

Stated preference survey for estimating passenger transfer penalties: design and application to Madrid

Rocio Cascajo¹  · Andres Garcia-Martinez² · Andres Monzon³

Received: 23 December 2016 / Accepted: 21 July 2017 / Published online: 1 August 2017
© The Author(s) 2017. This article is an open access publication

Abstract

Introduction Transfers in multimodal urban trips imply a disutility for travellers, who perceive them as a penalty when using public transport. It is therefore important to estimate the utility associated to transfers and the main factors affecting it, to allow policymakers to reduce users' perceived disutility and enhance their intention to use public transport. The aim of this paper is to develop a Stated Preference experiment to estimate the penalty perceived by commuters (workers and students) when making transfers in multimodal urban trips.

Method A web-based survey combining a Revealed Preferences and Stated Preferences survey was created using Ngen software, and an efficient design was applied to estimate multinomial logit models. We present here the first stage of the research: the design and results of the pilot survey of commuters travelling by metro or urban bus in the city of Madrid, Spain.

Results The findings reveal a pure transfer penalty, independent of in-vehicle time, walking and waiting time and crowding. This pure penalty increases with the number of transfers. Crowded transfers cause a high disutility for commuters, which rises with the number of transfers in the total trip.

Conclusions This paper highlights the importance of conducting a pilot survey when designing the final survey.

Transfer penalties vary between cities, so pilot studies are encouraged to obtain more accurate results. Further research is needed to consolidate the pilot results with those of a final survey.

Keywords SP experiment · Web-based survey · Transfer penalty · Efficient design · Utility functions

1 Introduction

Numerous studies have addressed the transfer phenomenon in recent decades. Passengers perceive a penalty associated to the fact of transferring, which has been defined by some authors as the disutility of one transfer option compared to a non-transfer alternative [1]. Others see it as an extra cost caused by the additional effort required to make connections [2]. Currie [3] and Iseki and Taylor [4] identified a pure transfer penalty, independent of walking and waiting time, but dependent on other factors such as the availability of adequate information, safety, security, comfort and convenience, familiarity with the public transport (PT) system, and frequency of PT use. Cascajo et al. [5] recently identified these factors through qualitative research based on focus groups (FGs). FGs have proved very useful during the design phase of the questionnaires [6], and have been used by authors to identify which variables to include in surveys of PT users. As a continuation of this research, the purpose of this paper is to design an experiment to estimate passenger transfer penalties in urban PT trips.

The estimation of pure transfer penalties is essential to enable policymakers to reduce, as far as possible, passengers' aversion to transfer, as some even choose not to travel by PT if it involves a transfer. Some researchers have estimated transfer penalties by designing stated preferences (SP) experiments

✉ Rocio Cascajo
rocio.cascajo@upm.es

¹ Transport Research Centre, TRANSyT, Universidad Politécnica de Madrid ETSI Caminos, Canales y Puertos, C/ Profesor Aranguren, s/n. 28040, Madrid, Spain

² Transport Research Centre, TRANSyT, Universidad Politécnica de Madrid, Madrid, Spain

³ Universidad Politécnica de Madrid, Madrid, Spain

[7–9]. The use of SP surveys has become widespread in a range of fields such as marketing, transport, health economics, and agricultural and environmental economics [10]. They can elicit responses regarding the behaviour of a single individual or group to estimate and identify their preferences. SP methods have been widely used in the transport field since the 1980s [11, 12], and can be effectively applied to estimate transfer penalties, as they measure perceptions and attitudes. Douglas and Jones [7] describe a SP survey undertaken in Sydney in 2012 to derive ‘pure’ transfer penalties that separate the cost of the penalty of walking and waiting times. The survey was designed to test the difference between ‘same platform’ and ‘up and down’ transfers involving escalators or elevators, as well as bus-to-bus and bus/rail transfers, and used pair-wise choices presented on computer tablets. A logit model isolated the pure transfer penalty and estimated penalties for different types of transfer. Navarrete and Ortúzar [8] investigated users’ subjective valuations of the transfer experience and the influences of certain variables on transferring between PT modes. Their analysis consisted of a qualitative study based on FGs and a quantitative study with a SP experiment and the estimation of advanced discrete choice models. Finally, Schakenbos [9] carried out SP research to determine the disutility of a transfer between bus/tram/metro and train, and estimated a general mixed logit error component model (ECL). The alternatives in the choice experiment are described by six attributes, and the results provide insights into their relative importance, expressed in generalized travel time.

This article focuses on the pure transfer penalty and its estimation; i.e. the penalty of an ideal transfer in which walking and waiting times are equal to zero. The experimental design is applied to Madrid, the capital of Spain. As the first step in the research, this paper develops a pilot survey to obtain more accurate and precise prior parameters for calibrating a MNL model. The results of the pilot survey are used to design the definitive survey. The questionnaire provides both RP data on current travel behaviour and SP data on route choices under scenarios with a different number of transfers. Bradley and Daly [13] highlight the desirability of combining the stronger features of RP and SP data. Many researchers have assessed the transfer penalty by designing a survey with a RP part and a SP part [7–9, 14, 15]. The survey sample was drawn from a diverse population of commuters travelling by urban bus and metro in the city of Madrid. In common with other researches [8, 16], the survey was restricted to commuters due to their higher transfer rates [17]. This research considers commuters to be people who periodically and recurrently travel between their place of residence and their place of work or study. This article is the first step in identifying the relevant attributes involved in the way travellers perceive transfers, and analyses the results of the pilot survey in order to propose preliminary policy recommendations.

The paper is structured as follows. Section 2 presents the methodological background of the research. Section 3 proposes an experimental design for estimating passenger transfer penalties applied to Madrid, and Section 4 presents the definitive redesign of the survey based on the results of the pilot survey. Finally, Section 5 discusses the paper’s main outcomes and conclusions.

2 Methodological background

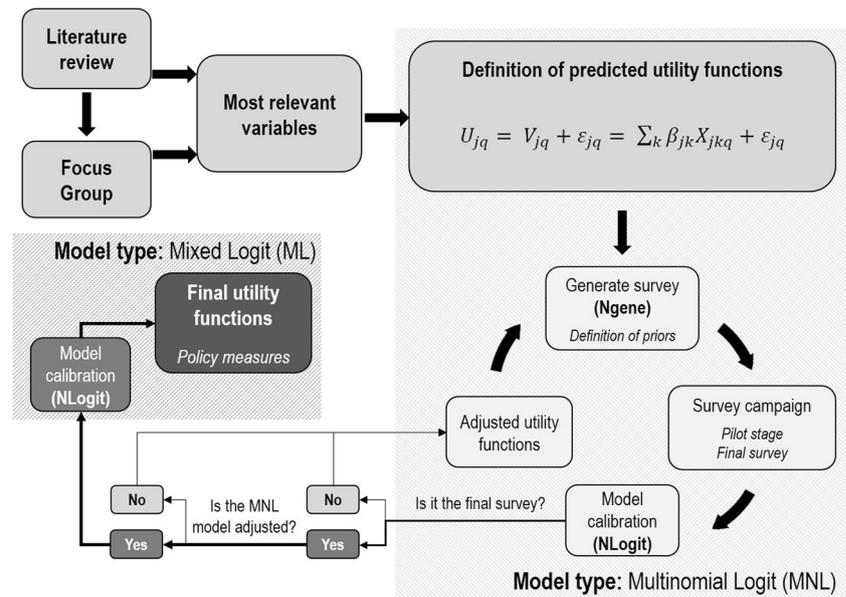
This section contains the methodology used to estimate the penalties associated to transfers. A SP survey is designed using a qualitative approach based on the variables obtained from a FG. SP surveys generally require the definition of the model and the parameters to be estimated in order to build the experimental design and create the questionnaire.

Figure 1 shows the methodology. The first step is to identify the most important known variables affecting the perception of transfers from a review of the literature. A series of FGs are then conducted with PT users to identify these variables in their multimodal trips in the proposed case study. According to Iseki and Taylor [4], transfer penalties vary between cities, so city-specific studies are required for a more accurate identification of the variables. A SP experiment was designed with an efficient design using Ngene software based on the variables identified in the literature review and FGs. A pilot survey is needed to obtain more accurate and precise prior parameters for calibrating a MNL model with Limdep NLogit software. Finally, the definitive survey was created with Ngene software. This means that the MNL model is adjusted; in other words, that almost all the parameters of the variables are significantly different from zero at the 95% level. The following steps in the research are also shown in Fig. 1 (although they are not part of the objective of this paper). After the final survey, the data were used as input to calibrate a ML model with NLogit software, which serve as a tool to propose policy recommendations and to quantify PT users’ perception of transfers.

2.1 Focus group approach

Qualitative methods can gather information directly from transport users in order to understand “why” user decisions are being made. Qualitative methods offer a powerful tool for understanding the complexities of travel behaviour. Different methods such as FGs and interviews have been applied in recent decades, and can be combined with quantitative approaches or used on their own to fill the gaps left by quantitative techniques [18]. When analysing the transfer phenomenon, some researchers have used FGs as a first step to identify the variables to include in the SP experiment [8, 9, 19]. Following these approaches, we decided to use FGs to identify

Fig. 1 Methodology chart



the factors most penalised by PT users when taking a route with transfers in urban areas.

The findings of this focus group study serve as the basis for designing the preliminary or pilot version of the survey.

2.2 Modelling utility functions

Random utility theory postulates that individuals (q) choose among different alternatives (A_j) on the basis of their utility (U_{jq}). The modeller assumes that the utility can be represented as a sum of two components: first, a representative utility function (V_{jq}), measurable from attributes (x) of the individual, alternative or choice situation, weighted by coefficients (β); while the second is a random term (ε) for the difference between the representative and the real utility. It is represented according to the following expression:

$$U_{jq} = V_{jq} + \varepsilon_{jq} = \sum_k \beta_{jk} X_{jkq} + \varepsilon_{jq} \tag{1}$$

Domencich and McFadden [20] postulated that if errors have an identical and independent Gumbel distribution, the probability of an individual choosing a particular alternative (A_j) from the available choice sets (A_q) is given by the multinomial logit (MNL) model.

$$P_{iq} = \frac{\exp(\beta' V_{iq})}{\sum_{A_j \in A_q} \exp(\beta' V_{jq})} \tag{2}$$

The vector (β') cannot be estimated separately from the other parameters defined in the (V_{jq}) function, and for many purposes it can be assumed without loss of generality that the first coordinate of (β') takes value one. However, the MNL model has some limitations. It assumes the independence of irrelevant alternatives; it does not consider the order (where

relevant); random tastes (differences in β) cannot be represented; it assumes the utility functions of the alternatives are homoscedastic, and that all observations are independent. Interactions can be added to the MNL model in order to capture the influence of other socioeconomic variables (such as age, gender, income and others) on attributes. In all cases, if any of the limitations are not satisfied, the results of the MNL model may not be realistic.

The mixed logit (ML) model is an effective solution to overcome the restrictions of the MNL model. It allows the coefficients (β) to vary randomly across individuals (taste variations), and offers an efficient estimation when there are repeated observations. It can also deal with the correlation between alternatives and heteroscedasticity. One specification of the ML model must be considered: the error component logit (ECL).

The ECL model is similar to the MNL model in Eq. (1). However, this model includes an additional error term (η_{jq}) at the end of the formula, which can be distributed to achieve a closer model fit (see Eq. 3). We assumed this term to be normally distributed. The ECL model is well known for its modelling potential, while also offering a multinomial logit kernel, inter-alternative and inter-observation correlation of random terms.

$$U_{jq} = V_{jq} + \varepsilon_{jq} = \sum_k \beta_{jk} X_{jkq} + \varepsilon_{jq} + \eta_{jq} \tag{3}$$

All these models are estimated using the simulated maximum likelihood method, which finds the parameter values that make the observed data most likely. These models (MNL and ECL) will be used with data obtained from a stated preference (SP) survey to analyse the impact of each attribute on the users' perception of transfers and the pure penalty transfer phenomenon.

2.3 Generating data for estimating transfer penalties in public transport

SP surveys are an essential tool for modelling utility functions (developing a Discrete Choice Model) to represent traveller decisions when facing different travel alternatives and policy measures. They present individuals with hypothetical choice situations and enable several responses for each one. However, SP data cannot be used in MNL models as each individual's responses to different choice situations are not independent [21].

Each design must take into consideration the real values of the attributes in the case study in order to present realistic choice situations. Attributes may vary under predefined values at different levels and again be consistent with reality. The latest software allows multiple types of surveys to be created with an orthogonal or an efficient design. In contrast to orthogonal designs, efficient designs are not intended to minimise the correlation of the data for estimation purposes, but to generate parameter estimates with the lowest possible standard errors [22]. These designs are especially interesting when the sample size is small, when they can provide a very efficient estimation. Efficient designs were not used until recently as they required an estimation of the prior parameter values [23]. These prior parameters are not easy to replicate (and in some cases even the sign is unknown), as the users' perception of attributes varies in each case study. However, this trend has now changed, mainly due to increased research containing information about these prior parameters. Efficient designs will therefore always outperform orthogonal designs [24] if there is any information available on the parameters.

Since prior parameter values substantially influence the final outcome, a pilot study is strongly recommended to obtain more accurate and precise prior parameters. Each design requires the introduction of the number of alternatives for each choice situation, the number of choice situations, prior parameters, the levels at which attributes may vary, and restrictions in the generation of certain variables. Prior parameters and levels may be modified to comply with the utility balance criterion [25]. This refers to the choice probabilities, which, if not balanced, would lead respondents to answer the same option systematically. The number of required respondents would also increase if the choices are not balanced. The software then iterates and indicates the required number of respondents (*S-estimate*) to ensure the sample significance [24].

3 Application to Madrid

Madrid is the capital of Spain and has a population of 3.5 million, a metropolitan area of 6 million inhabitants, and an area of 8030 km². About 12.93 million displacements take

place in its metropolitan area every working day, so it is crucial to ensure an efficient transport network in a city of this size.

PT in the central core of Madrid includes buses, metro, light rail, suburban train and a bicycle sharing system. The urban bus service covers the entire central core with over 200 lines, 3562 km in length, and some 1900 vehicles. The metro network also plays a key role in this area, with 13 lines and a total length of 287 km. This study focuses on the bus and metro systems, as they are the main urban modes of transport, and in combination account for almost 85% of total trips in the central core (2.6 million trips every working day). Our estimates show that approximately 56% of PT users make a single transfer, while 21% transfer more than once, highlighting the importance of optimising transfers to achieve an efficient and high-quality PT system.

3.1 Focus groups conducted in Madrid

The aim of the FGs is to identify the factors affecting users' perception of transfers in their multimodal trips. We conducted three FGs in Madrid: the first with university students, the second with middle-aged workers, and the third with retired people over 65. The decision to separate the respondents by age was due to the differences in the literature on the physical limitations of the elderly [19, 26, 27], which influence the perception of transfers.

A total of 20 people participated in the three FGs. A €10 gift voucher was offered as an incentive for participation in all cases. The FGs were conducted according to the recommended methodologies [27–29]. For detailed information on the process, see Cascajo et al. [5].

The results revealed several factors affecting users' perception of the transfer penalty. The most important –mentioned by over 50% of the participants in the FGs– are “time” (in its different components: walking time, waiting time, transfer time, and total travel time); real-time information; crowding; mode; and different levels in the transfer. Two factors emerged from the FGs that did not appear in the literature, related to the pure transfer penalty [5]: mental effort and activity disruption. Mental effort refers to the extra work required by passengers when making a transfer and the need to remain alert throughout the whole journey in order not to miss their transfer stop. Activity disruption concerns the utility of in-vehicle time, especially on longer trips when the time on-board can be used for activities such as reading, listening to music or even sleeping. Activity disruption is in some way related to mental effort; travellers must be aware of the stop where they need to get off, and can therefore not immerse themselves wholly in their chosen on-board activity, which must also be interrupted when alighting to make the transfer.

3.2 Perceived transfer utility in Madrid

The three FGs conducted in Madrid identified the following quantitative and qualitative variables warranting inclusion in the preliminary utility functions:

- Mode: takes value 1 if metro and 0 if bus.
- In-vehicle time: time (min) elapsed while a person is inside a mode of transport.
- Walking time: time (min) elapsed from the moment a traveller gets off a vehicle and walks to reach the next stop or station.
- Waiting time: time (min) elapsed from the instant a traveller arrives at a stop or station and waits until the next mode.
- Stairs: takes value 1 if there are stairs (or a difference in level) while transferring and 0 otherwise (it is always 0 in bus-bus transfers).
- Real-time information: takes value 1 if there are panels with dynamic time arrival for the intended trip as a whole and 0 otherwise.
- Crowding: takes value 1 if the transfer is overcrowded (involving walking and waiting stages) and 0 otherwise.

Utility functions also captured mental effort and activity disruption (pure penalty transfer) through their constants (α_1 and α_2), which refer not only to this phenomenon but to others not included in the functions. As both utility functions contain the most relevant variables related to transfers, constants will mainly capture the pure penalty transfer.

The first linear utility functions were based on these representative transfer variables. Utility functions had the structure of a MNL model. In our case study, we defined three utility functions depending on the number of transfers (T_0 –no transfers–, T_1 –one transfer– or T_2 –two transfers–). As a 21% of PT users transfer more than once, we decided to show these three alternatives, which is another point of interest of this research.

$$U(T_0) = \beta_{tveh00} \cdot tveh0 + \beta_{mode00} \cdot mode0 \tag{4}$$

$$U(T_1) = \alpha_1 + \beta_{tveh01} \cdot tveh0 + \beta_{mode01} \cdot mode0 + \beta_{twalk11} \cdot twalk1 + \beta_{twait11} \cdot twait1 + \beta_{stair11} \cdot stair1 + \beta_{tveh11} \cdot tveh1 + \beta_{mode11} \cdot mode1 + \beta_{info1} \cdot info + \beta_{crowd1} \cdot crowd \tag{5}$$

$$U(T_2) = \alpha_2 + \beta_{tveh02} \cdot tveh0 + \beta_{mode02} \cdot mode0 + \beta_{twalk12} \cdot twalk1 + \beta_{twait12} \cdot twait1 + \beta_{stair12} \cdot stair1 + \beta_{tveh12} \cdot tveh1 + \beta_{mode12} \cdot mode1 + \beta_{twalk22} \cdot twalk2 + \beta_{twait22} \cdot twait2 + \beta_{stair22} \cdot stair2 + \beta_{tveh22} \cdot tveh2 + \beta_{mode22} \cdot mode2 + \beta_{info2} \cdot info + \beta_{crowd2} \cdot crowd \tag{6}$$

It should be noted that cost is not included in the utility functions. About 73% of PT users in Madrid have a monthly or annual flat-rate travel card (in the commuter group this number is even higher). The remaining 27% use single or multiple tickets which allow transfers within the metro network at no extra cost (except in the case of metro-bus or bus-bus transfer). There is therefore no extra cost to transfer in most cases, and nor was this cost variable found to be significant in the FGs. Commuters assume cost as an unchanging variable when travelling on PT, regardless of the number of transfers. These predicted and preliminary utility functions serve as the basis for designing the SP survey (both the pilot and final versions).

3.3 Pilot survey design

The pilot test is one of the most important components of the survey procedure. A pilot survey is a useful fail-safe precaution to take before conducting the main survey [30]. The main benefits of pilot surveys from our point of view are the following: they test the survey structure and the validity of the experimental design; they allow the fieldwork to be refined; and they identify certain behaviours by the respondents.

The pilot survey was web-based. The questionnaire was broadly divided into three main parts: a) trip characteristics regarding current travel behaviour (RP data); b) SP choice scenarios; and c) socio-economic/personal information. Below we explain parts a) and c) in the RP and socioeconomic questions, and part b) in SP choice situations.

3.3.1 RP and socioeconomic questions

Aside from the SP part of the survey, in part a) we asked participants about the characteristics of their regular journey, including all the variables described, plus others. They were asked their occupation (as participants have to commute), trip purpose, type of ticket used, trip origin and destination, trip start time, number of transfers, and trip features such as total travel time, waiting and walking time, modes of transport used, in-vehicle time, whether passengers used mobile apps to see the waiting time for the next vehicle, existence of real time information panels during transfers, and whether they engage in any activity during the trip (listening to music, reading, studying, sleeping and others).

Part c) gathered socio-economic and personal information. The questions concerned gender, age, level of studies completed, household income and household size. There were also some questions about the importance and satisfaction with certain aspects related to transfers (real-time information and mobile network coverage during transfers, or sheltered stops and seats at transfer points). There was an open question at the end for noting additional comments.

3.3.2 SP choice situations

Both our pilot and final SP questions were designed using Ngene software based on a multi-criteria approach, which compares a number of alternatives in different choice situations on the basis of attributes obtained from the literature review and FGs. All the variables described and included in the predicted utility functions are used to design choice situations.

We opted to use an efficient design to estimate MNL models, which involved introducing the predicted utility functions. Although the desirable final output of the study is a ML model, and an efficient design can be applied to estimate random parameter models, we chose the MNL option as Ngene strongly recommends first generating a non-Bayesian design with this model [24]. This identifies any potential problems with the design specifications more quickly.

In the particular case of Madrid there are no previous studies calculating prior components, so they were all unknown. We obtained the average values and signs for each one, and the common levels of each attribute from the literature review. The prior parameter values were then slightly modified to ensure the utility balance criterion [25]. Table 2 shows the values of the first prior parameters and the levels of attributes considered.

Some restrictions were also applied to avoid Ngene generating unreal alternatives and failing to ensure the principle of utility balance. For example, total trip time in T_1 and T_2 should be less than 3 to 10 min compared to T_0 and T_1 , respectively.

After an iterative process, Ngene generated 18 choice situations (the number is a multiple of the levels of attributes) with three alternatives each (54 alternatives in total). Respondents took a long time to understand and choose between these choice situations, so there was a risk that the survey would be only half completed. To avoid this occurring, Ngene generated three blocks, each one of which would be completed by

a different respondent; so three respondents would complete a whole SP survey. Each participant therefore answered the block of choices that had generated the fewest responses at that particular time. This has the added advantage of making the answers less correlated between individuals than if only one respondent had completed the survey. On the other hand, more respondents are required to comply with the value of the *S-estimate*.

Figure 2 shows a screenshot of a choice situation. Total trip time, total walking time and total in-vehicle time are indicated for each alternative. This is because of comments made by the participants in the first test of the survey; they declined to add these times together so they could be considered as a whole.

The way choice situations were shown to participants depends on the habitual trip revealed in the RP questionnaire:

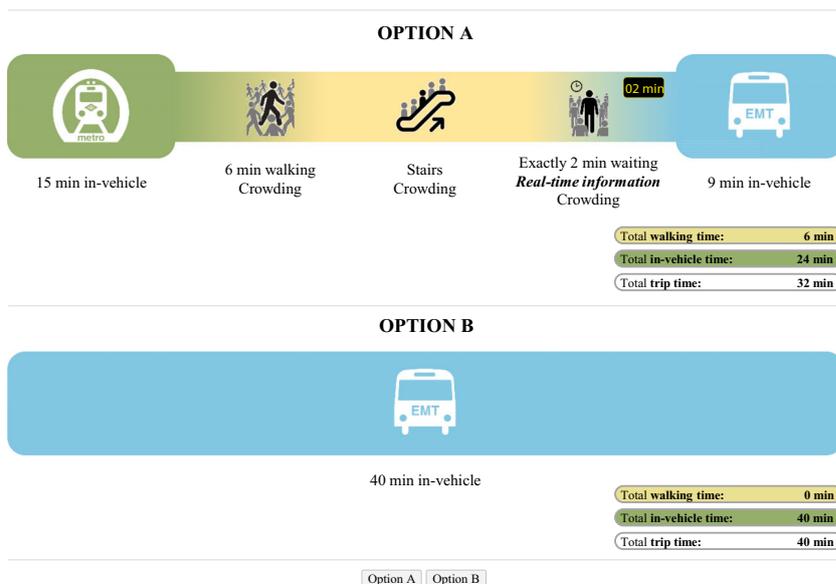
- If respondents did not usually transfer, they were given 6 simple choice situations between 0 and 1 transfer.
- If respondents usually transferred once or more, they were given 6 double choice situations between 0 and 1 transfer, and then between the same alternative of 1 and 2 transfers.

These rules produce more realistic data, as participants did not have to choose between unfamiliar scenarios. Finally, the minimum number of required surveys was established at 24.60, corresponding to the maximum value of the *S-estimate* parameter for all the attributes considered. The minimum number of individuals surveyed is obtained by multiplying the rounded-up *S-estimate* parameter and the number of blocks; i.e. 75 people were required in our case study.

3.4 Conducting the pilot survey

Once the survey had been designed, it was uploaded to a web page. It was decided to use a web-based format as this has

Fig. 2 Survey screenshot: choice situation between no-transfer and one-transfer route



some advantages over paper surveys. The main benefits of web-based surveys are [31]: they are cheaper than other survey setups; they reduce the time required for implementation; data from web-based surveys can easily be imported into data analysis programs; and the respondents can answer at their own convenience (some online surveys even allow respondents to start and then return to the question where they left off earlier).

Respondents were mainly recruited by handing out flyers with all the information required for the survey: website address, personal password for filling in the survey, information about the project and the option of entering a draw for a gift voucher if the survey was completed. The flyers were distributed at metro and bus stations during the morning peak commuter period (7–10 am) on five consecutive weekdays, seeking to obtain a representative sample. As indicated before, target participants were limited to commuters. Participation was voluntary. The answer ratio using this method was around 15%. A €200 gift voucher was offered as an incentive for participation.

Ten respondents completed the survey with personal assistance to test whether there were comprehension problems and whether the icons used in the SP part of the survey (see Fig. 2) could be understood.

The pilot survey involved 79 commuters, of whom 65% were workers and the remaining 35% were students. The average age was 31, and 56% were women. All respondents regularly commuted via PT and their average number of transfers was 0.72 (44% of respondents did not transfer). The average door-to-door trip time was about 35 min, and initial and final walking times were around 5 and 7 min respectively. 20% and 80% of participants chose the no-transfer and one-transfer options respectively in the SP choice situations. When they were asked to choose between making one or two transfers, 64% opted for the first option and the remaining 36% for the second.

3.5 Limitations of the survey

It should be noted that the number of respondents surveyed is sufficiently representative to obtain prior parameters, but small enough to gather meaningful results. However the main aim of this paper is to describe the methodology to conduct a successful SP experimental survey. The MNL model suggests some preliminary policy recommendations, which must be confirmed by the results of the final survey.

3.6 MNL model calibration

After conducting the pilot SP survey and cleaning up the data base, we used Limdep NLogit software to run the utility functions (Table 1). The results in Table 1 offer great potential for

analysing trends and providing preliminary policy recommendations, and for understanding users' perception of transfers.

All significant variables have the expected signs and values. Time-related variables (walking time, waiting time and in-vehicle time) are negative as expected, and almost all significantly. In order to estimate comparisons in equivalent in-vehicle times (IVT), we set an average value of 0.3576 and 0.3472 when making one and two transfers respectively. This shows that on average, in-vehicle times are perceived almost the same, regardless of whether one or two transfers are made. Walking and waiting times, however, are more poorly perceived on average in $U(T_2)$ than in $U(T_1)$ in this pilot study.

Constants in $U(T_1)$ and $U(T_2)$ are also significant, and explain all the unobserved variables not included in the model, and particularly the pure transfer penalty phenomenon. The constant in $U(T_2)$ is clearly higher than in $U(T_1)$, so the penalty perceived on trips increases with the number of transfers. The impact of the pure transfer penalty is perceived as 10.9 and 16.7 equivalent in-vehicle minutes when making one and two transfers respectively. Stairs also have a negative sign, but this result must be treated with caution as it is close to –but not significantly different from– zero at the 90% level.

It is worth noting the disutility produced by crowding scenarios, which is higher in the case of two transfers than only one. This variable, which implies a large number of people gathered together in a limited space, influences transfer perceptions, and is one of the most significant in utility functions after constants. Its impact is comparable to an increase of 2.9 min in equivalent IVT when making one transfer, and more than double when making two transfers (6.8 min equivalent IVT), highlighting its importance. Nor is mode significantly different from zero at the 95% level, which can be explained by the way the variable was introduced in the model. This can be resolved by expressing the variable mode using dummy variables indicating the absence or presence of transfers between bus-metro, metro-bus and metro-metro (assuming transfer between bus-bus as the reference group). In the next step of the research, we separate out the effects of the different modal transfer combinations in the ECL model when analysing the results of the final survey.

Utility functions indicated that the most severely penalised time varied between alternatives. If only one transfer was made, in-vehicle time produced the maximum disutility (greater than walking or waiting times). However, when transferring twice, walking time was more poorly perceived in the second transfer, followed by in-vehicle time in the second vehicle, and waiting times. These results can be compared to those of other studies. Navarrete and Ortúzar [8] reported that waiting time was the most severely penalised, followed by initial and final walking times. These variations highlight the differences between cities, as stated by Iseki and Taylor [4].

Finally, parameter values from Table 1 were introduced as new prior components to design the definitive survey.

Table 1 Results of the MNL models of utility functions. Pilot survey

Utility function	Attributes		Parameter value	<i>p</i> -value
U (T ₀): no-transfer	In-vehicle time	tveh0	-.3573	.0000*
	Mode	mode0	-.1547	.6389
U (T ₁): one transfer	Constant	constant1	-3.9160	.0053*
	In-vehicle time in first vehicle	tveh0	-.3705	.0001*
	First mode	mode0	.6236	.0961
	Walking time	twalk1	-.3407	.0010*
	Waiting time	twait1	-.3320	.0009*
	Difference in level	stair1	-.4515	.1113
	In-vehicle time in last vehicle	tveh1	-.3448	.0008*
	Last mode	mode1	.3647	.3869
	Real-time information	info	.1691	.6370
	Crowding	crowd	-1.0272	.0038*
U (T ₂): two transfers	Constant	constant2	-5.7962	.0375*
	In-vehicle time in first vehicle	tveh0	-.3255	.0264*
	First mode	mode0	-.0133	.9803
	Walking time in first transfer	twalk1	-.3570	.2880
	Waiting time in first transfer	twait1	-.4379	.0127*
	Difference in level in first transfer	stair1	-1.1385	.0277*
	In-vehicle time in second vehicle	tveh1	-.4372	.0006*
	Second mode	mode1	1.5490	.0059*
	Walking time in second transfer	twalk2	-.9148	.0044*
	Waiting time in second transfer	twait2	-.3712	.0662
	Difference in level in second transfer	stair2	-.7091	.1267
	In-vehicle time in last vehicle	tveh2	-.2790	.0153*
	Last mode	mode2	1.0633	.0605
Real-time information	info	.3593	.3911	
Crowding	crowd	-2.3642	.0008*	

*Significantly different from zero at the 95% level

4 Definitive redesign of the survey

The pilot survey enabled the definitive survey to be generated by applying an efficient design for estimating a random parameter model. However, this option was rejected in favour of optimising the design for the MNL model, assuming that this design is also efficient for estimating a ML model. Its behaviour was then verified under a ML model (using the `eval` command in Ngene software). This decision has a negligible influence on the survey design, as a ML model can subsequently be adopted for utility functions.

With the exception of the newly adopted prior parameters (Table 2), all other settings remained unchanged (number of alternatives, number of choice situations, number of blocks, attribute levels, and restrictions). It should be noted that if any of these settings were changed, another pilot survey would be required to obtain accurate prior parameters. The attributes also remained unchanged since they were identified in the focus group stage. If the final model estimates insignificant parameters (we obtained 11 in Table 1, which is high), then it

might be advisable to propose new attributes and change the number of alternatives and choice situations, attribute levels and restrictions in the SP part of the survey, and perform another pilot study. Ngene yielded a similar value for the *S-estimate* as in the pilot survey (25.43). This number must be multiplied by the number of blocks to obtain the minimum number of respondents (i.e. 78 respondents required). However, it is recommended to collect a higher number of answers in the final stage in order to gather meaningful results.

The results of Table 2 indicate that the prior parameters found in the literature for the disutility of times (walking, waiting and in-vehicle) do not differ substantially from those obtained by calibrating the predicted utility functions. However, the influence of other parameters such as stairs and mode on travellers' perception of transfers varies considerably between different case studies (although the signs of the parameters significant at the 90% level always match). As observed in Table 2, some prior parameters differ from the results of the MNL model (Table 1), as the parameters were not significantly different from zero at the 95% level. The

Table 2 Attribute levels considered and prior parameters (pilot and final surveys) as input of Ngene software

Utility function	Coefficients	Prior parameters		Attributes	Levels
		Pilot survey	Final survey		
U (T ₀): no-transfer	tveh00	-2.000	-.3573	tveh0	35, 40, 45
	mode00	1.0000	.5000*	mode0	0, 1
U (T ₁): one transfer	constant1	-1.0000	-3.9160	constant1	-
	tveh01	-2.000	-.3705	tveh0	9, 12, 15
	mode01	2.0000	.6236	mode0	0, 1
	twalk11	-.2500	-.3407	twalk1	2, 4, 6
	twait11	-2.000	-.3320	twait1	2, 5, 8
	stair11	-1.0000	-.7000*	stair1	0, 1
	tveh11	-2.000	-.3448	tveh1	9, 12, 15
	mode11	2.0000	.6000*	mode1	0, 1
	info1	2.0000	.7000*	info	0, 1
	crowd1	-2.0000	-1.0272	crowd	0, 1
U (T ₂): two transfers	constant2	-1.1000	-5.7962	constant2	-
	tveh02	-2.000	-.3255	tveh0	4, 7, 10
	mode02	1.1000	1.5000*	mode0	0, 1
	twalk12	-4.000	-.3570	twalk1	2, 3, 4
	twait12	-3.000	-.4379	twait1	1, 3, 5
	stair12	-1.0000	-1.1385	stair1	0, 1
	tveh12	-2.000	-.4372	tveh1	4, 7, 10
	mode12	1.1000	1.5490	mode1	0, 1
	twalk22	-4.000	-.9148	twalk2	2, 3, 4
	twait22	-3.000	-.3712	twait2	1, 3, 5
	stair22	-1.0000	-.7091	stair2	0, 1
	tveh22	-2.000	-.2790	tveh2	4, 7, 10
	mode22	1.1000	1.5000*	mode2	0, 1
	info2	1.1000	.7000*	info	0, 1
crowd2	-1.1000	-2.3642	crowd	0, 1	

*Prior parameters differ from the results of the MNL model as the parameters were not significantly different from zero at the 95% level

prior parameters associated to the non-significant parameters differ from the pilot survey due to the utility balance criterion [25]. As some prior parameters (those statistically significant in the MNL model), levels and number of choices, among others, are fixed, we could only change non-significant prior parameters to ensure that all choices provide a similar utility to the respondents. We sought to vary these parameters by complying with the utility balance criterion and using similar values to those found in the literature research.

Based on the feedback from the ten respondents who completed the survey in our presence, we changed some questions and visual items (real-time information and crowding) to avoid misunderstandings. The preliminary findings from the research in this area so far are of great interest, and will be further enhanced by the analysis of the results of the definitive survey, the adjusted ML model and a more accurate study of the pure transfer penalty, which is the goal of a future research work (see Fig. 1).

5 Discussion and conclusions

This research summarises the methodology used to design a SP experimental survey to estimate different utility functions based on the number of transfers in multimodal PT trips. The experiment was applied to commuters in Madrid. This research contributes to the existing literature by measuring the transfer penalties perceived by PT users when transferring twice. The results from the FGs and the MNL model suggest some preliminary policy recommendations. First, the FGs provide evidence of the importance of two variables included in the pure transfer penalty and not yet identified in the literature: mental effort and activity disruption. These variables are negatively perceived while transferring. This pure transfer penalty was captured by the constants in the utility functions, which were found to be significant and higher in the case of two transfers. This means that the more transfers made, the greater the associated penalty. This finding suggests a traveller would

prefer a non-transfer alternative to a single transfer, even in an ideal transfer in which walking and waiting times are equal to zero.

Similarly, the disutility produced by transferring is the same as if the walking or waiting times increase by ten minutes. Although walking and waiting times are known to negatively affect transfer utility, some measures can be implemented to increase transfer utility, such as offering real-time information, or providing services or activities to pass the time.

Crowding was found to be the most important variable in utility functions after constants. An overcrowded scenario is even worse when users transfer twice, as its effects on their transfer perception is double the effect of transferring once. It is therefore essential to optimise users' flows to avoid crowding.

The design of the definitive survey revealed the paramount importance of conducting a pilot survey. The main advantage of the pilot stage is that more accurate prior parameters can be obtained before designing the definitive survey. Otherwise, variables in the final model may be insignificant when in reality they are not. A pilot study allows the design's structure and validity to be tested, and the RP and SP parts of the survey to be debugged. It is also essential to personally assist a group of respondents to ensure they understand the choice scenarios. In our case, the definitive survey was enhanced by improving the visual design of the alternatives (we increased the visibility of some details such as real-time information and crowding) and the wording of some of the questions. Another lesson learned was the advisability of designing a survey based on a MNL model in Ngene, although the final model was intended to be a ML. This has a negligible effect on the survey design, but confers some advantages, including a decrease in the required computational time and resources, and ease of understanding. Prior parameters related to transfer times (walking, waiting and in-vehicle) are similar to those found in the literature, but this was not the case of the prior parameters associated to variables such as mode and stairs. Users therefore perceive certain attributes differently depending on their location. A pilot survey enables certain prior parameters to be refined and allows values to be obtained for the remaining parameters.

The final survey will shortly be conducted as part of our future research. The ECL model will be tested and the results will provide more accurate policy recommendations. The ML model will incorporate the variables included in utility functions, in addition to more variables from the RP part of the survey. Constants will therefore reflect the pure transfer penalty more accurately (as they capture all the remaining effects overlooked by the variables in the model). Efforts will focus on calibrating the final models and confirming the trends observed in the pilot survey. This later stage of the research will provide even greater insights and offer a better understanding of how travellers perceive transfers, which will

be of considerable interest for planning transfer modes and managing connecting PT services so as to improve the competitiveness of PT.

Acknowledgements The authors would like to thank the Spanish Ministry of Economy and Competitiveness and its State Program for Research, Development and Innovation oriented to Challenges in Society (TRANSFER Project; TRA2013-45239-R) for providing the necessary funding. The authors also acknowledge the collaboration of Elena Lopez and Luigi Dell'Olio who helped in the design of the pilot survey. This research was also possible thanks to the research grant awarded to the second author [BES-2014-069223].

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Guo Z (2003) Assessment of the transfer penalty for transit trips in downtown Boston. A GIS-based Disaggregate Modeling Approach, Doctoral Dissertation, Massachusetts Institute of Technology, Massachusetts
2. Horowitz AJ, Zloesel DJ (1981) Transfer penalties: another look at transit riders' reluctance to transfer. *Transportation* 10(3):279–282
3. Currie G (2005) The demand performance of bus rapid transit. *J Public Transport* 8:41–55
4. Iseki H, Taylor BD (2009) Not all transfers are created equal: towards a framework relating transfer connectivity to travel behaviour. *Transp Rev* 29:777–800
5. Cascajo R, Herrero F, Monzon A (2016) Identificación de las variables que afectan a la percepción de las transferencias en los viajes multietapa: focus groups en Madrid y Vitoria. *TRANSyT Working Paper 2016-01-ES*, Universidad Politécnica de Madrid. <http://www.transyt.upm.es/index.php/es/mnu-difusion/working-papers.html>. Accessed 7 Dec 2016
6. Roetting M, Yueng-Hsiang H, McDevitt J, Melton D (2003) When technology tells how you drive-truck drivers attitudes towards feedback by technology. *Transport Res F-Traf* 6:275–287
7. Douglas NJ, Jones M (2013) Estimating transfer penalties and standardised income values of time by stated preference survey. *Australian transport research forum*, pp. 1–21
8. Navarrete FJ (2010) Valoración Subjetiva del Transbordo en Transporte Público: caso Santiago de Chile. M.Sc. Thesis, Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Santiago de Chile
9. Schakenbos R (2014) Valuation of a transfer in a multimodal public transport trip. A stated preference research into the experienced disutility of a transfer between bus/tram/metro and train within the Netherlands. Master Thesis, University of Twente, Enschede
10. Cherchi E, Hensher DA (2015) Workshop synthesis: stated preference surveys and experimental design, an audit of the journey so far and future research perspectives. *Transport Res Proc* 11:154–164
11. Bates J (1988) Econometric issues in stated preference analysis. *J Transp Econ Policy* 22(1):59–69
12. Hensher DA (1994) Stated preference analysis of travel choices: the state of practice. *Transportation* 21(2):107–133
13. Bradley MA, Daly AJ (1997) Estimation of logit choice models using mixed stated preference and revealed preference information.

- Proceedings of Sixth International Conference on Travel Behaviour, Quebec 1:117–133
14. Chowdhury S, Ceder A, Sachdeva R (2014) The effects of planned and unplanned transfers on public transport users' perception of transfer routes. *Transport Plan Tech* 37(2):154–168
 15. Espino R, Martín JC, Román C (2007) Análisis de la calidad del transporte de pasajeros. Un estudio de la disposición a pagar por no trasladar. *X Encuentro de Economía Aplicada*, Logroño
 16. Chowdhury S, Ceder A, Schwalger B (2015) The effects of travel time and cost savings on commuters' decision to travel on public transport routes involving transfers. *J Transp Geogr* 43:151–159
 17. Currie G, Loader C (2010) Bus network planning for transfers and the network effect in Melbourne, Australia. *Transp Res Rec* 2145: 8–17
 18. Clifton KJ, Handy SL (2003) Qualitative methods in travel behaviour research. In: Stopher P, Jones P (eds) *Transport Survey Quality and Innovation*, Pergamon, pp 283–302
 19. Navarrete FJ (2010) Valoración Subjetiva del Transbordo en Transporte Público: caso Santiago de Chile. M.Sc. Thesis, Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Santiago de Chile
 20. Domencich T, McFadden D (1975) Urban travel demand: a Behavioural analysis. North-Holland, Amsterdam
 21. Ortúzar JD, Roncagliolo DA, Velarde UC (2000) Interactions and independence in stated preference modelling. In: *Stated preference Modelling techniques, perspectives 4*, PTRC, London
 22. Rose JM, Bliemer MCJ (2009) Constructing efficient stated choice experimental designs. *Transp Rev* 29(5):587–617
 23. Bradley M (2009) Important stated preference experimental design issues in recent transportation applications. *Transp Rev* 29(5):657–663
 24. ChoiceMetrics (2014) *Ngene 1.1.2 User Manual & Reference Guide*, Sydney, Australia
 25. Huber J, Zwerina K (1996) The importance of utility balance in efficient choice designs. *J Mar Res* 33:307–317
 26. Kitzinger J (1995) Qualitative research. Introducing focus groups. *Brit Med J* 311(7000):299–302
 27. Morgan DL (1996) Focus groups. *Annu Rev Sociol* 22:129–152
 28. Krueger RA, Casey MA (2014) *Focus groups: a practical guide for applied research*. SAGE Publications, Los Angeles
 29. Tracy SJ (2010) Qualitative quality: eight «big-tent» criteria for excellent qualitative research. *Qual Inq* 16(10):837–851
 30. Richardson AJ, Ampt ES, Meyburg AH (1995) *Survey methods for transport planning*. Eucalyptus Press, Melbourne
 31. Evans JR, Mathur A (2005) The value of online surveys. *Internet Res* 15(2):195–219