Step versus slope model for valuing waiting time unreliability

Baiba Pudane, TUM CREATE (Singapore), baiba.pudane@tum-create.edu.sg

Andreas Rau, TUM CREATE (Singapore)

Abstract

This work addresses the problem of designing the choice experiment to test the way public transport passengers perceive unreliability. Two methods from the scheduling family of the reliability measurement functions are selected – step and slope model. Both of them have good arguments for being well suited to describing the decision making process of passengers, as well as both have been shown to be reducible to closed forms. However, there has been little research to guide practitioners in choosing between these two models.

This paper presents a reliable way of designing a stated choice experiment, which would allow deciding whether the perception of waiting time unreliability by passengers is closer to the step or the slope model. The design procedure is explained step-by-step, allowing future studies to replicate it. Then, the experimental design of the stated choice experiment is tested in simulations for the ability to distinguish between the step and slope model preferences. The results show acceptable model performance in several scenarios.

Keywords: travel time variability, reliability, valuation, public transport, experimental design, stated preference

1 Introduction

The purpose of studies of passenger valuations of travel time components is benefit assessment of transport improvements. If the proposed improvement concerns travel time reliability, then two methods of formulating perceived utility are dominant in literature: scheduling model (found by Small, 1982) and centrality-dispersion method (as named by Carrion & Levinson, 2012). The scheduling model measures the disutility from travel time unreliability with variables such as earliness and lateness experienced at the destination. Through explicit departure and arrival times it enables departure time choice modelling, which is conceivably the first instrument for passengers to adjust to unreliable travel time. Therefore, it can be argued that scheduling model approximates the decision making process of passengers very well. However, centrality-dispersion method, which has been the most widely used method in reliability studies, is advantageous, because it includes in the utility function statistical measures to describe the spread of the travel times, like standard deviation of travel times. Such measures are often readily available in practice unlike the scheduling model measures, which depend on the preferred arrival time of each traveller and are therefore hard to observe.

Therefore it would seem very difficult to apply the scheduling approach for any real world project assessment. However, Fosgerau and Karlström (2010) contributed greatly to the solution of this problem by translating the most common scheduling function (later referred to as the "step model") into a closed form[1], that is, a form which contains only statistics of the travel time distribution. Later, Fosgerau and Engelson (2011) added on their result by introducing a new form of the scheduling function named "slope model", which, as they demonstrated, can also be reduced to a closed form. The names of these two models stem from the underlying assumption about the value of time at origin and destination – constant for the step model and changing linearly for the slope model. Both models will be shown later in chapter 2.

Most recently, Benezech and Coulombel (2013) developed a framework to use the step model scheduling function for reliability assessment of public transport waiting times. They introduced two indicators – Value of Service Headway (VoSH, meaning the value of the mean headway of a service) and Value of Service Reliability (VoSR, meaning the value of the standard deviation of a service headway), which are derivatives of a reduced form of the step model for waiting times. These indicators can be a useful instrument for practitioners to compare, for example, the benefits of introducing more buses versus improving the headway regularity. It is possible to obtain the VoSH and VoSR for the slope model as well, as will be shown in chapter 2.

In the light of the recent contributions of the step and slope models, a question of interest arises: which model represents the passenger decision making better? An answer to this question will have a direct influence on the benefit estimation of transport improvement

[1] Previously, the result was developed only for specific travel time distributions, for example, by Noland and Small (1998) and Bates et al. (2001). Fosgerau and Karlström (2010) presented the result in a general case.

proposals. In this paper this question is addressed by proposing a technique to create stated choice experimental designs, which would enable judgement of the fit of these two models.

The outline of the paper is following: first, the step and slope models are introduced and their application in public transport reliability modelling shown. Then, a brief qualitative comparison between the two models is given. After, a step-by-step methodology is developed to create an experimental design to test the suitability of the two models for waiting time reliability modelling. Following, a methodology is explained for testing the created design in different scenarios via a simulation. Finally, the results are presented.

2 Introduction of the models

As mentioned in the introduction, two models of the scheduling family are prominent in literature: step and slope model. These models share the advantage of being reducible to closed forms. That is, although they depend on the departure and arrival times of passengers in the formulation, by assuming optimal behaviour of travellers, the utility can be expressed in closed form: depending only on the distribution statistics of travel times. Furthermore, both models can lay claim on possibly intuitive representation of utilities at home and work. Step model assumes rigid constraints at the origin and destination, while slope model assumes gradual decrease and increase of the importance of being at home and work respectively. There have been other scheduling models proposed as well, e.g., Li et al. (2012), Jou et al. (2008), which could be studied in future work.

In the remainder of this chapter three aspects of each model are presented: formulation, reduction to a closed form and application to public transport. All these points are introduced very briefly, for more detailed derivations, cited publications should be consulted.

In application to public transport, it is assumed that waiting times are variable and in-vehicle times are constant. In this way, the model is adjusted for evaluating waiting time reliability. This split is necessary, as explained by Benezech and Coulombel (2013), because the data necessary for estimating the distribution of waiting times are often easier accessible as compared to total travel times. This is especially true in cities like Singapore, where public transport passengers use smart cards, which record travel times and itineraries and deducts fare. Furthermore, the split also facilitates easier usage of the values by operators, who can control the headway (and thus waiting time) variability easier as compared to the total travel time variability.

Alternatively, one could set waiting time constant and analyse the value of in-vehicle time reliability. Analysing both waiting time and in-vehicle time reliability simultaneously is problematic due to the complexity of analysis as well as presentation to the survey respondents. Vincent and Hamilton (2008) addressed this problem by carrying out two surveys – one exploring the attitudes towards the variable departure times, the other – towards variable in-vehicle times. In this paper, waiting time reliability is chosen as preferable because it enables estimation of value of service headway and value of service headway reliability (as defined by Benezech & Coulombel, 2013). These indicators allow an easy assessment and comparison of benefits from improved bus service reliability and/or frequency, which is a desired outcome of the envisioned case study.

Step model in general case and derivation of the closed form

The simplest form of the step scheduling function contains only travel time, earliness and lateness at destination:

$$U(t) = \alpha T + \beta(t^* - t - T)^+ + \gamma(t + T - t^*)^+ \tag{1}$$

Here t^* stands for preferred arrival time, t – for departure time and T – for travel time. Travel time is a stochastic variable. Term $(t^* - t - T)^+$ is defined as maximum between $t^* - t - T$ and 0. Therefore $(t^* - t - T)^+$ represents earliness and similarly $(t + T - t^*)^+$ represents lateness.

This function is the most widely used form of scheduling preferences. It was originally proposed by Vickrey (1969) and Small (1982). Börjesson et al. (2012) named it a "step model" due to the sudden change of utility as the arrival time passes the preferred arrival time. An illustration of the step model in case of earliness (taken from Börjesson et al., 2012) is depicted in the Figure 2.1. The step utility (1) can be calculated as the loss of utility either at home or at work due to travelling.

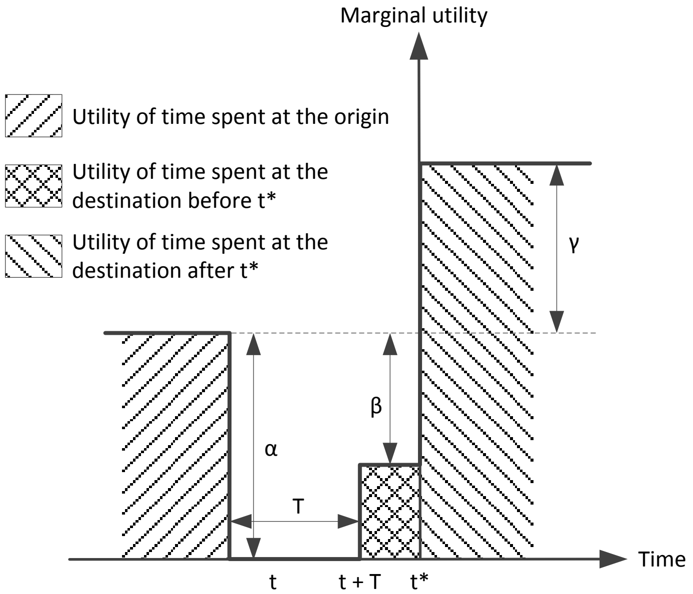


Figure 2.1 Visualisation of the step model - adapted from Börjesson et al. (2012)

Fosgerau and Karlström (2010) used function (1) as a starting point to obtain the reduced form, i.e., the form which expresses the utility as a function of the distribution of the travel time. Following is a brief explanation of their procedure summarised in three steps.

Firstly, the expected utility is calculated from the function (1). Means of the stochastic utility function components – travel time, earliness and lateness – are taken. The expected utility is[2]:

$$E\big(U(t)\big) = \alpha\mu + \beta\int_{0}^{t^{*}-t}(t^{*}-t-x)\phi(x)dx + \gamma\int_{t^{*}-t}^{+\infty}(t+x-t^{*})\phi(x)dx \qquad (2)$$

Here μ is the mean travel time, $\phi(x)$ is the probability density function of the total travel time. Consequently, the integrals represent the mean earliness (multiplied by β) and mean lateness (multiplied by γ). All the parameters α, β, γ are negative.

Secondly, the expected utility function is used to obtain the optimal departure time. This is done by differentiating the expected utility (2), and setting the derivative equal to 0. Consequently, the optimal departure time, which maximises the utility, is obtained. In essence this step equals an assumption that the passengers, who perform the trip regularly, are able to intuitively estimate the stochastic travel time distribution and optimise their departure time based on their preferences.

Here it should be noted that the function of the optimal departure time is continuous; therefore the traveller is able to realise the optimal departure time only if any departure time is a potential departure time for him. This is the case for private transport travellers.

Thirdly, if the passenger chooses optimal departure time then he experiences the maximum utility U^{*}. This is obtained by inserting the optimal departure time in the expected utility function. Executing the calculations analogously to Fosgerau and Karlström (2010) and Fosgerau and Engelson (2011), but for a non-standardised travel time, the maximum expected utility is:

$$EU^{*} = E\left(U\big(t_{opt}\big)\right) = (\alpha-\beta)\mu + (\beta+\gamma)\int_{\frac{\gamma}{\beta+\gamma}}^{1}\Phi^{-1}(s)ds \qquad (3)$$

This is the reduced form of the step function of total travel times. Here $\Phi^{-1}(s)$ is a quantile function of the cumulative distribution function of travel times.

It can be seen that the maximum expected utility no longer depends on the departure and arrival time, but only on the statistics of the travel time. Therefore it can be conveniently applied in practice, for example, by calculating the value of the mean of total travel time or the value of the standard deviation of total travel time as derivations of (3).

[2] Here the travel time is not standardised. But other authors like Fosgerau and Karlström (2010) and Börjesson et al. (2012) operated with standardised travel time, in which case the travel time x is replaced with $\mu + \sigma x$, where μ and σ are mean and standard deviation of the travel time. The standardised travel time approach is beneficial for some purposes. But, considering the desired separation of the total travel time in the waiting time and in-vehicle time, the standardisation would create complications. Non-standardised travel time approach was also chosen by Benezech and Coulombel (2013).

Application of the step model in public transport

Having the basic model outlined, it is possible to adjust it for the use of waiting time modelling in public transport. Following the classification of Fosgerau and Engelson (2011), the adjustment is done for unplanning passengers, i.e., passengers who do not time their arrival at bus stop but arrive at random times. Therefore the unplanning passengers experience random waiting time, while planning passengers have no waiting time. The model for unplanning passengers is suitable if the headways are short or there is no information about the bus arrival times and buses arrive randomly, which is typical in Southeast Asia.[3]

The adjustment of the formula is done by separating the total travel time into waiting time and in-vehicle time, and assigning different weights to them. Next, in-vehicle time is set to be constant, so that the waiting time variability is analysed. Then the expected utility function (2) becomes:

$$E\big(U(t)\big) = \alpha_v T_v + \alpha_w \mu_w + \beta \int_0^{t^*-t-T_v} (t^* - t - T_v - x)\phi(x)dx + \gamma \int_{t^*-t-T_v}^{+\infty} (t + T_v + x - t^*)\phi(x)dx \tag{4}$$

Here T_v is the in-vehicle time, μ_w is the mean waiting time and $\phi(x)$ is the probability density function of waiting times. This function will be used for the creation of the experimental design. It is the view of the authors that the expected utility function should be used for reliability studies. Otherwise, the respondents would perceive the delay as known a priori and available for rearrangements. This would likely lead to the valuation of earliness and lateness being much reduced – which was the result in the study of Börjesson et al. (2012).

To obtain the closed form of the step function in public transport case, the cumulative distribution function, which enters (3), needs to be transformed. The cumulative distribution function is equivalent for the total travel time and waiting time, if the in-vehicle time is constant:

$$\Phi(Tot) = \Phi_w(Tot - T_v) = \Phi_w(w)$$

The quantiles of the total distribution are:

$$\Phi^{-1}(s) = \{x + T_v \in R{:}\, s = \Phi(x + T_v) = \Phi_w(x)\} = \Phi_w^{-1}(s) + T_v \tag{5}$$

The maximum expected utility (3) becomes:

[3] However, the real time information about bus arrival times is nowadays accessible in many locations. In this way, it could be argued that the model of planning passengers is suitable too. However, it is unclear how large is the share of passengers who use these information services. Moreover, the information is often available only in short term, which means that the waiting time in the bus stop is transferred to the waiting time at the origin, which is not exactly no waiting time, as understood in the planning passengers category.

$$EU^* = (\alpha - \beta)(T_v + \mu_w) + (\beta + \gamma)\int_{\frac{\gamma}{\beta+\gamma}}^{1} (\Phi_w^{-1}(s) + T_v)ds \tag{6}$$

If different weights are assigned to in-vehicle time and waiting time then the maximum expected utility can be expressed:

$$EU^* = (\alpha_w - \beta)\mu_w + \alpha_v T_v + (\beta + \gamma)\int_{\frac{\gamma}{\beta+\gamma}}^{1} \Phi_w^{-1}(s)ds \tag{7}$$

By renaming some variables these results are equivalent to Benezech and Coulombel (2013).[4] They subsequently differentiated (7) to obtain the afore-mentioned VoSH and VoSR.

Slope model in general case and derivation of the closed form

Vickrey (1973) first considered and Fosgerau and Engelson (2011) analysed an alternative type of scheduling function, which was named the "slope model" by Börjesson et al. (2012). In contrast to the step model, the slope model assumes that utility in the origin and destination changes in a linear way with time: $h(x) = \beta_0 + \beta_1 x$ for the utility at home and $w(x) = \gamma_0 + \gamma_1 x$ for the utility at work. The utility can be expressed as an integral over the time spend at home and at work in an arbitrary interval$(A; B)$, which includes the travel time:

$$U(t,T) = \int_A^t (\beta_0 + \beta_1 s)ds + \int_{t+T}^{B} (\gamma_0 + \gamma_1 s)ds \tag{8}$$

Here t is the departure time, $t + T$ is the arrival time, β_1 – assumed negative – is the slope parameter of the home utility function and γ_1 – assumed positive – is the utility of the work utility function. The signs of the slope parameters mean that the marginal utility of time at home is decreasing and the marginal utility of the time at work is increasing. Furthermore, the 0 point is conveniently defined at the time when work and home utility functions cross. In this way it holds that $\beta_0 = \gamma_0$. The intercept of functions is assumed positive, which will lead to the value of travel time being negative. Börjesson et al. (2012) illustrated the slope model as in Figure 2.2.

[4] They introduced the notion of head start, which is defined as the time which passengers allocate for waiting: $t^* - T_v - t$.

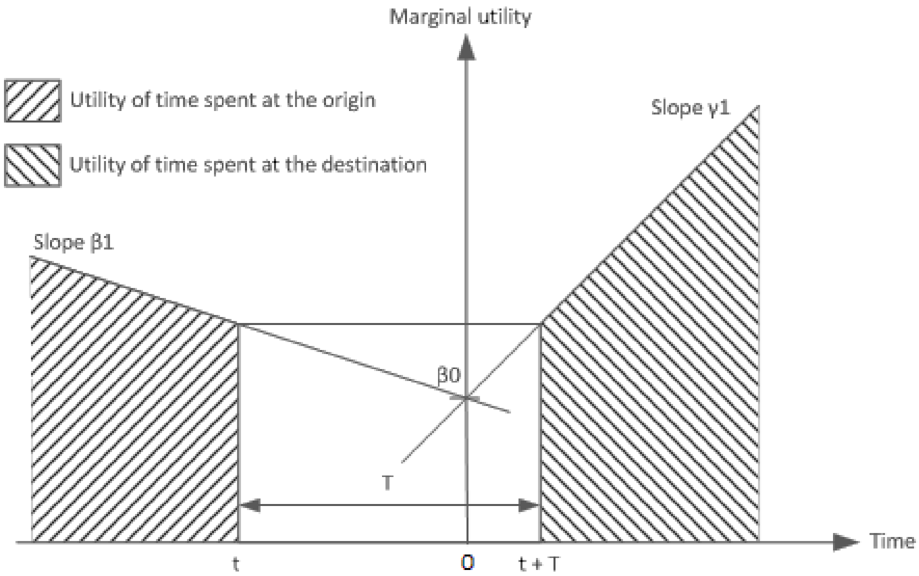


Figure 2.2 Visualisation of the slope model - adapted from Börjesson et al. (2012)

They explained that in contrast to the step model there is no one preferred arrival time. The optimal departure time is chosen such that the lowest utility time of length T is spent travelling. This is illustrated in Figure 2.2, where the travelling time T is located in a way which minimises utility loss, which means that the marginal utility at home at the departure time is equal to the marginal utility at work at the arrival time.

A consequence of this is that the most preferred arrival time is not set constant but varies with the length of travel time, which was discussed by Börjesson et al. (2012). This characteristic may apply to some situations, but likely not in all.

Subtracting the maximum possible utility (when travel time equals 0) $U(0,0) = \int_A^0 (\beta_0 + \beta_1 s)ds + \int_0^B (\gamma_0 + \gamma_1 s)ds$ from (8), the disutility of travelling is obtained:

$$
\begin{aligned}
U(t,T) &= -\int_t^0 (\beta_0 + \beta_1 s)ds - \int_0^{t+T} (\gamma_0 + \gamma_1 s)ds \\
&= \beta_0 t + \frac{\beta_1}{2} t^2 - \gamma_0 (t+T) - \frac{\gamma_1}{2}(t+T)^2 \\
&= -\beta_0 T - \frac{\gamma_1}{2}(t+T)^2 + \frac{\beta_1}{2} t^2
\end{aligned}
\tag{9}
$$

If the travel time T is stochastic with mean μ and standard deviation σ then the arrival time becomes $a = t + \mu + \sigma x$, where x is a random variable representing the standardised travel time. Then expected arrival time becomes $E(a) = t + \mu$, but expected squared arrival time is:

$$
\begin{aligned}
E(a^2) = E((t + \mu + \sigma x)^2) &= E(t^2 + 2t(\mu + \sigma x) + (\mu + \sigma x)^2) \\
&= t^2 + 2t\mu + \mu^2 + \sigma^2
\end{aligned}
\tag{10}
$$

Then expected utility of slope model can be obtained from (9) and (10):

$$EU(t) = -\beta_0\mu - \frac{\gamma_1}{2}(t^2 + 2t\mu + \mu^2 + \sigma^2) + \frac{\beta_1}{2}t^2 \tag{11}$$

This equation will be later modified for the use in the creation of the experimental design. Therefore it should be noted that it depends on the previously defined reference or 0 point through the departure time t. Since in reality the reference point is unknown it will act similarly as model parameter in the sense that it will be estimated from the (simulated) choices.

Finalising the derivation of closed form, following is found by Fosgerau and Engelson (2011) and Börjesson et al. (2012):

$$\begin{aligned} t_{opt} &= \frac{-\gamma_1\mu}{\gamma_1 - \beta_1} \\ EU^* = EU(t_{opt}) &= -\beta_0\mu - \frac{\beta_1\gamma_1}{2(\beta_1 - \gamma_1)}\mu^2 - \frac{\gamma_1}{2}\sigma^2 \end{aligned} \tag{12}$$

This corresponds to the case when travellers can choose their departure times freely, i.e., are travelling by private transport. The result shows that utility does not depend on the distribution of travel times, but only on the mean and standard deviation of the distribution. Therefore, if this model was true then it would possess a significant convenience in application as compared to the step model: no knowledge about travel time distribution would be necessary to use the results. Also, all computation would always be contained in closed forms independent of the travel time distribution.

Application of the slope model in public transport

Fosgerau and Engelson (2011) also showed that slope model could be applied to scheduled services. For unplanning passengers the travel time is split in waiting time and in-vehicle time. Similarly as for the step model, in-vehicle time is set to be constant, and different weights are assigned to waiting time and in-vehicle time. The expected utility function becomes:

$$\begin{aligned} EU(t) &= \alpha_v T_v + \alpha_w \mu_w + \frac{\beta_1}{2}t^2 - \frac{\gamma_1}{2}E(a^2) \\ E(a^2) &= E((t + \mu_w + T_v + \sigma_w x)^2) \\ &= t^2 + 2t\mu_w + \mu_w^2 + \sigma_w^2 + 2tT_v + 2\mu_w T_v + T_v^2 \end{aligned} \tag{13}$$

This function will be used for the creation of the experimental design.

The optimal expected utility can be obtained from (12):

$$EU^* = EU(t_{opt}) = \alpha_v T_v + \alpha_w \mu_w - \frac{\beta_1\gamma_1}{2(\beta_1 - \gamma_1)}(\mu_w + T_v)^2 - \frac{\gamma_1}{2}\sigma_w^2 \tag{14}$$

Qualitative comparison of the step and slope model

The aim of this paper is to develop a methodology to decide between the step and slope models. Given that both models provide a plausible explanation for passenger perception of reliability and that both are useable in practice, since they can be converted to closed forms as shown before, more criteria for the selection should be developed. Some aspects along with their assessment for both models ("+" for advantages and "-" for disadvantages) are given in the table below.

Table 2.1 Comparison between step and slope model

	Criteria	**Step model**	**Slope model**
1.	Knowledge about travel time distribution necessary	(-) Yes	(+) No
2.	Preferred arrival time	Fixed	Dependent on the travel duration
3.	Function additive over parts of the trip	(-) No	(+) Yes
4.	Treatment of departure and arrival time constraints	(-) Only arrival time constraints included	(+) Symmetric
5.	Sensitivity to travel time spread	(-) Low	(+) Higher
6.	Possibility to use VoSH and VoSR	(-) Only with a known travel time distribution	(+) Always

The criteria 1-4 are discussed by Fosgerau and Engelson (2011). Based on criteria 1, 3, 4 the slope model seems to be more advantageous in terms of implementation ease (1), application to links in a network (3) and inclusion of constraints in the origin (4). The second criterion, in the view of the authors of this paper, does not infer an advantage to either model type, but present the suitability of the models to different application cases. The step model explicitly includes a fixed preferred arrival time at the destination and models the inconvenience from unreliability as the costs of arriving too early or too late with respect to that time. The slope model does not include a fixed preferred arrival time, but it can be inferred given the travel duration. This leads to a flexible preferred arrival time. Both models could be representative of different work requirements and culture.

The fifth criterion was included inspired by the example given by Li et al. (2012). A service with a delay of 10 minutes in two trips would be evaluated in the same way as a service with a 5 minutes delay in one day and 15 minutes delay in another day using the step function; the step model cannot differentiate between these two services because the mean lateness is the same. This is apparent also in the expected utility function (see equation (2)). The expected slope function (see equation (11)) includes the σ^2 term, therefore the first service would be evaluated as superior to the second service by the slope function. In this way slope model displays higher sensitivity to travel time differences, which is one of the goals of a reliability valuation study.

The sixth criterion addresses the ease of computing $VoSH$ and $VoSR$. These variables, introduced by Benezech and Coulombel (2013), present a convenient way to assess and compare improvements in the mean and spread of headways. The $VoSH$ and $VoSR$ are defined as derivatives of the expected utility at an optimal departure time situation EU^* with respect to the mean and standard deviation of headways, respectively.

In the case of the step model and exponentially distributed headways, Benezech and Coulombel (2013) found that the $VoSH$ and $VoSR$ do not depend on the mean or standard deviation of headways separately, but only on their relationship in the coefficient of variation $cv_H = \frac{\sigma_H}{\mu_H}$. This finding is likely due to the exponential distribution being characterized by one parameter (rate). This characteristic enables a convenient visualisation of the functions in a graph (see Appendix A.1.). However, for other distributions the independence of mean or standard deviation of headways separately may not hold. In this case separate graphs would need to be created for any fixed μ_H and σ_H. Additionally, model adjustment is necessary for each travel time distribution.

For the slope model the independence of $VoSH$ and $VoSR$ from μ_H and σ_H separately does not hold as well. However, the functions can be easily computed for any known or unknown travel time distribution. This eliminates the work and potential errors of testing the distribution of headways and adopting the model accordingly. Equation 15 and 16 shows the derivation of the $VoSH$ and $VoSR$ from the slope function (14) [5].

$$\begin{gathered}\frac{dEU^*}{d\mu_w} = \alpha_w - \frac{\beta_1\gamma_1}{\beta_1-\gamma_1}(\mu_w + T_v)\\ \frac{dEU^*}{d\sigma_w} = -\gamma_1\sigma_w\end{gathered} \tag{15}$$

$$\begin{aligned}VoSH = \frac{dEU^*}{d\mu_H} &= \frac{dEU^*}{d\mu_w} * \frac{d\mu_w}{d\mu_H} + \frac{dEU^*}{d\sigma_w} * \frac{d\sigma_w}{d\mu_H}\\ &= \left(-\beta_0 - \frac{\beta_1\gamma_1}{\beta_1-\gamma_1}(\mu_v + \mu_w)\right) * \frac{1 - \frac{\sigma_H^2}{\mu_H^2}}{2}\\ &- \gamma_1\sqrt{\sigma_w}\left(\frac{\mu_H}{12} - \frac{\eta_H\sigma_H^3}{6\mu_H^2} - \frac{\sigma_H^4}{4\mu_H^3}\right)\end{aligned} \tag{16}$$

[5] The formulas connecting mean and standard deviation of waiting times with mean and standard deviation of headways are presented in Benezech and Coulombel (2013).

$$VoSR = \frac{dEU^*}{d\sigma_H} = \frac{dEU^*}{d\mu_w} * \frac{d\mu_w}{d\sigma_H} + \frac{dEU^*}{d\sigma_w} * \frac{d\sigma_w}{d\sigma_H}$$
$$= \left(-\beta_0 - \frac{\beta_1\gamma_1}{\beta_1 - \gamma_1}(\mu_v + \mu_w)\right)\frac{\sigma_H}{\mu_H} - \frac{\gamma_1\sqrt{\sigma_w}}{2}\left(\sigma_H + \frac{\eta_H\sigma_H^2}{\mu_H} - \frac{\sigma_H^3}{\mu_H^2}\right)$$

Similarly, it is possible to visualise the slope model $VoSH$ and $VoSR$ functions (16), see Figure 2.3[6]. The functions depend on the mean and standard deviation of headway separately, therefore different graphs are obtained if the mean headway is held constant and standard deviation varies (solid lines) and if the standard deviation is held constant and mean headway varies (dashed lines).[7]

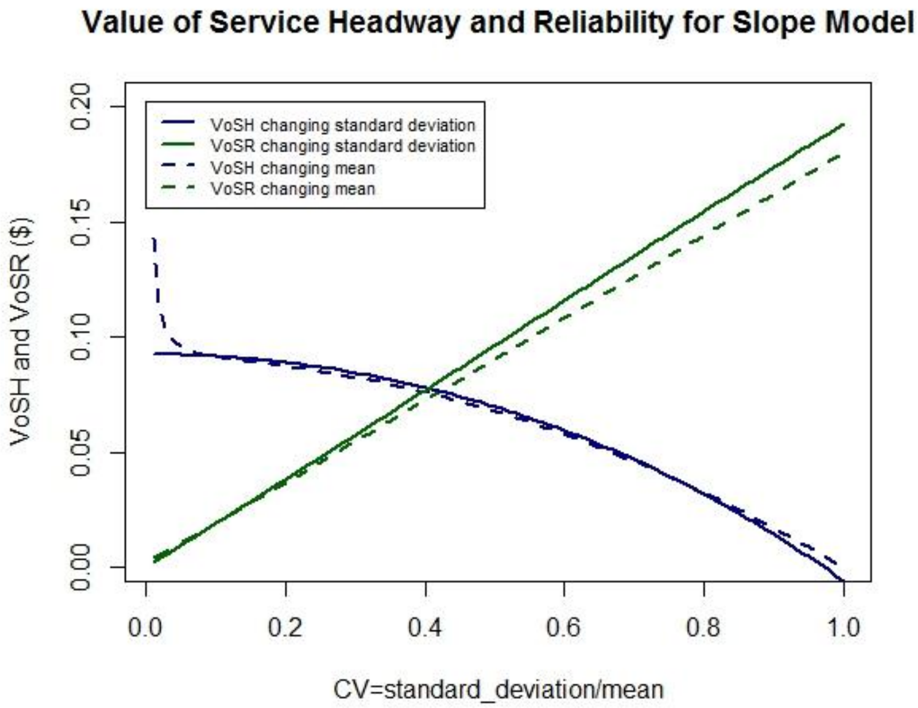


Figure 2.3 Value of Service Headway and Value of Service Reliability for the Slope Model

It seems from the overview above that slope model possesses many advantages compared to the step model and is convenient in applications. But to be able to use it, it is necessary that passenger preferences are indeed explained reasonably well by the model. Therefore it is desirable to test which of the models (step or slope) better approximates the average passenger perception. For this reason, a good survey design should be developed – a design, which can recover the parameters of both models. This will be done in the following chapter.

[6] Passenger preferences from Börjesson et al. (2012) are used for the creation of the graph.
[7] This explains the spike of the $VoSH$ dashed line for small cv_H. Standard deviation is set equal to 1 (minute). If, for example, $cv_H = 0.01$ then $\mu_H = 100$ (minutes), which would increase the $VoSH$ drastically.

3 Methodology for creation and testing of experimental design

A survey design has two dimensions: the mathematical experimental design and the visual design. The focus of this paper is the mathematical experimental design, although the visual design is not less important, and some studies have investigated the ways to visualise the reliability for the respondent (an in-depth study was done by Tseng et al., 2009). Specifically for this study, it would be important that the visual design allows the respondent to intuitively measure the components of step and slope model, including elements such as departure time, expected arrival time, earliness and lateness.

The remainder of this chapter is dedicated to presenting a step-by-step creation of the experimental design in the first section, and testing of the design in different simulated scenarios in the second section. The workflow of the methodology is visualised in Figure 3.1.

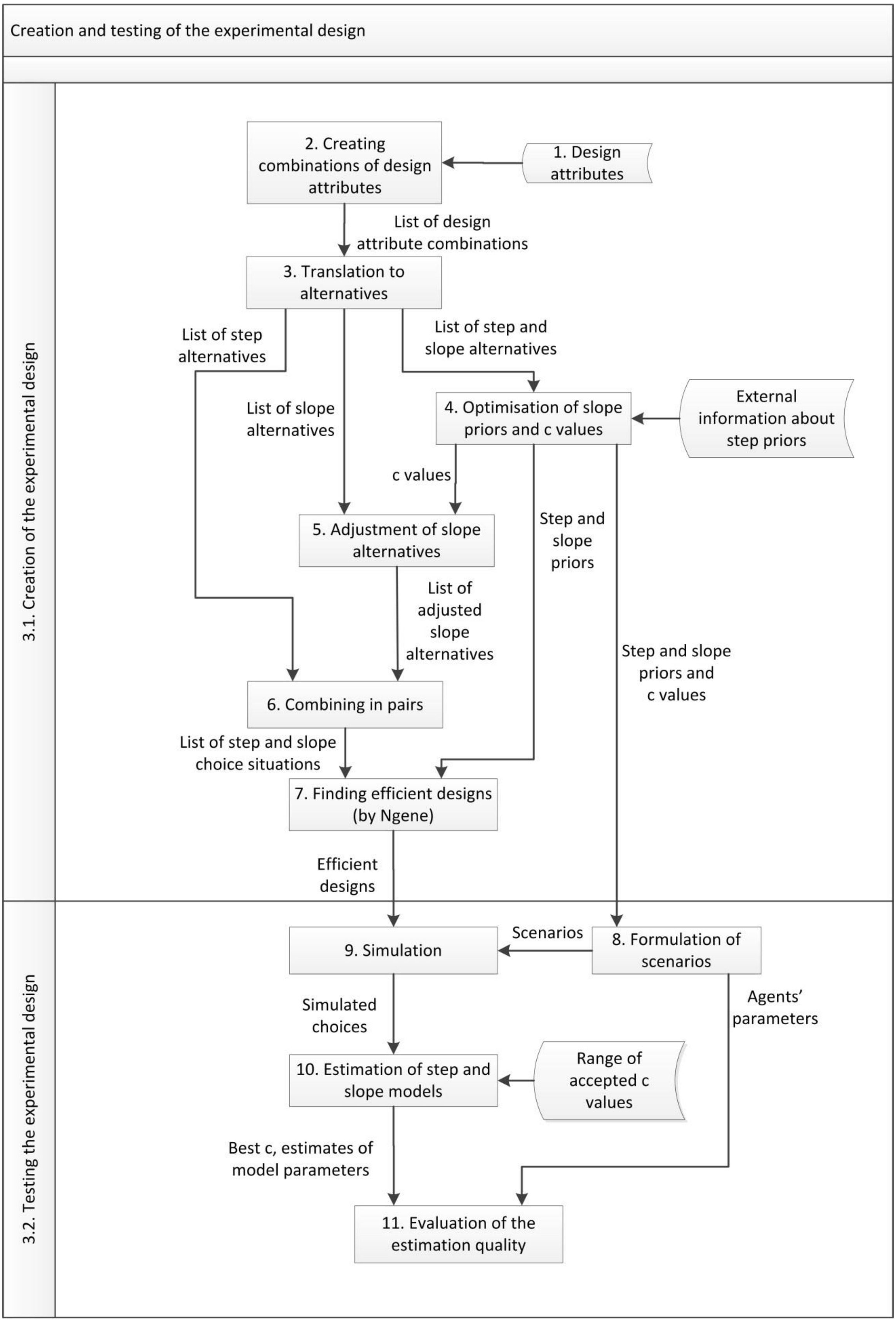


Figure 3.1 Workflow of design creation and testing

Creation of the experimental design

The creation of the experimental design is presented in this section (see steps 1 to 7 in Figure 3.1).

1. Design attributes

As mentioned in previous chapters, the expected utility functions (4) and (13) are used for creation of the design. However, it can be seen that the utility function components are not independent and unrelated as it is required by standard experimental design procedures. Instead they depend on other attributes, which are not evaluated by respondents; here they are named the design attributes.

Similar problem was encountered by Koster and Tseng (2010), who also created design for scheduling type functions. They developed a simulation procedure to find the best design that would minimise set efficiency criterions (S-efficiency and WTP-efficiency). The current study employs similar logic, but relies on Ngene software to find the best design, as will be explained later.

Given the dependency of utility function components on the design attributes, the first step in creating the design is specifying the levels for the design attributes. The information about possible levels can be taken from external sources or otherwise the levels need to be created artificially. In this study, the attribute levels are artificial, given in Appendix A.2. Each design attribute has three levels. Three sets of levels are created for all attributes containing questions with short, medium and long in-vehicle travel times. One of the three sets is assigned to a respondent based on the reported travel time. Therefore, the procedure includes partial adjustment to the respondent.

2. Creating combinations of design attributes

Set of all possible design attribute level combinations is created and filtered for fulfilment of three rules:

a) Earliest departure time should not be combined with shortest in-vehicle travel time, as it would result in too high levels of earliness;
b) Latest departure time should not be combined with longest in-vehicle travel time, as it would result in too high levels of lateness;
c) Shortest mean headway should not be combined with highest variability to avoid situations when $cv_H = \frac{\sigma_H}{\mu_H} > 1$. According to Singapore bus headway data[8], the coefficient of variation cv_H is typically less than 0.8.

[8] Bus headways in Singapore are obtained from the data of EZ-link card, the smart card of Singapore public transport users. Bus departure times in busy bus stops are extracted from the tap-in and tap-out times. From the departure times headways are obtained.

3. Translation to alternatives

A design attribute combination corresponds to a choice alternative. The translation is done via the equations (4) and (13) for the step and slope model respectively. For an overview, a table of relationships between design attributes and utility function components is presented in Table 3.1.

Table 3.1 Time related components of the utility functions and their design attributes

Utility function components		**Design attributes**						
		Departure time t	In-vehicle time T_v	Mean headway μ_H	Standard deviation of headway σ_H	Skewness of headway η_H[9]	Distribution of head-way ϕ_H[10]	Reference point for slope model c
Common	Mean waiting time μ_W			+	+			
	In-vehicle time T_V		+					
	Longest waiting time			+	+		+	
Step	Mean earliness	+	+	+	+		+	
	Mean lateness	+	+	+	+		+	
Slope	Squared departure time t^2	+						+
	Mean of squared arrival times $E(a^2)$	+	+	+	+	+		+

The first four design attributes are varied in the previous step. The next two – skewness and distribution of headway - are constant design attributes. The last attribute – reference point for slope model c – is defined as difference between the preferred arrival time of the step model t^* and the crossing point of home and work utility functions in the slope model. It serves both as a design attribute – because the utility function components depend on it – and a parameter, because its value is obtained by optimisation, as will be explained in step 4.

Additionally two more components are included in the experiment design – travel cost and longest displayed waiting time. The latter was implemented due to initial interviews with respondents. Interpretation of a sample questionnaire, which displayed several waiting times and their corresponding frequency, showed that the longest waiting time had a noticeable impact on the decision making.

[9] Skewness of the headway distribution is necessary for calculating the standard deviation of waiting times (formula can be found in Benezech and Coulombel, 2013), which in turn is used for calculating $E(a^2)$ in (13). The parameter is set equal to 0.6, a value which was obtained from Singapore bus headway data.
[10] Weibull distribution is used for calculating the mean earliness and lateness. The distribution was tested to be the closest headway distribution for a number of bus routes in Singapore.

Due to three sets of attribute levels as in step 1, resulting are three lists of possible choice alternatives, where each alternative is expressed in the step and slope model components, from here on called step alternatives and slope alternatives.

4. Optimisation of slope priors and c values

The priors are rough estimates of the utility function parameters and are necessary for creation of an efficient design, but approximate values are sufficient for that purpose (ChoiceMetrics, 2012).

It is important that the priors in both models correspond and are not in conflict; otherwise no feasible design can be found. In this study, the priors for the step model were taken from the study of Benezech and Coulombel (2013). The overlapping attributes in step and slope model – waiting and in-vehicle times, travel cost and longest displayed waiting time – received identical priors[11]. The priors for slope model parameters β_1, γ_1 and the reference point c were optimised for the best match with the step model for each travel time segment, because later the designs are found for the travel time segments separately.

Optimisation was done by minimizing the sum of differences of utilities from the same alternative expressed in the step and slope parameters. Excel Solver was used. Additionally, constraints on the parameters were imposed. The parameter β_1 was set to be negative because the slope for the utility at home was assumed to be negative; the parameter γ_1 is negative, because the slope for the utility at work is positive but the parameter has absorbed the minus sign from the utility function (13).

The values of priors and c are shown in Table 3.2. The names of the parameters correspond to the parameters in equations (4) and (13), with addition of ω for the travel cost and δ for the longest displayed waiting time.

Table 3.2 Priors for the step and slope models

For the common attributes				For step model attributes		For slope model attributes		
α_w	α_v	ω	δ	β	γ	β_1	γ_1	c
-2	-1	-3	-0.5	-0.8	-3	-0.003	-0.043	10

5. Adjustment of slope alternatives

The slope model attributes squared departure time and expected squared arrival time depend on c, which is optimised in the previous step. Therefore the attribute levels of slope alternative lists are updated accordingly.

[11] The prior for travel cost in Singapore was obtained from the study of Mueller (2013). To the knowledge of the authors, no study has evaluated the parameter for longest displayed waiting time. The values for these parameters were chosen such that the utility contribution from all attributes would be comparable for the designed alternatives in previous steps.

6. Combining in pairs

To obtain the choice situations, all possible combinations of two alternatives from the step and slope alternatives are created. Three lists with 8646 possible choice situations each are created, where each alternative is expressed in step and slope attributes. The dominating choice situations are filtered out.

7. Finding efficient designs (by Ngene)

The final step is supplying the priors and the 3 lists of possible choice situations to Ngene to find an efficient design for multinomial logit model for each travel time segment[12]. Supplying the choice situations of the step model and slope model separately, the optimal designs delivered by the program were different, and the best design for estimating the step model was unsuitable for estimating the slope model and vice versa. Therefore it was decided to supply choice situations and use model averaging property to find the best design. Doing this it was ensured that the alternatives, which were picked by the optimisation program, are reasonably suitable for estimating both step and slope models in an efficient way. The resulting designs for all three travel time segments are presented in the Appendix A.3.

Testing the experimental design

The goals of testing the experimental design in this study are twofold:

a) Assessing the ability of the design to recover the true parameters of respondents;

b) Assessing the ability of the design to reveal the underlying decision making model – step or slope.

The testing is done via a simulation, which is described next according to the steps 8 to 11 presented in Figure 3.1.[13]

8. Formulation of the scenarios

Four scenarios for simulation are defined. They can be characterised by a matrix as shown in Table 3.3. The scenarios are introduced in an order of increasing complexity and realism. The simplest scenarios – Scenarios 1 and 2 – are most useful to test the ability of the design to recover the true parameters of agents. This is done to assess the quality and adequacy of the procedure used for the creation of experimental designs. The last two scenarios serve the purpose of testing the ability of the design to differentiate between models which underlie the agents' preferences.

[12] This is possible in the newest update of the Ngene software (from 2015).
[13] All simulations and model estimations were performed in program R.

Table 3.3 Scenarios for the testing of the design

<table>
<tr><td colspan="2" rowspan="2"></td><th colspan="2">True parameters for agents</th></tr>
<tr><td>Original values – equal to priors of the efficient design</td><td>Parameters changed systematically</td></tr>
<tr><th rowspan="2">Mix of agents using step and slope models</th><td>No, homogeneous sample</td><td>Scenario 1
(2 cases)</td><td>Scenario 2
(26 cases)</td></tr>
<tr><td>Yes, mix in different proportions</td><td>Scenario 3
(100 cases)</td><td>Scenario 4
(90 cases)</td></tr>
</table>

Scenario 1 is obtained by assigning to 100% of the agents either the step or slope model with the parameters equal to the prior values used for the creation of the design. Scenario 3 has the same parameter values but is evaluated at 100 different mix situations between step and slope agents (1% step agents and 99% slope agents; 2% step agents and 98% slope agents and so on). Scenario 4 was evaluated at only five mix situations of step and slope agents: (10%;90%), (25%;75%), (50%;50%), (75%;25%) and (90%;10%).

The parameters for Scenario 2 and Scenario 4 are changed in a following way:

Each parameter is increased and decreased by 50% of its absolute value, i.e. increase of the parameter is understood as increase in its absolute magnitude and thus impact on the utility. The 50% value was selected as large enough to impact the behaviour significantly, yet maintain the balance among contributions of the attributes to the utility value. The number of cases in Scenario 2 is thus 26 (6 parameters for the step model and 7 parameters for the slope model, each increased and decreased).

In Scenario 4 it was assumed that the agents have similar preferences, even if they belong to different models. The similarity was implemented as adjustment of the opposite model's parameters to the changes in any parameter. The overlapping parameters in step and slope model (waiting time, in-vehicle time, travel cost and last waiting time) were changed together. The different parameters (earliness, lateness, squared departure time, squared expected arrival time and c) were optimised with Excel Solver (same as for finding the priors for the slope model) to reflect the change. This means that the number of cases for Scenario 4 is 90 (9 non-overlapping parameters for step and slope model, each increased and decreased for 5 scenarios).

The result of the optimisation was, for example, that decreased value of lateness leads to decreased value of c and vice versa. This connection is intuitive, because a traveller who penalises late arrival less would tend to depart and arrive later. Such behaviour would be interpreted by the slope model as movement of the crossing point of home and work functions to a later time, i.e., decreasing the distance to the preferred arrival time used in step model, which is measured with variable c. It should be noted also, that same parameters were assumed for all agents who have the same utility function.

9. Simulation

Having created the simulation scenarios and design, the simulation is performed by taking random draws from the alternative probabilities in each choice situation. Step-by-step the simulation procedure is following:

a) Assign the utility function (step or slope) and respective parameters to the agents;

b) Calculate logit probability of each alternative of a choice situation according to agent's parameters and utility function; create split of an interval accordingly. E.g., probability of the first alternative 0.3 would generate an interval split [0;0.3); [0.3;1];

c) Generate a random number between 0 and 1;

d) Assign the choice based on the interval where the random number belongs;

e) Repeat a-d for all choice questions and all agents. The number of agents is 50 from each travel time segment – a number, which is a reasonable sample size for the real survey;

f) Simulate each case of each scenario 10 times.

10. Estimation of step and slope models

Having the simulated choices, both step and slope models are estimated. Multinomial logit is used in all estimations. The estimation of slope model is more challenging than step model, because in reality the analyst would not know the true parameter c of the respondents. And with the current model formulation it also cannot be directly estimated from the choices, because the alternative levels of slope model depend on c. To deal with this problem, the simulated choices were paired with different sets of slope alternatives, which were obtained by varying the c. The c, which gave the best model in terms of log-likelihood was chosen.

It was assumed that c belongs to the interval $[-20; 50]$ meaning that the crossing point of the home and work functions could lie between 50 minutes before the preferred arrival time of step model and 20 minutes after the preferred arrival time. This interval was fixed because most of the optimised c values using this interval were lying within this interval and not equal the threshold values -20 and 50. In essence, a local optimum was found within this interval most of the time. The range of this interval also has a direct impact on the computational time.

11. Evaluation of the estimation quality

For each estimation (10 simulations and estimations were performed for each case of each scenario) three indicators describing the fit quality of estimators, as well as the fit of the model to the simulated choices were computed: mean absolute percentage error, maximum p-value and log-likelihood. They are presented in Table 3.4.

.

Table 3.4 Indicators describing model fit

Indicator and formula	**Comments**
Mean absolute percentage error (MAPE): $$\frac{1}{n}\sum_{i=1}^{n}\left\|\frac{true[i]-estim[i]}{true[i]}\right\|*100\%$$ n – number of parameters in the model; $true[i]$ – value of the i^{th} parameter assigned to the agent; $estim[i]$ – estimated value of the i^{th} parameter.	This error is especially important if the agents have homogeneous preferences (same utility function, same parameters). Then the estimates should be close to the true parameters. But it was applied also for not homogeneous scenarios. The mean absolute percentage error was chosen instead of the more common mean squared error, because the magnitude of the parameters is different. Therefore error as a percent of the true parameter value describes the impact of the error better.
Maximum p-value	This indicator describes if the experimental design and the assigned true parameters of agents allow finding parameter estimates that are all significantly different from 0 (based on any preferred significance level).
Log-likelihood	This parameter describes how well the model explains the choices made by the agents.

4 Results

The results from the simulations of the four scenarios are presented here.

The ability of the design to recover different parameters, if all agents either behave in step or slope way (Scenario 1 and Scenario 2), is illustrated in Figure 4.1. In Scenario 1, the MAPE is below 10% for step and below 25% for slope model for all 10 simulations (see the column of “Original” parameters). Full results for Scenario 1 are given in the Appendix A.4. For Scenario 2, the parameter changes sometimes lead to high discrepancies between the true parameters and estimations, for example, as in the case of increased c parameter. However, for the most cases the true parameters could be recovered reasonably well – the MAPE is below 30% most of the time.

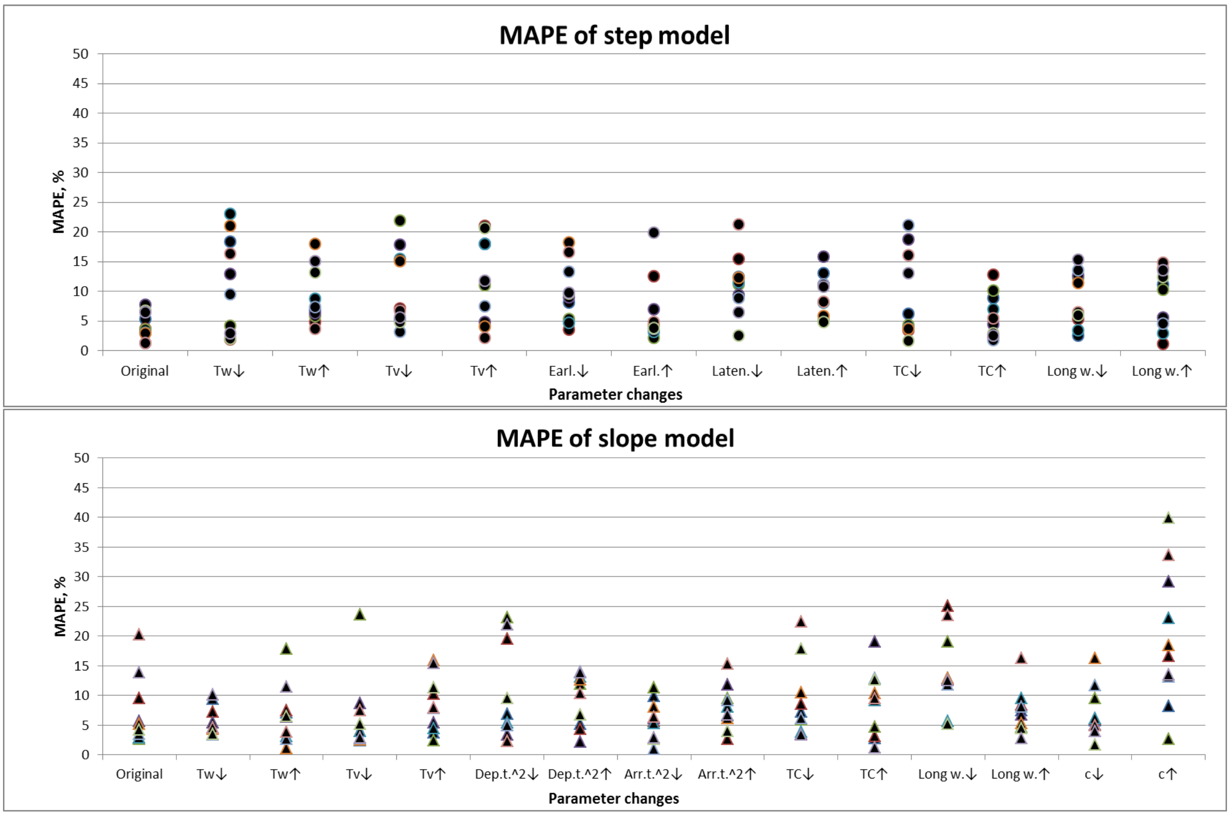


Figure 4.1 MAPE of step and slope model parameters in Scenarios 1 and 2

The ability of the design to distinguish between step and slope models with changed parameters, if the entire sample belongs to the step or slope model, is very good. This is illustrated in Figure 4.2 with differences of log-likelihood between the true model and the opposite model. Hence, positive difference means that the true model had a higher log-likelihood. It can be seen that the true model was recognised all of the time and mostly with a substantial difference.

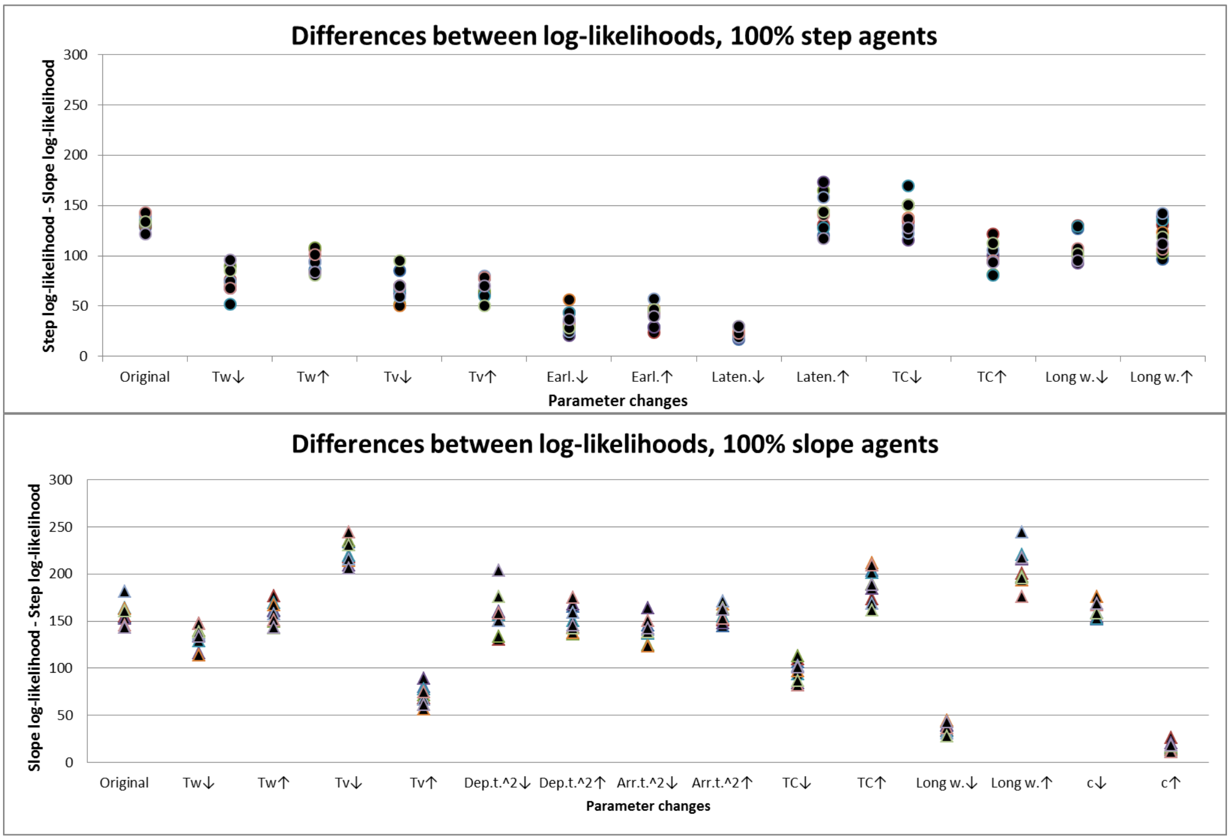


Figure 4.2 Log-likelihood of step and slope model in Scenarios 1 and 2

The ability of the design to recover the true parameters is lost quite quickly, once there is a mixture of step and slope agents (Scenario 3 and 4). This means that the preferences of step and slope model agents are very different, even having the parameters optimised for maximum match. In the Appendix A.4., the full result is presented for mixtures of 90% step agents and 10% slope agents and vice versa, if the agents have the original parameters (cases from Scenario 3). The ability to recover the true parameters in presence of small mixtures (10% of the opposite model agents) is stronger impaired for the slope model – MAPE goes up to 39% on average, while the step model still has all MAPE values below 30%. However, these errors are due to the decrease in the magnitude of the estimates once a mixture is introduced; the relationships between the estimates are still reasonably well maintained. This indicates that the MAPE criterion is also not an ideal one for describing the estimation quality.

However, even if the parameters cannot be recovered well, the design still maintains a very good ability to differentiate between agents of step and slope model. In Figure 4.3 the average log-likelihood of step and slope models are presented, calculated over 10 simulations for each mixture (1% step and 99% slope, 2% step and 98% slope, and so on). The step model would be recognised superior, if the percentage of step is larger than 40%. The slope model is recognised superior if the share of slope agents exceeds 60%.

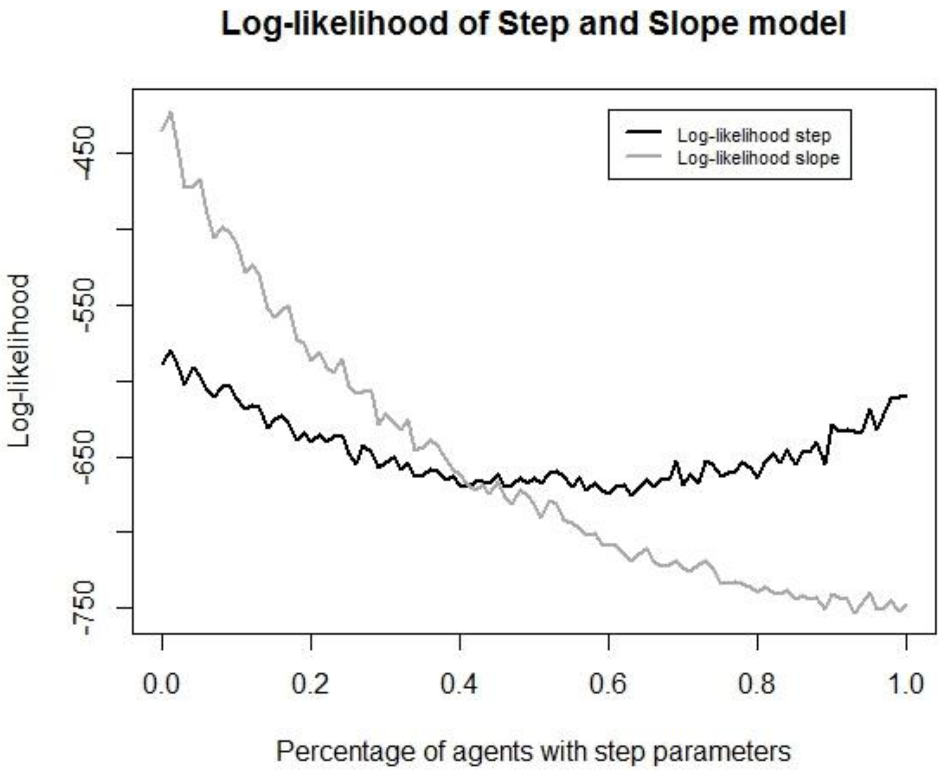


Figure 4.3 Log-likelihood of step and slope model in Scenario 3

Finally, the design was tested for the quality in the most realistic case – mixture of step and slope models with varied parameters. Due to a large number of combinations, the results are summarised concisely in Table 4.1. For each step and slope mix scenario (lines in the table), all the parameter change cases including the original parameters were simulated, each 10 times. The percentages in the table correspond to the cases when the results fulfil the respective criteria in the columns. I.e., 100% corresponds to all scenarios fulfilling the criteria and 0% corresponds to none of the scenarios fulfilling the criteria.

Table 4.1 Summary of testing results in Scenario 4

Scenario:	MAPE<=50		Max p-value<=0.05		Loglik relation	Expected signs	
step %	Step	Slope	Step	Slope		Step	Slope
0.1	56%	87%	92%	97%	99%	95%	89%
0.25	69%	38%	92%	86%	100%	95%	85%
0.5	87%	11%	94%	78%	46%	97%	67%
0.75	97%	1%	98%	57%	94%	99%	25%
0.9	100%	1%	97%	42%	100%	100%	16%

As discussed before, the design quickly loses the capability to recover the true parameters, which can be seen as low percentage of cases with MAPE lower than 50% especially for the slope model. The columns of p-values show the percentage of cases, where all model parameters could be estimated significantly at a 5% significance level. The results indicate that significance of parameters is generally higher for the step model, which could be due to higher absolute parameter values in the step model. As could be expected, the number of significantly estimated models increases with the percentage of the corresponding model in the scenario. The log-likelihood relation column indicates the percentage of cases when the model of the majority of agents also showed a higher log-likelihood of the estimates. The results are good for 10% and 25% mix situations; in case of the 50%:50% split the slope model was recognised as superior 46% of the time. Finally, an alternative criterion for the estimation quality of the model is the ability to recover the correct signs of the parameters. Interestingly, the step estimates maintain the correct sign much of the time even if the majority of agents have slope preferences. The slope model estimates are much more unstable even for small percentages of step agents.

Overall, it can be argued that the design performs reasonably well, especially with respect to the second goal – recognising the underlying decision model. But, it should be noted that different results are possible if different efficient designs are used. Therefore it is very important to simulate the design before carrying out a survey, so that later the application of both models would be possible.

5 Discussion

Along the way, several assumptions and simplifications are done that should be highlighted. Firstly, an assumption in the testing of the design is that all agents who have the same utility function also have the same parameter values. Clearly, this is unlikely in a realistic case; more simulations could have been done to test the performance of the design under different conditions.

Secondly, assumption about maximum possible offsets of the slope model with respect to the step model (parameter c) has to be made in order to estimate the slope model in the present formulation. Although it is possible to cover a wide range of c values, it would be desirable in future work to reformulate the model to avoid this procedure.

Thirdly, a fundamental simplification in this work is waiting time being separated from the in-vehicle travel time, while the latter is held constant. This decision has both advantages and disadvantages. On the good side, such separation results in data requirements, which are often easier to fulfil. Waiting time reliability can be obtained from headway reliability, while total travel time reliability would likely involve travel time surveys of a large number of travellers. Also, the separation is likely to benefit the operationalisation of the results. On the negative side, it is known that variable waiting times cause variable in-vehicle times due to uneven load of passengers. Therefore the full effect of changing the reliability is likely to be underestimated in this work. Also a design with an assumption of perfectly reliable in-vehicle times is likely to cause incredulity in the stated preference survey respondents. However, this is a necessary simplification for the sake of carrying out the stated preference survey. In future work another survey could be carried out, which would test the effects of variable in-vehicle time.

Finally, a simplified design was created for the simulation of the results. In practice, it is usually necessary to include socio-economic variables and estimate the results for different segments of the population. A full simulation should be performed to ensure that the design maintains a satisfactory quality in such case.

6 Conclusion

This research addresses the question of deciding between step and slope model for reliability measurement. This decision may be important for several reasons: it may help to describe the true passenger preferences better and it may reduce the workload of the analysts in case the slope model is true, because no knowledge about travel time distribution is necessary in such case. This paper contributes to the solution of the decision problem by creating an experimental design, which is capable of estimating both models. A methodology is described that completes the picture of necessary steps to be taken for creation of such designs. Thanks to the recent advances of the Ngene software, such complex designs can now be created more easily and reliably. Results show that the obtained design could withstand several tests of different response scenarios.

Further research could be conducted to develop designs for other utility specifications. For example, a combination of step and slope model elements in one utility function could be of interest. A successful implementation of the survey, which allows reliable estimation of several utility functions, may open the doors for better understanding of the passenger preferences and perception of reliability. This in turn, can help the practitioners to address the problem of unreliability in the most efficient way. Finally, similar methodology may be applicable also in other cases and fields where the underlying model of travellers' or customers' preferences is unknown.

Acknowledgements

This work was financially supported by the Singapore National Research Foundation under its Campus for Research Excellence And Technological Enterprise (CREATE) programme. The views expressed herein are solely the responsibility of the authors and do not necessarily represent the official views of the Foundation.

The authors would like to thank ChoiceMetrics team and specifically Dr. Andrew Collins and Prof. Michiel Bliemer for providing the newest Ngene update.

The first author would like to specifically thank Tharsis Teoh for continuous support in all stages of this research.

References

Bates, J., Polak, J., Jones, P., & Cook, A. (2001). The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review*, *37*(2-3), 191-229.

Benezech, V., & Coulombel, N. (2013). The value of service reliability. *Transportation Research Part B: Methodological*, *58*(2013), 1-15.

Börjesson, M., Eliasson, J., & Franklin, J. P. (2012). Valuations of travel time variability in scheduling versus mean–variance models. *Transportation Research Part B: Methodological*, *46*(7), 855-873.

Carrion, C., & Levinson, D. (2012). Value of travel time reliability: A review of current evidence. *Transportation Research Part A: Policy and Practice*, *46*(4), 720-741.

ChoiceMetrics (2012) Ngene 1.1.1 User Manual and Reference Guide. http://www.choice-metrics.com/documentation.html

Fosgerau, M., & Karlstrom, A. (2010). The value of reliability. *Transportation Research Part B: Methodological*, *44*(1), 38-49.

Fosgerau, M., & Engelson, L. (2011). The value of travel time variance. *Transportation Research Part B: Methodological*, *45*(1), 1-8.

Jou, R. C., Kitamura, R., Weng, M. C., & Chen, C. C. (2008). Dynamic commuter departure time choice under uncertainty. *Transportation Research Part A: Policy and Practice*, *42,* 774–783.

Koster P., & Tseng, Y.-Y. (2010). Stated choice experimental designs for scheduling models. *Choice Modelling: The State-of-the-art and the State-of-practice : Proceedings from the Inaugural International Choice Modelling Conference*

Li, Z., Tirachini, a., & Hensher, D. a. (2012). Embedding Risk Attitudes in a Scheduling Model: Application to the Study of Commuting Departure Time. *Transportation Science*, *46*(July 2015), 170–188.

Mueller S. (2013). Revisiting the First-Last Mile Problem in the Singaporean Context and Evaluating the Potential Net User Benefit of both Innovative and Practical Solutions by way of a Discrete Choice Model. *Master thesis*

Noland, R. B., Small, K. a., Koskenoja, P. M., & Chu, X. (1998). Simulating travel reliability. *Regional Science and Urban Economics*, *28*(5), 535-564.

Small, K., (1982). The scheduling of consumer activities: Work trips. *American Economic Review 72*(3), 467-479.

Tseng, Y.-Y., Verhoef, E., de Jong, G., Kouwenhoven, M., & van der Hoorn, T. (2009). A pilot study into the perception of unreliability of travel times using in-depth interviews. *Journal of Choice Modelling*, *2*(1), 8-28.

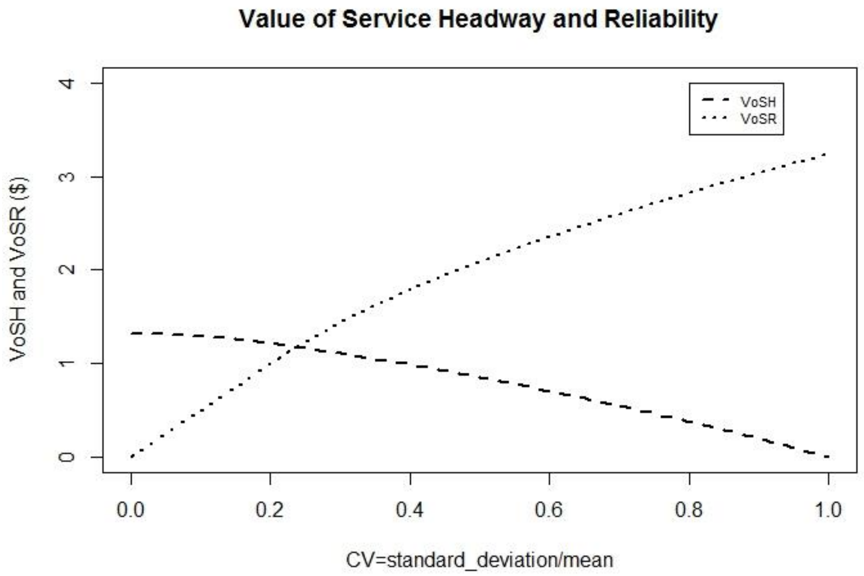
Vickrey, W.S. (1969). Congestion theory and transport investment. *American Economic Review*, *59*(2), 251-260.

Vickrey, W.S. (1973). Pricing, metering, and efficiently using urban transportation facilities. *Highway Research Record 476*, 36–48.

Vincent, M., & Hamilton, B. A. (2008). Measurement valuation of public transport reliability. *Land Transport New Zealand Research Report 339.*

Appendix

Value of Service Headway and Reliability in Exponential case of Step model



Levels for design attributes

Travel time code	Departure time (min)[14]	Mean headway (min)	Standard deviation of headway (min)	Travel cost (SGD)	In-vehicle travel time (min)
1 - respondents with short reported travel time (≤15 min)	-10 -15 -20	2 5 7	1 2 4	0.5 0.9 1.3	5 10 15
2 - respondents with medium long reported travel time (20-40 min)	-30 -40 -55	4 7 10	2 4 7	0.8 1.4 2	20 30 40
3 - respondents with long reported travel time (≥45 min)	-60 -70 -80	6 10 13	3 5 8	1 1.7 2.4	45 60 70

[14] Here preferred arrival time is set to 0, but with the reported preferred arrival time this column is shifted accordingly.

Efficient designs for all travel time segments

Short travel time (travel time < 20 min)

D error	0.06
A error	6.33
B estimate	62.35
S estimate	1136.03

D error	0.22
A error	3.22
B estimate	77.07
S estimate	3.48

Prior	b1	b2	b3	b4	b5	b6
Fixed prior value	-2	-1	-0.003	-0.043	-3	-0.5
Sp estimates	9.48	14.24	1136.03	8.69	10.15	10.01
Sp t-ratios	0.64	0.52	0.06	0.66	0.62	0.62

Prior	b7	b8	b9	b10	b11	b12
Fixed prior value	-2	-1	-0.8	-3	-3	-0.5
Sp estimates	2.95	2.65	2.82	2.87	3.48	3.03
Sp t-ratios	1.14	1.2	1.17	1.16	1.05	1.13

Design	Alternative 1 (slope)						Alternative 2 (slope)						Alternative 3 (step)						Alternative 4 (step)					
Choice situation	Tw	Tv	Dep 2	Arr 2	Tc	Long w	Tw	Tv	Dep 2	Arr 2	Tc	Long w	Tw	Tv	Early	Late	Tc	Long w	Tw	Tv	Early	Late	Tc	Long w
1	3	5	25	13.6	1.3	5	0.5	10	100	1.1	1.3	5	3	5	7	0	1.3	5	0.5	10	9.5	0	1.3	5
2	3	15	25	173.6	0.9	5	4.5	10	25	100.3	1.3	15	3	15	0	3	0.9	5	4.5	10	2	1.5	1.3	15
3	0.5	10	100	1.1	0.9	5	3	10	25	68.6	1.3	5	0.5	10	9.5	0	0.9	5	3	10	2	0	1.3	5
4	0.5	15	25	111.1	0.5	5	1.5	5	25	4.4	1.3	5	0.5	15	0	0.5	0.5	5	1.5	5	8.5	0	1.3	5
5	2.5	10	0	160.5	0.5	5	0.5	15	100	31.1	1.3	5	2.5	10	0	2.5	0.5	5	0.5	15	4.5	0	1.3	5
6	2.5	10	0	158.9	1.3	5	4.5	10	100	32.9	0.5	10	2.5	10	0	2.5	1.3	5	4.5	10	5.5	0	0.5	10
7	2.5	10	0	160.5	1.3	5	4	10	25	87.2	0.5	10	2.5	10	0	2.5	1.3	5	4	10	1.5	0.5	0.5	10
8	4.5	5	0	100.3	0.5	15	4.5	10	100	32.9	0.5	10	4.5	5	2	1.5	0.5	15	4.5	10	5.5	0	0.5	10

Medium travel time (travel time 20 - 40 min)

D error	0.01
A error	3.85
B estimate	54.56
S estimate	7.96

D error	0.09
A error	3.89
B estimate	80.15
S estimate	5.23

Prior	b1	b2	b3	b4	b5	b6
Fixed prior value	-2	-1	-0.003	-0.043	-3	-0.5
Sp estimates	7.96	5.25	4.89	5.2	5.62	4.19
Sp t-ratios	0.69	0.86	0.89	0.86	0.83	0.96

Prior	b7	b8	b9	b10	b11	b12
Fixed prior value	-2	-1	-0.8	-3	-3	-0.5
Sp estimates	3.4	3.46	3.42	3.5	4.18	5.23
Sp t-ratios	1.06	1.05	1.06	1.05	0.96	0.86

Design	Alternative 1 (slope)						Alternative 2 (slope)						Alternative 3 (step)						Alternative 4 (step)					
Choice situation	Tw	Tv	Dep 2	Arr 2	Tc	Long w	Tw	Tv	Dep 2	Arr 2	Tc	Long w	Tw	Tv	Early	Late	Tc	Long w	Tw	Tv	Early	Late	Tc	Long w
1	5.5	30	400	257.2	2	15	7.5	30	2025	89.9	0.8	20	5.5	30	0	5.5	2	15	7.5	30	17.5	0	0.8	20
2	2	30	2025	172.5	0.8	5	5.5	30	900	40.7	2	10	2	30	23	0	0.8	5	5.5	30	4.5	0	2	10
3	2	40	2025	12.5	0.8	5	2	30	2025	172.5	2	5	2	40	13	0	0.8	5	2	30	23	0	2	5
4	5.5	30	400	257.2	0.8	15	5.5	30	2025	100.7	2	10	5.5	30	0	5.5	0.8	15	5.5	30	19.5	0	2	10
5	2	30	2025	172.5	1.4	5	7.5	20	900	39.9	1.4	20	2	30	23	0	1.4	5	7.5	20	12.5	0	1.4	20
6	3	30	2025	152.5	2	10	5.5	40	2025	10.7	0.8	10	3	30	22	0	2	10	5.5	40	9.5	0	0.8	10
7	5.5	30	400	250.7	1.4	10	3	40	2025	12.5	2	10	5.5	30	0	5.5	1.4	10	3	40	12	0	2	10
8	2	40	900	147.5	2	5	7.5	20	900	39.9	0.8	20	2	40	0	2	2	5	7.5	20	12.5	0	0.8	20

Long travel time (travel time > 40 min)

D error	0.02
A error	11.39
B estimate	38.52
S estimate	26.11

D error	0.10
A error	3.99
B estimate	82.38
S estimate	4.80

Prior	b1	b2	b3	b4	b5	b6
Fixed prior value	-2	-1	-0.003	-0.043	-3	-0.5
Sp estimates	26.11	12.62	12.06	16.66	15.87	10.87
Sp t-ratios	0.38	0.55	0.56	0.48	0.49	0.59

Prior	b7	b8	b9	b10	b11	b12
Fixed prior value	-2	-1	-0.8	-3	-3	-0.5
Sp estimates	3.45	3.56	3.63	3.75	4.16	4.8
Sp t-ratios	1.06	1.04	1.03	1.01	0.96	0.89

Design	Alternative 1 (slope)						Alternative 2 (slope)						Alternative 3 (step)						Alternative 4 (step)					
Choice situation	Tw	Tv	Dep 2	Arr 2	Tc	Long w	Tw	Tv	Dep 2	Arr 2	Tc	Long w	Tw	Tv	Early	Late	Tc	Long w	Tw	Tv	Early	Late	Tc	Long w
1	7.5	45	2500	70.9	1	20	4	45	3600	140.2	2.4	10	7.5	45	8	0.5	1	20	4	45	21	0	2.4	10
2	5.5	60	2500	259.1	1.7	10	5.5	60	4900	53.7	2.4	15	5.5	60	0	5.5	1.7	10	5.5	60	14.5	0	2.4	15
3	3.5	60	2500	190.1	2.4	10	5.5	60	4900	39.1	1	10	3.5	60	0	3.5	2.4	10	5.5	60	14.5	0	1	10
4	4	70	3600	215.2	1	10	7.5	60	3600	120.9	2.4	20	4	70	0	4	1	10	7.5	60	4	1.5	2.4	20
5	3.5	45	3600	140.1	2.4	10	5.5	60	3600	49.1	1	10	3.5	45	21.5	0	2.4	10	5.5	60	4.5	0	1	10
6	7.5	45	3600	120.9	1	20	5.5	60	3600	49.1	2.4	10	7.5	45	17.5	0	1	20	5.5	60	4.5	0	2.4	10
7	3.5	60	4900	50.1	1.7	10	8.5	45	3600	129.3	1.7	20	3.5	60	16.5	0	1.7	10	8.5	45	16.5	0	1.7	20
8	3.5	70	3600	190.1	1	10	5.5	60	4900	39.1	2.4	10	3.5	70	0	3.5	1	10	5.5	60	14.5	0	2.4	10

Simulations of the survey using the generated efficient designs

Scenario 1. 100% agents with step parameters:

Assigned true parameters to the agents					
Wait. time	In-veh. time	Earl.	Laten.	Trav. Cost	Long. wait. time
-2	-1	-0.8	-3	-3	-0.5

						Parameter estimates of slope model							Parameter estimates of step model					
Surv. no.	Mape	Max. p-value step	Max. p-value slope	Log-likelih. step	Log-likelih. slope	Wait. time	In-veh. time	Dep. time sq.	Arr. time sq.	Trav. cost	Long wait. time	c	Wait. time	In-veh. time	Earl.	Laten.	Trav. cost	Long. wait. time
1	5.4%	0.00	0.56	-603.18	-731.55	-0.316	-0.158	0	-0.008	-0.864	-0.086	6	-1.901	-0.934	-0.75	-2.841	-2.911	-0.468
2	3.7%	0.00	0.44	-585.61	-715.61	-0.315	-0.178	0	-0.009	-0.842	-0.108	6	-1.906	-0.957	-0.781	-2.92	-2.86	-0.484
3	3.6%	0.00	0.05	-634.15	-765.36	-0.338	-0.17	0	-0.008	-0.684	-0.068	5	-1.962	-0.964	-0.778	-2.908	-2.892	-0.466
4	7.8%	0.00	0.00	-624.66	-757.80	-0.278	-0.159	0	-0.008	-0.704	-0.088	7	-1.846	-0.932	-0.743	-2.79	-2.753	-0.451
5	3.0%	0.00	0.00	-601.69	-737.13	-0.338	-0.202	0	-0.01	-0.827	-0.139	7	-2.023	-1.037	-0.815	-3.047	-3.082	-0.535
6	3.0%	0.00	0.86	-624.86	-767.08	-0.26	-0.17	0	-0.007	-0.605	-0.096	5	-1.911	-0.989	-0.793	-2.92	-2.751	-0.502
7	1.2%	0.00	0.00	-612.58	-751.79	-0.276	-0.159	0	-0.008	-0.746	-0.117	7	-1.991	-1.006	-0.798	-2.939	-2.947	-0.512
8	1.3%	0.00	0.00	-613.77	-756.82	-0.266	-0.139	0	-0.008	-0.693	-0.085	7	-2	-0.995	-0.804	-2.997	-2.867	-0.489
9	6.7%	0.00	0.02	-620.88	-754.96	-0.28	-0.166	0	-0.006	-0.734	-0.082	4	-1.86	-0.938	-0.741	-2.754	-2.814	-0.473
10	6.5%	0.00	0.00	-608.48	-730.40	-0.348	-0.209	0	-0.007	-0.816	-0.083	3	-1.877	-0.95	-0.746	-2.736	-2.848	-0.463

Scenario 1. 100% agents with slope parameters:

												Assigned true parameters to the agents						
												Wait. time	**In-veh. time**	**Dep. time sq.**	**Arr. time sq.**	**Trav. cost**	**Long wait. time**	**c**
												-2	-1	-0.003	-0.043	-3	-0.5	10
						Parameter estimates of step model						**Parameter estimates of slope model**						
Surv. no.	**Mape**	**Max. p-value slope**	**Max. p-value step**	**Log-likelih. slope**	**Log-likelih. step**	**Wait. time**	**In-veh. time**	**Earl.**	**Laten.**	**Trav. cost**	**Long. wait. time**	**Wait. time**	**In-veh. time**	**Dep. time sq.**	**Arr. time sq.**	**Trav. cost**	**Long wait. time**	**c**
1	3.0%	0.00	0.00	-416.63	-571.10	-1.281	-0.654	-0.305	-1.312	-1.816	-0.251	-2.117	-1.043	-0.003	-0.045	-2.945	-0.499	10
2	9.6%	0.00	0.00	-441.31	-596.80	-1.1	-0.551	-0.242	-1.058	-1.63	-0.21	-1.831	-0.901	-0.003	-0.04	-2.627	-0.452	10
3	2.7%	0.00	0.00	-430.42	-593.71	-1.105	-0.553	-0.234	-1.03	-1.654	-0.201	-2.058	-0.997	-0.003	-0.043	-2.968	-0.468	10
4	5.7%	0.00	0.00	-421.75	-574.58	-1.001	-0.55	-0.218	-0.964	-1.54	-0.218	-1.86	-0.967	-0.003	-0.041	-2.693	-0.491	10
5	2.9%	0.00	0.00	-409.54	-573.79	-1.238	-0.655	-0.324	-1.345	-1.928	-0.281	-1.897	-0.956	-0.003	-0.043	-2.979	-0.516	10
6	5.2%	0.00	0.00	-425.31	-588.90	-0.945	-0.488	-0.181	-0.824	-1.504	-0.195	-1.904	-0.942	-0.003	-0.041	-2.81	-0.488	10
7	3.5%	0.00	0.00	-411.34	-592.86	-1.193	-0.613	-0.295	-1.182	-1.724	-0.216	-2.017	-0.992	-0.003	-0.044	-2.926	-0.485	10
8	20.3%	0.00	0.00	-454.82	-600.85	-0.938	-0.523	-0.226	-0.931	-1.568	-0.215	-1.471	-0.775	-0.003	-0.034	-2.398	-0.425	10
9	4.2%	0.00	0.00	-415.45	-575.93	-1.154	-0.587	-0.254	-1.118	-1.703	-0.215	-2.142	-1.066	-0.003	-0.045	-3.116	-0.507	10
10	13.9%	0.00	0.00	-452.47	-596.01	-1.1	-0.589	-0.267	-1.112	-1.729	-0.24	-1.65	-0.855	-0.003	-0.036	-2.595	-0.45	10

Scenario 3. 90% agents with step parameters:

							Assigned true step parameters to the agents						Assigned true slope parameters to the agents						
							Wait. time	**In-veh. time**	**Earl.**	**Laten.**	**Trav. cost**	**Long wait. time**	**Wait. time**	**In-veh. time**	**Dep. time sq.**	**Arr. time sq.**	**Trav. cost**	**Long wait. time**	**c**
							-2	-1	-0.8	-3	-0.8	-0.5	-2	-1	-0.01	-0.05	-0.8	-0.5	11
							Parameter estimates of step model						**Parameter estimates of slope model**						
Survey no.	**Mape step**	**Mape slope**	**Max. p-value step**	**Max. p-value slope**	**Log-likelih. step**	**Log-likelih. slope**	**Wait. time**	**In-veh. time**	**Earl.**	**Laten.**	**Trav. cost**	**Long wait. time**	**Wait. time**	**In-veh. time**	**Dep. time sq.**	**Arr. time sq.**	**Trav. cost**	**Long wait. time**	**c**
1	16.2%	82.1%	0.00	0.00	-637.94	-740.93	-1.57	-0.782	-0.547	-2.205	-0.765	-0.406	-1.601	-0.556	0.001	-0.01	-0.437	-0.301	15
2	22.0%	81.7%	0.00	0.00	-656.39	-739.21	-1.567	-0.864	-0.873	-2.725	-0.524	-0.428	-1.558	-0.507	0.001	-0.01	-0.184	-0.258	15
3	3.5%	81.1%	0.00	0.00	-613.75	-750.94	-1.731	-0.878	-0.74	-2.629	-0.82	-0.421	-1.941	-0.616	0	-0.014	-0.536	-0.305	16
4	25.8%	83.0%	0.00	0.14	-653.96	-738.00	-1.747	-0.91	-0.799	-2.668	-0.746	-0.446	-1.707	-0.564	0	-0.011	-0.385	-0.285	15
5	17.7%	82.2%	0.00	0.00	-644.95	-750.11	-1.568	-0.757	-0.592	-2.214	-0.601	-0.365	-1.704	-0.553	0	-0.013	-0.343	-0.275	14
6	20.3%	82.9%	0.00	0.00	-662.22	-766.16	-1.491	-0.787	-0.618	-2.24	-0.695	-0.397	-1.605	-0.545	0	-0.011	-0.44	-0.295	16
7	16.1%	84.6%	0.00	0.70	-650.38	-759.10	-1.543	-0.839	-0.889	-2.675	-0.688	-0.409	-1.424	-0.444	0.001	-0.007	-0.268	-0.221	17
8	12.9%	81.5%	0.00	0.00	-637.21	-747.59	-1.519	-0.854	-0.935	-2.69	-0.811	-0.415	-1.414	-0.462	0.001	-0.007	-0.373	-0.229	16
9	15.9%	82.8%	0.00	0.48	-625.87	-727.38	-1.515	-0.801	-0.7	-2.436	-0.599	-0.384	-1.552	-0.519	0	-0.011	-0.294	-0.259	15
10	12.0%	81.6%	0.00	0.00	-634.22	-743.40	-1.467	-0.724	-0.534	-2.092	-0.64	-0.37	-1.29	-0.403	0	-0.007	-0.316	-0.233	19

Scenario 3. 90% agents with slope parameters:

							Assigned true step parameters to the agents						Assigned true slope parameters to the agents						
							Wait. time	In-veh. time	Earl.	Laten.	Trav. cost	Long wait. time	Wait. time	In-veh. time	Dep. time sq.	Arr. time sq.	Trav. cost	Long wait. time	c
							-2	-1	-0.8	-3	-3	-0.5	-2	-1	-0.003	-0.043	-3	-0.5	10
							Parameter estimates of step model						Parameter estimates of slope model						
Survey no.	**Mape step**	**Mape slope**	**Max. p-value step**	**Max. p-value slope**	**Log-likelih. step**	**Log-likelih. slope**	**Wait. time**	**In-veh. time**	**Earl.**	**Laten.**	**Trav. cost**	**Long wait. time**	**Wait. time**	**In-veh. time**	**Dep. time sq.**	**Arr. time sq.**	**Trav. cost**	**Long wait. time**	**c**
1	45.2%	37.0%	0.00	0.00	-610.92	-515.07	-1.204	-0.636	-0.335	-1.392	-1.824	-0.278	-1.124	-0.599	-0.002	-0.027	-1.851	-0.352	10
2	56.1%	42.9%	0.00	0.00	-619.18	-525.10	-0.998	-0.522	-0.243	-1.035	-1.563	-0.221	-1.071	-0.552	-0.002	-0.024	-1.762	-0.311	10
3	51.1%	36.3%	0.00	0.00	-634.09	-519.20	-1.086	-0.576	-0.299	-1.262	-1.632	-0.238	-1.138	-0.6	-0.002	-0.029	-1.838	-0.343	10
4	57.2%	44.3%	0.00	0.00	-623.70	-523.67	-0.971	-0.519	-0.236	-1.025	-1.48	-0.218	-0.949	-0.495	-0.002	-0.022	-1.627	-0.303	11
5	54.3%	27.2%	0.00	0.00	-609.90	-481.69	-1.056	-0.564	-0.263	-1.113	-1.57	-0.213	-1.402	-0.721	-0.002	-0.032	-2.102	-0.367	10
6	53.9%	44.9%	0.00	0.00	-609.29	-520.51	-0.962	-0.545	-0.249	-1.107	-1.633	-0.258	-0.861	-0.494	-0.002	-0.022	-1.675	-0.329	11
7	50.6%	38.7%	0.00	0.00	-609.60	-512.06	-1.215	-0.589	-0.295	-1.214	-1.728	-0.209	-1.276	-0.6	-0.002	-0.026	-1.915	-0.282	10
8	53.8%	41.8%	0.00	0.00	-618.34	-522.39	-1.103	-0.554	-0.272	-1.107	-1.653	-0.202	-1.138	-0.56	-0.002	-0.024	-1.834	-0.284	10
9	47.6%	39.3%	0.00	0.00	-605.48	-516.07	-1.188	-0.619	-0.311	-1.325	-1.763	-0.256	-1.144	-0.595	-0.002	-0.026	-1.852	-0.323	10
10	45.8%	40.1%	0.00	0.00	-590.58	-508.56	-1.231	-0.641	-0.326	-1.337	-1.856	-0.263	-1.136	-0.582	-0.002	-0.025	-1.892	-0.312	10