium. Transport. Sci. 40 (4), 409–420.
Avineri, E., Bovy, P.H., 2008. Identification of parameters for a prospect theory model for travel choice analysis. Transport. Res. Rec. 2082 (1), 141–147.
Avineri, E., Prashker, J.N., 2005. Sensitivity to travel time variability: travelers' learning perspective. Transport. Res. C: Emerg. Technol. 13 (2), 157–183.
Bates, J., Polak, J., Jones, P., Cook, A., 2001. The valuation of reliability for personal travel. Transport. Res. Part E: Logist. Transport. Rev. 37 (2), 191–229.
Ben-Akiva, M.E., Lerman, S.R., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press.
Bhat, C.R., Sardesai, R., 2006. The impact of stop-making and travel time reliability on commute mode choice. Transport. Res. B: Methodol. 40 (9), 709–730.
Carrion, C., Levinson, D., 2012. Value of travel time reliability: a review of current evidence. Transport. Res. A: Pol. Pract. 46 (4), 720–741.
Carrion, C., Levinson, D., 2013. Valuation of travel time reliability from a GPS-based experimental design. Transport. Res. C: Emerg. Technol. 35, 305–323.
Chen, C., Xia, J.C., Smith, B., Olaru, D., Taplin, J., Han, R.J.T.r.p., 2017. Influence of travel time variability on train station choice for park-and-rider users. Vol. 25, pp. 2473–2489.
Connors, R.D., Sumalee, A., 2009. A network equilibrium model with travellers' perception of stochastic travel times. Transport. Res. B: Methodol. 43 (6), 614–624.
Cramer, J., 1986. Econometric Applications of Maximum Likelihood Methods Cambridge University Press, New York, New York.
De Palma, A., Ben-Akiva, M., Brownstone, D., Holt, C., Magnac, T., McFadden, D., Moffatt, P., Picard, N., Train, K., Wakker, P., 2008. Risk, uncertainty and discrete choice models. Market. Lett. 19 (3–4), 269–285.
Ellsberg, D., 1961. Risk, ambiguity, and the savage axioms. Quart. J. Econ., 643–669.
Gao, S., Frejinger, E., Ben-Akiva, M., 2010. Adaptive route choices in risky traffic networks: a prospect theory approach. Transport. Res. C: Emerg. Technol. 18(5), 727–740.
Gaver Jr, D.P., 1968. Headstart strategies for combating congestion. Transport. Sci. 2 (2), 172–181.
Huang, C., Burris, M., Shaw, W.D., 2017. Models of transportation choice with risk: an application to managed lanes. Transport. Lett. 9 (5), 276–295.
Jackson, W.B., Jucker, J.V., 1982. An empirical study of travel time variability and travel choice behavior. Transport. Sci. 16 (4), 460–475.
Jou, R.-C., Chen, K.-H., 2013. An application of cumulative prospect theory to freeway drivers' route choice behaviours. Transport. Res. A: Pol. Pract. 49, 123–131.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica: J. Econom. Soc. 263–291.
Kahneman, D., Tversky, A., 2013. Choices, values, and frames. Handbook of the Fundamentals of Financial Decision Making: Part I. World Scientific.
Kim, S., Coifman, B., 2014. Comparing Inrix speed data against concurrent loop detector stations over several months. Transport. Res. C: Emerg. Technol. 49, 59–72.
Li, H., Tu, H., Hensher, D., 2016. Integrating the mean-variance and scheduling approaches to allow for schedule delay and trip time variability under uncertainty. Transport. Res. A: Pol. Pract. 89, 151–163.
Li, Z., Hensher, D., 2011. Prospect theoretic contributions in understanding traveller behaviour: a review and some comments. Transp. Rev. 31 (1), 97–115.
Li, Z., Hensher, D.A., Rose, J.M., 2010. Willingness to pay for travel time reliability in passenger transport: a review and some new empirical evidence. Transport. Res. E: Logist. Transport. Rev. 46 (3), 384–403.
Michea, A., Polak, J., 2006. Modelling Risky Choice Behaviour: Evaluating Alternatives to Expected Utility Theory. Research, Kyoto.
Mishra, S., Tang, L., Ghader, S., Mahapatra, S., Zhang, L., 2017. Estimation and valuation of travel time reliability for transportation planning applications. Case Stud. Transp. Pol.
Noland, R.B., Polak, J.W., 2002. Travel time variability: a review of theoretical and empirical issues. Transp. Rev. 22 (1), 39–54.
Prashker, J.N., 1979. Direct analysis of the perceived importance of attributes of reliability of travel modes in urban travel. Transportation 8 (4), 329–346.
Quiggin, J., 1982. A theory of anticipated utility. J. Econ. Behav. Organ. 3 (4), 323–343.
RCore, T., 2012. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0 < http://www.R-project.org > .
Schmeidler, D., 1989. Subjective probability and expected utility without additivity. Econometrica: J. Econom. Soc. 571–587.
Senbil, M., Kitamura, R., 2004. Reference points in commuter departure time choice: a prospect theoretic test of alternative decision frames. Intell. Transport. Syst. Taylor & Francis.
Small, K.A., 1999. Valuation of Travel-Time Savings and Predictability in Congested Conditions for Highway User-Cost Estimation. Transportation Research Board.
Tang, L., Ghader, S., Mishra, S., Zhang, L., 2015. Impact of travel reliability on mode choice: empirical analysis with observed travel time variation data in Washington, DC. Transportation Research Board 94th Annual Meeting.
Tversky, A., Kahneman, D., 1992. Advances in prospect theory: cumulative representation of uncertainty. J. Risk Uncertain. 5 (4), 297–323.
Wang, W., Liu, X., Ding, L., Gao, G., Zhang, H., 2019. Stochastic network user equilibrium and traffic system evolution based on reference-dependent utility theory. IEEE Access 7, 19866–19880.
Xu, H., Zhou, J., Xu, W., 2011. A decision-making rule for modeling travelers' route choice behavior based on cumulative prospect theory. Transport. Res. C: Emerg. Technol. 19 (2), 218–228.
Yang, C., Liu, B., Zhao, L., Xu, X., 2017. An experimental study on cumulative prospect theory learning model of travelers' dynamic mode choice under uncertainty. Int. J. Transp. Sci. Technol. 6 (2), 143–158.
Yang, J., Jiang, G., 2014. Development of an enhanced route choice model based on cumulative prospect theory. Transport. Res. Part C: Emerg. Technol. 47, 168–178.
Zhang, C., Liu, T.-L., Huang, H.-J., Chen, J., 2018. A cumulative prospect theory approach to commuters' day-to-day route-choice modeling with friends' travel information. Transport. Res. C: Emerg. Technol. 86, 527–548.
Zhou, L., Zhong, S., Ma, S., Jia, N., 2014. Prospect theory based estimation of drivers' risk attitudes in route choice behaviors. Accid. Anal. Prev. 73, 1–11.