Identification of cumulative prospect theory parameters for mode choice model

Sylvia Indriany[1,*], *Ade* Sjafruddin[2], *Aine* Kusumawati[2], and *Widyarini* Weningtyas[2]

[1]Doctoral Study Program of Civil Engineering, Faculty of Civil and Environmental Engineering, Institut Teknologi Bandung, Bandung, Indonesia

[2]Transportation Engineering Research Group, Faculty of Civil and Environmental Engineering, Institut Teknologi Bandung, Bandung, Indonesia

Abstract The use of Cumulative Prospect Theory (CPT) in decision making related to transportation risk is still much debated. Mainly because of the travel and socio-economic characteristics of the traveller it possible for different responses to the specified Reference Point (RP) as well as the loss aversion. This difference can be seen from the value of Cumulative Prospect Theory parameters. Therefore, this paper will discuss about the determination of parameters CPT which affect public transportation mode choice model in the course of work trip activity. The reference point as an essential part of this study is determined based on the average travel time of commuter worker from South Tangerang City to Jakarta. Data obtained from stated preference survey, Feeder Busway/Busway and Commuter Line Jabodetabek as mode alternative and travel time attribute as a risk factor. The Binomial Logit model which has transformed utility distribution and probability with CPT and the Least Square Method to be obtained the parameters. Finally, some conclusions can be drawn that the CPT parameters produced by this study, have closed the range of value requirements in the CPT theory. So that the parameter value can be used to model the probability of mode choice with the risk of travel time in the study area. .

1 Introduction

Travel choices in urban areas are influenced by several factors including public transport services and perceptions of travellers on network conditions. With the increase in private vehicles, causing congestion which resulted in increased travel time. For this reason, the government encourages travellers to move to mass public transportation, especially for commuter workers from the DKI Jakarta buffer zone. But this must be accompanied by an increase in public transport services. Because [1] states that improving the quality of public transport services will attract private transport users. However, this increase depends on the context, especially the motivation of people to use private transport

On the other hand travel demand theory also develops along with people's behaviour in making decisions. The maximum utility concept that assumes the traveller knows his journey and make rational decisions is being abandoned. Because with the complexity of the network, travellers do not always know the conditions that will be faced. So those travel choice decisions are influenced by personality, psychological conditions, preferences about risk, environmental factors and other factors that cannot be accommodated by the maximum utility theory. It causes the use of EUT / RUT to have the potential to get a greater prediction from reality in providing transportation facilities. With evidence of the experimental results of behavioural researchers, finally, the idea that travellers behave rationally is opposed by many transportation researchers. [2,3,4,5]

To accommodate the shortcomings of the EUT / RUT, prospect theory was proposed [6] and furthermore cumulative prospect theory [7] which is a descriptive method that considers the risk of assuming that people will risks averse in the gain domain and risks seeking in the domain of loss relative to the reference point. The magnitude of this reference point will produce a difference between positive and negative utilities. So if applied in the mode choice with attributes of travel time as a risk factor, it will affect the magnitude of the probability of choosing a mode. Thus it is critical in this method that there is a Reference Point. That is the reference value of a trip attribute is determined from the average value of the attribute in real conditions.

However, one of the disadvantages of using CPT is the requirement to carry out empirical estimates of parameter values in the different travel behaviours context, because the existing CPT parameters estimate from risk behaviour in the financial field. The CPT parameters cannot be measured directly such as reference points, and there is no standard method in determining reference points in travel

---

*Corresponding author: syllfa.umb@gmail.com

choice so that further analysis of the use of parameters is necessary. According to [8] mode differences, the ability to use modes and the purpose of travel may have some effect on the traveller attitude in response to risks and uncertainties in travel choices. The above may result in various parameter values, and somewhat diffuse and complex distributions of parameter values overpopulation, time and other travel dimensions. So it is difficult to estimate a set of parameter values that represent the general decision maker.

Although estimates of parameter values have not been widely explored, some studies in economics and transportation have been carried out. In the field of economic, [9] get γ values that are not much different from [7]. Likewise with Wu and Gonzalez (1996), but for the parameter α obtained a smaller value. Whereas in the field of transportation, the value of the CPT parameters obtained is more varied. [10] estimate CPT parameters for more than one value except for the value α. [11] in the route choice study, the results of the lower CPT parameters were obtained. Whereas [12] in the mode choice gets a value of λ which is much larger than the parameters obtained by [7]. These results indicate that the determination of CPT parameters in the field of transportation results in more diverse values due to differences in the characteristics of travellers and their sensitivity to risk. For this reason, the paper aims to examine the suitability of existing CPT parameters if applied to the choice of commuter public transport modes for work trips with more heterogeneous respondents. As well as the possibility of using the new CPT parameter for the case.

2 Literature review

2.1. Discrete choice model

The Utility is something that can be maximised by an individual and an indicator in making a decision. In the context of the mode choice, the traveller will choose the alternative mode with the highest utility. Nevertheless, such utility models do not always correspond to the reality because the individual decision can vary no discernible reason. [13] stated that this model contains errors both from the decision maker and the model maker's assumptions. In the development, the probability model (RUT) accommodated the lack, so the formula for this model becomes:

$$U_{it} = V_{it} + \varepsilon_{it} \qquad (1)$$

where:

U_{it}	=	The true utility of the alternative i to the decision maker t, is equivalent to $U(Xi, St)$
V_{it}	=	The deterministic or observable portion of the utility estimated,
	=	$v(S_t) + v(X_i) + v(S_t, X_i)$

ε_{it}	=	The error or the part of the utility unknown.

Furthermore, in the development of the travel choice as the disaggregated model known as the Discrete Choice Model, in which the traveler is choosing an alternative as a function of socio-economic characteristics and the attraction of that choice. The concept of utility is used to declare the magnitude of an option. Some models that follow the maximum utility concept are logit models. Logit models commonly used in the mode choice as binomial logit and multinomial logit

In the binomial logit model, the decision maker is confronted with a pair of discrete alternatives in which the selected alternatives are those with the highest utility and as random variables. Assuming that ε_n distributes Gumbell independently and identically, then it is equal to assume that $\varepsilon_n = \varepsilon_{jn} - \varepsilon_{in}$ is logistically distributed, thus the choice probability for alternative i is given by:

$$P_n(i) = \frac{e^{V_{in}}}{e^{V_{in} + e^{V_{jn}}}} \qquad (2)$$

where $V_{in} = \beta X_{in}$

Furthermore, the multinomial logit model is used to model the mode choice that consists of three or more alternatives and is a frequently used discrete model. The general form of the probability of selecting alternative i ($i = 1,2, \dots, j$) from one alternative set j is:

$$Pn(i) = \frac{\exp(V_{in})}{\sum_{j=1}^{j} \exp(V_{jn})} \qquad (3)$$

where :

$Pn_{(i)}$ = The probability of the decision-maker choosing alternative i

V_{jn} = The systematic component of the utility of alternative j.

= βX_{jn}

β = The parameter which defines the direction and importance of the effect of attribute k on the utility of an alternative j

X_{jn} = the value of attribute k for alternative j

2.2. Cumulative prospect theory

Prospect theory is an advanced approach in explaining people's behavior when dealing with uncertainty in choices. So that if the Expected Utility Theory is initially to help people reach better decisions, Prospect Theory provides additional decisions that are not risky and risky. The theory developed into Cumulative Prospect Theory proposed by [7]. CPT is a decision-making method with a descriptive approach and sees uncertainty from a traveler's point of view. So that the decision depends very much on how the choice is prepared against a reference point or the way the traveler views risk with the complexity of the transportation network

The CPT method has a value function curve to represent people's behavior in the face of risk that results in a profit or loss against a reference point. It also produces a weighting probability function that shows a reduction in marginal sensitivity due to benchmark uncertainty. [7] apply cumulative functional rather than separating weights for gains and losses. To determine cumulative functions, the results of each prospect are arranged in order to increase their values. A prospect is f then represented as a sequence of pairs (x_i, A_i), which produces x_i if A_i occurs, where $x_i > x_j$ if $i > j$ and A_i are part of S

The positive part f, denoted f^+, is obtained by stating $f^+(s) = f(s)$ if $f(s) > 0$, and $f^+(s) = 0$ if $f(s) \leq 0$. The negative part of f, denoted f^-, is defined as equivalent. CPT confirms that there is a strict increase in the function value v: $X \to Re$. so that $v(x_0) = v(0) = 0$ and the weight of the decision function w^+ and w^-, so that for f = (x_i, A_i), -m $\leq i \leq$ n,

$$V(f) = v(f^-) + v(f^+) \qquad (4)$$

$$V(f^+) = \sum_{i=0}^{n} \pi_i^+(x_i) \qquad (5)$$

$$V(f^-) = \sum_{i=-m}^{0} \pi_i^-(x_i) \qquad (6)$$

where $v(f^+)$ is the prospect of profit value, $v(f^-)$ad is the prospect of loss value, $\pi^+(f^+) = (\pi^+_0, ..., \pi^+ n)$ is the weight of the decision on profits, and $\pi^-(f^-) = (\pi^-_{-m}, .., \pi^-_0)$ is the weight of the decision of the loss. If the prospect f = (x_i, A_i) is given by p probability distribution $(A_i) = p_i$, it can be seen as a probabilistic or risky prospect (x_i, p_i). In this case, the decision weight is determined by

$$\pi_n^+ = w^+(p_n) \qquad (7)$$
$$\pi_m^- = w^-(p_{-m}) \qquad (8)$$
$$\pi_i^+ = w^+(p_i + \cdots + p_n) - w^-(p_{1+i} + \cdots + p_n)$$
$$0 \leq i \leq n-1 \qquad (9)$$
$$\pi_i^- = w^-(p_{-m} + .. + p_i) - w^-(p_{-m} + .. + p_{i-1})$$
$$1 - m \leq i \leq 0 \qquad (10)$$

where w^+ and w^- strictly increase the function of the unit interval into itself so $\pi^+(0) = w^+(0) = w^-(0) = 0$ and $\pi^-(1) = w^+(1) = w^-(1) = 1$. Thus the value function that is in accordance with Cumulative Prospect Theory is:

$$\Delta x = x = x_i - x_0 \qquad (8)$$
$$v(x) = x^\alpha \quad if \ x \geq 0 \qquad (9)$$
$$v(x) = -\lambda * (-x)^\beta \quad if \ x < 0 \qquad (10)$$

Parameter $\lambda \geq 1$ describes the individual level loss aversion and parameter $0 < \alpha, \beta \leq 1$ measures the diminishing sensitivity. The proposed weighting functions for the gain and loss are, respectively:

$$w^+(p) = p^\gamma / ((p^\gamma + (1-p^\gamma)^{\frac{1}{\gamma}})) \qquad (11)$$

$$w^-(p) = p^\delta / (p^\delta + (1-p^\delta)^{\frac{1}{\delta}}) \qquad (12)$$

Where p is the stated probability for the result to i, $0 < \gamma < 1$ illustrates the phenomenon that people always overestimate the small probability and underestimate the high probability. [7] estimated the parameters α, β, λ,

γ and δ, and obtained the value α= β=0.88, λ=2.25, γ=0.61 and δ=0.69.

Furthermore, the cumulative prospect value is calculated by accumulating profits and losses. Suppose the lottery consists of m + n + 1, $x_{-m} < ... < x_0 < ... < x_n$, with the probability that each $p_{-m}, ..., p_n$ occurs. Then, the corresponding risk can be represented by the pair (x; p), where x = $(x_{-m}, .., x_n)$ dan p= $(p_{-m}, .., p_n)$. Cumulative decision weight is defined as follows

$$\pi^+(p_i) = w(\sum_{n=1}^{n} p_k) - w(\sum_{n=1+i}^{n} p_k),$$
$$0 \leq i \leq n \qquad (13)$$
$$\pi^+(p_{-j}) = w(\sum_{k=-m}^{-j} p_k) - w(\sum_{k=-m}^{-j-1} p_k),$$
$$m \leq -j \leq 0 \qquad (14)$$

Accordingly, Cumulative Prospect Value (CPV) of (x; p) is calculated as

$$CPV(x,p) = \sum_{i=0}^{n} v(\Delta x_i).\pi^+(p_i) + \sum_{i=-m}^{-1} v(\Delta x_i).\pi^-(p_i) \qquad (15)$$

CPV values indicate preferences for choices and should be positive because they are expected to be profitable. But when the value between two choices is negative then the one chosen is the one that gives the least loss.

3 Research method and data collection

3.1. Data collection

The Cumulative Prospect Theory method is sensitive to the Reference point value to measure changes in traveller preferences towards the gain and loss domain. So that, the revealed preference survey was conducted to obtain the average travel time of commuter workers from South Tangerang to Central Jakarta. Reference Point in this study used the average travel time. Furthermore, with this value two scenarios of travel time and a probability of occurrence are made. The first scenario consists of 5 questions with RP 70 minutes. Travel time in each subject creates less than or equal to 70 minutes (Gain domain). But the probability of bus more variation than Jakarta Commuter lines. So that, the respondent will choose one of the two modes offered.

In the second scenario, RP = 45 minutes used to answer five next question. Similar to the first scenario, the mode alternatives are offered a bus and commuter line. The difference that the scene of travel time offered is more significant than 45 minutes, or the respondent in the domain loses. With these ten questions stated preference surveys were carried out to 100 respondents and later become input to the parameter determination program.

3.2. Data structure in mathlab R2016a

To get the CPT parameters created the program with the help of Matlab 2016a software. And as input to the program are:

a. Time and probability scenario.

b. Stated preference survey results

Then it is created in the excel data structure like Table 1 below. The B1-B10 shows ten scenarios for Bus, while the K1-K10 is ten scenarios for commuter lines. With a time range of 45-70 minutes, for the B1-B5 scene, and the K1-K5 reference point is 70 minutes or on the gains domain. Whereas reference point 45 minutes in the loss domain for K6-K10 and B6-B10. Furthermore, because the method used in determining this CPT parameter is the least square, the percentage of bus mode choices is required from the survey results symbolised by % Bus.

Tabel 1. Data structure

Range of T.Time	Probability Scenario in each mode										%Bus
	B1	B2	B3	B4	B5	K1	K2	K3	K4	K5	
45	0	0	0	0.1	0.1	0	0	0	0	0	0.55
47.5	0.1	0.1	0.1	0	0	0	0	0	0	0.5	0.53
50	0	0	0	0	0	0	0	0	0	0	0.52
52.5	0	0	0	0	0	0	0	0	0	0	0.51
55	0	0.5	0.7	0	0	0.2	0.7	0.9	0.3	0	0.55
57.5	0	0	0	0	0	0	0	0	0	0	
60	0	0	0	0	0.3	0	0	0	0	0	
65	0	0	0	0	0	0	0	0	0	0	
69.5	0.9	0.4	0.2	0.9	0.6	0.8	0.3	0.1	0.7	0.5	

Range of T.Time	Probability Scenario in each mode										%Bus
	B6	B7	B8	B9	B10	K6	K7	K8	K9	K10	
45.5	0.9	0.6	0.1	0.3	0.5	0.7	0.5	0	0.2	0.4	0.43
47.5	0	0	0	0	0	0	0	0	0	0	0.49
50	0	0	0	0	0	0	0	0	0	0	0.69
52.5	0	0	0	0	0	0	0	0	0	0	0.55
55	0	0.3	0.8	0.6	0.4	0.3	0.5	1	0.8	0.6	0.43
57.5	0	0	0	0	0	0	0	0	0	0	
60	0	0	0	0	0	0	0	0	0	0	
65	0	0	0	0	0	0	0	0	0	0	
70	0.1	0.1	0.1	0.1	0.1	0	0	0	0	0	

3.2. Determination CPT parameters

To determine the CPT parameters, the steps in the Matlab 2016a program are made according to the following diagram. There are two step for estimates CPT parameter. First step to obtaine α and γ values. The second, with the γ value that has been obtained before, then the other parameters are determined, namely δ and λ.
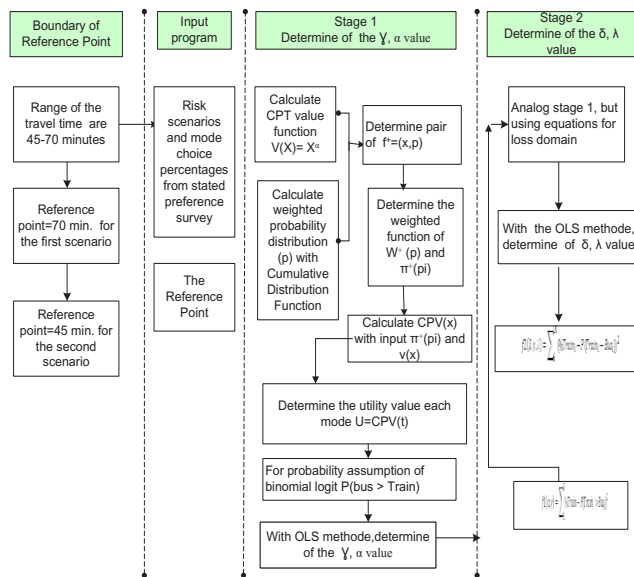
Fig. 1. Determination of parameter

4 Result and discussion

Based on the data that has been done to determine the CPT parameters, with the attribute of travel time as a risk factor, the parameters that are in accordance with the previous theory are obtained. However, for δ which is a elevation of weighting function parameter, the value is smaller than the previous studies. The results of all these parameters are:

Table 2. Result of CPT parameters

Parameters	Symbol	Value of CPT parameters	Reference parameters for previous researchers
Diminishing of Sensitivity	$\alpha = \beta$	0,78	$\alpha, \beta \leq 1$
the elevation of the weighting function(+)	γ	0,50	$\gamma > 0$
loss aversion	λ	1,31	$\lambda > 1$
the elevation of the weighting function(-)	δ	0,1	$\delta > 0$

The parameters α and β have a maximum limit of 1, which shows the real risk aversion. Whereas $\alpha \leq 1$ indicates the risk aversion for gain and the risk-seeking for loss. Thus the higher the value α means the more significant of the risk aversion.

While the value of λ as a parameter of risk aversion, then the value is greater than 1 indicating that individuals are more sensitive to losses than gain. From the value of the CPT parameter based on the travel time above, it can be made an image of the weighting function compared to the previous research, as shown in the following Figure 2.
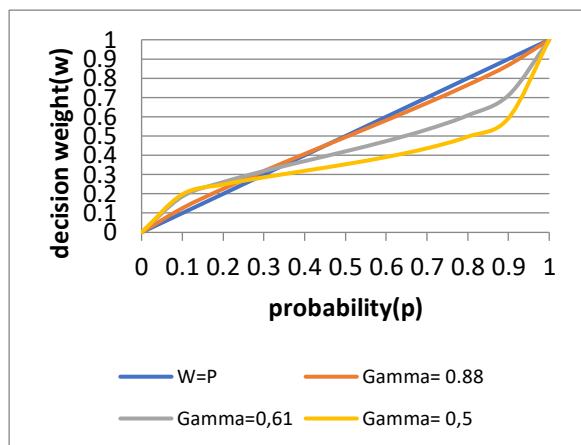
Fig. 2. Comparison of weighting function curve

Referring to the previous research, the fields of economic psychology [7] give a value of $\gamma = 0.61$. Furthermore in the field of transportation, [10] get a value of $\gamma = 0.82-0.88$. In 2014, [112] received $\gamma = 0.88$ in the case of bus mode choice with limited respondents. The parameter results γ obtained in this study are 18% smaller than the values obtained by [7]. However, they have the same pattern, that is, the small probability that people tend to be overestimated and on a high probability people tend to under estimate the risks to be obtained. This difference can be influenced by the determined Reference Point value and stated preference scenario and respondent's behavior in responding to risky attributes. Thus in the transportation parameter estimates in the scope of transportation can be very varied**.**

Furthermore, based on the parameters of the results of this study, the proportion of bus and commuter line modes can be calculated. Overall for the travel time attribute can be seen the difference in the portion of the mode choice of the model results with the survey results, as shown in Table 3. From the table, it can be calculated that the difference in the average of each model to the survey results. In the model using the CPT parameter, this study obtained a difference of 6.2% for bus mode and 10% for commuter line mode. Whereas by using the CPT parameter, previous research obtained quite significant results, namely 17.8% and 22.6% respectively for buses and commuter line JABODETABEK.

Table 3. Comparison of proportion of moda choice.

Scenario	New Parameter (%)		Tversky and KahnemanParameters(%)		Result of survey(%)	
	BUS	CJ	Bus	CJ	Bus	CJ
1	46.86	53.14	44.71	55.29	55	45
2	49.42	50.58	49.01	50.99	53	47
3	48.16	51.84	46.61	53.39	52	48
4	49.17	50.83	48.47	51.53	51	49
5	49.48	50.52	49.02	50.98	55	45
6	48.00	52.00	39.61	60.39	43	57
7	48.32	51.68	36.43	63.57	49	51
8	48.44	51.56	37.75	62.25	69	31
9	48.41	51.59	37.30	62.70	55	45
10	48.36	51.64	36.75	63.25	43	67

5 Conclusion

From this study, some conclusions can be drawn that the CPT parameters produced by this study, with the attribute of travel time as a risk factor have closed the range of value requirements in the CPT theory. However, there is still one parameter whose value is small compared to previous researchers. In the scope of transportation, it can occur because of differences in behaviour of the traveller in response to network complexity. So it has the potential to have a long range in the results of the CPT parameters. On the weighting function curve, it has the same pattern as the previous research so that the parameter value can be used to model the probability of mode choice with the risk of travel time in the study area

References

1. L. Redman, Margareta Friman, Tommy G̈arling and Terry Hartig. Transport Policy **25**, 119 –127 (2013)
2. S. Fujii and R. Kitamura, Networks and Spatial Economics **4 (3**), 243–256. (2004)
3. E. Avineri and J.N. Prashker. Transportation Research Record **1894**, 222–229 (2004)
4. E. Avineri & J.N. Prashker. Transportation Research Record **1854**, 90–98(2003)
5. A. Ceder, S. Chowdhury,N. Taghipouran and J.Olsen, Transport Policy **27** 112–122(2013)
6. D. Kahneman and A. Tversky, Econometrica, **Vol. 47, No. 2.** 263-292 (1979)
7. A. Tversky and D. Kahneman, Journal of Risk and Uncertainty **5 (4)**, 297–323 (1992)
8. E. Avineri, Transportation Research Record **2082**: Travel Behavior Analysis, 141-147 (2008)
9. C.F. Camerer and T.H Ho. Journal of uncertainty and uncertainty, 8(2), 167-196 (1994)
10. T. Schwanen and D.Ettema,. Transportation Research Part A: Policy and Practice, **43(5),** 511 - 525 (2009)
11. H. Xu, J. Zhou and W. Xu, Transport Research Part C **19 (2),** 218–228 (2011)
12. L. Zhao and C. Yang, The 10th International Conference of Eastern Asia Society for Transportation Studies, EASTS_ISC-D-13-00265, 2013)
13. F.S. Koppelman and C. Bhat, U.S. Department of Transportation Federal Transit Administration (2006)
14. E. Avineri, Transportation Science **40 (4),** 409–420 (2006)

15. L. Xiaowei, W. Wei, X. Chengcheng, L. Zhibin and W. Baojie, J Syst Sci Complex **28** 661–678 (2015)

16. P. Wakker, Fennema, H., Journal of Behavioural Decision Making (1997)