Innovative approaches for an interactive stated choice survey

REITER Thomas, VÖLKLI Andrea & FELLENDORF Martin

Institute of Highway Engineering and Transport Planning,
Graz University of Technology
Rechbauerstrasse 12
8010 Graz
Austria

Phone: +43 316 873 6221,
Fax: +43 316 873 4119

Corresponding author:
Thomas Reiter
Email: reiter@tugraz.at

Submission date:
August 01, 2012

Words including abstract
4227 words, 3 Tables, 4 Figures.

Abstract

Stated preference surveys provide an excellent source of information for the development and targeted implementation of transport planning measures. This paper describes two innovative methods used in a study of road pricing’s impact on travel behavior in Austria. First, the research used an innovative process to generate realistic route choice alternatives for use in the survey. These alternatives were generated using online resources (e.g. Google Maps) and created alternatives that were familiar to the survey respondents. Second, the research used three different methods for data collection: tablet computer assisted interviews, web assisted interviews and traditional pencil and paper interviews. The use of these different methods helped the researchers obtain a representative sample of the Austrian population. While tablet-assisted interviews appealed to the younger generation, older people were incorporated with the pen-and-paper method. Additionally, via systematic sampling people from “hard-to-reach” groups became accessible. An important benefit of using the tablet-based data collection was that survey participant demographic data was available in real time enabling the survey manager to direct surveyors to target underrepresented groups immediately. The tablet computers also improved data quality and slightly increased willingness to participate in the survey.
1. INTRODUCTION

Travel surveys are fact of life for transport researchers, absolutely necessary but difficult to prepare and implement. There are two main types of travel survey: classic paper and pencil surveys, and computer-based data collection. Computer-based surveys are used often today because of their practical utility and their ability to integrate detailed information (e.g. GPS-data) into the survey, thus enhancing survey quality. However, the use of computer-based surveys can make it difficult to obtain a representative sample since some groups have limited access to the Internet or lack the motivation to engage with new technologies. Therefore it is advantageous to apply different survey methods to facilitate access to a broader range of representatives and thus include more elusive parts of society in the survey population.

This paper describes the survey methods used in a stated preference experiment on route choice behavior in Austria completed during spring 2011. The study goal was to learn about route choice behavior of road users under hypothetical road user charges. In the survey participants were given information about four alternative travel choices for a specific trip. Participants were asked to choose their preferred alternative, which provided information on individual preferences. The stated preference results are described in a separate paper, this paper focuses on the specific techniques used for developing the personalized surveys and for collecting the data.

The survey was highly customized for each participant. First, the specific trips began and ended with origins and destinations familiar to the participants (e.g. actual home to work trips). Second, the travel choices were developed using real data from on-line journey planners (three roadway-based trips and one public transport trip were offered for each trip). The true information on travel time, road type and variable costs obtained from the journey planners was supplemented by hypothetical road user charges. Each survey participant was asked to complete six choice sets, allowing researchers to assess individual preferences accurately. Section 4 of this paper describes the process and software developed as part of this research to generate these personalized choice sets.

Four different methods were used for data collection: three computer-based and the classic paper-and-pencil interview (PAPI). The computer-based methods were Tablet Assisted Interview (TAI) and Web Assisted Self Interview (WASI), i.e. completion of the questionnaire online using one’s own PC. Two versions of Tablet Assisted Interviews were conducted. Tablet Assisted Personal Interviews (TAPI) are interviews where respondents are asked questions by a surveyor using a tablet computer. Tablet Assisted Self Interviews (TASI) are interviews where respondents operate the tablets themselves. The software system developed to generate choice sets was used in four types of surveys. Section 5 presents an assessment of the data collection methods.

Section 2 describes stated preference surveys in travel demand research. Section 3 summarizes the stated preference experiment and data collection methods used in the research. Section 6 summarizes the main findings.

2. STATED PREFERENCE SURVEYS IN TRAVEL DEMAND RESEARCH

Stated preference experiments are designed to examine the choice behavior of persons. In stated preference travel surveys the participants face a specific transport situation with different alternatives and are asked to choose between different options (e.g. route choice), providing researchers with information on their preferences.

The fundamental method of stated preference experiments is providing participants with several options from which they can choose. It is vital for the modeling and evaluation of the survey that these options are based on real facts in order to generate a realistic scenario. Rose and Hensher [4] assume that a realistic stated preference experiment is guaranteed only if the alternatives, attributes and characteristics of the choice sets reflect the participants’ real experiences. This means that information about the participants has to be gathered before designing the choice sets.

Postal surveys can meet this requirement. In the first step participants are found e.g. by phone, mail-out/mail-back or interviews [2] and provide initial data. Then a customized survey containing choice sets based on information from the initial data can be sent to the participants. In this manner postal surveys can be a very efficient method.
Another efficient method is the computer- and internet-assisted survey. The choice sets can be generated via software in real-time while the interview is being conducted. Realistic route models are generated by short cuts of web-based traffic data and the demographic parameters of the participants.

3. THE CASE STUDY

This research goal was to assess the impact of variable road user charges in Austria. However, in addition to that specific goal, the research compared four different methodologies for data collection and investigated the benefits of a customized software program designed to provide more realistic trip choice options for the stated preference experiment. This section summarizes the stated preference experiment and the data collection methodologies used in the research. The rest of the paper focuses on the data collection methodologies (results of the stated preference experiment are presented in a separate paper).

Stated Preference Experiment

The stated preference experiment conducted in this research was designed to investigate the impact of elementary changes to travel costs on personal travel behavior. The focus was put on the better understanding of road users and on learning about selected deciding factors.

The research was completed using an interactive stated preference experiment to evaluate the potential behavioral effects of variable road charges. The model developed in this research aims at the detailed analysis of the willingness of road users to pay, depending on variables (factors) such as costs, purpose of journey etc. The main focus was to estimate the route choice behavior of road users. This was evaluated using a discrete choice analysis. Table 1 presents the factors considered in the research.

Table 1: Factors Considered for Route Choice Behavior

<table>
<thead>
<tr>
<th>Influencing factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 mode of transport (car / public transport)</td>
</tr>
<tr>
<td>2 travel time</td>
</tr>
<tr>
<td>3 road category</td>
</tr>
<tr>
<td>4 punctuality / reliability of travel time forecast</td>
</tr>
<tr>
<td>5 delay</td>
</tr>
<tr>
<td>6 toll</td>
</tr>
<tr>
<td>7 fuel costs</td>
</tr>
<tr>
<td>8 total variable travel costs</td>
</tr>
</tbody>
</table>

In the experiment, the participants choose from four possible choices for how they would get from their home to work or to a different well-known destination.

An innovative aspect of this research was the development of software designed to generate customized trip choices for each participant. Participants provided their home address and selected from a list of familiar destinations (e.g. shopping centers) called points of interest (POI). This data was used by the software to calculate trip duration, costs, etc. using on-line information (e.g. Google maps and public transport journey planners). This personalized information was then presented to participants in their personalized survey. Thus, the choices are realistic and aligned with the participant’s familiar environment. This process and software is described in detail in Section 4.

Survey Methodologies

The stated preference experiment was completed using four data collection methods: traditional pencil and paper (PAPI), web-assisted self interviews (WASI) and tablet-assisted interviews (TAI). Two versions of TAI were used: tablet assisted personal interviews (TAPI), where surveyors used a tablet computer to ask questions, and tablet assisted self interviews (TASI), where interviewees handled the tablet themselves. All four surveys were based on an interactive questionnaire, which enabled a standardized and personalized survey. Figure 1 summarizes the four methods.
4. SURVEY IMPLEMENTATION

As outlined above, a key aspect of the research was developing customized and realistic trip alternatives for participants to choose from in the survey. This section describes the process used to generate these alternatives including the software developed as part of this research. The same software was used to generate alternative for all the surveys used in the research. The only difference was that in the PAPI surveys the questionnaires were printed and sent by mail based on origin and destination address data obtained in the first contact with survey respondents.

Information Collection and Display

In the case of the tablet and web-based surveys the technical implementation of the survey is based on client-server communication using asynchronous data transfer. An interview is started when the user selects the start page on the user’s web browser. This start page is generated on the server using PHP and served to the client as an HTML document. To uniquely identify a particular interview, a unique ID is generated server-side and is embedded into the HTML document form as a hidden value, along with a reference ID pertaining to the client, a start timestamp for later calculation of the interview duration and geolocation data.

The HTML document’s structure and technical concept is such that the HTML form is used throughout the whole interview. Navigation between interview pages and steps, as well as input validation and other tasks are done by dynamically applying JavaScript routines. For example, form validation is invoked each time the user interacts with form elements, and navigation buttons are enabled or disabled in accordance with the validation state. At certain steps or intervals during the interview, asynchronous requests to the server are used to submit form data and thus create interview result data rows.

The advantages of this methodology are numerous, for instance, no page reloads are required between interview steps and the complete interview data is always present in the current state of the form.

Several personalized features of the survey such as POI recommendations or route variants based on user input (e.g. home address) require dynamic generation, retrieval and presentation of form content. This is also accomplished through asynchronous requests to the server. In this case the client provides necessary data (i.e. addresses) so that the server can develop appropriate information (e.g. specific routes), generate HTML blocks presenting this information, and return this information to the user. Several specific aspects of the software developed as part of this research are presented in the following sections.
Routing Mechanism

As outlined above the route alternatives presented to participants in the interview are based on addresses the interviewee provides. This requires dynamic calculation of routing data between two points at the time of the interview. The most important factors of the resulting routing information are the route distance and trip duration, which are the basis for later calculation of numerous characteristics for route variants. To accomplish this, the server uses Google’s Directions API for road-based travel and on-line public transport journey planners to generate realistic routing scenarios for the user.

Points of Interest (POI) recommendations

The survey considered both journey to work trips and leisure and shopping trips. The software developed in this research was designed to provide participants with a selection of points of interest (POI) to choose from in the survey. Nine POIs were generated for each survey: three nearby, three medium-distance and three further away. These POIs were selected for users based on their home address.

Figure 2 shows a screenshot of the POI selection page. The decision to use this graphic choice presentation method was made after pre-testing showed that numerous test persons did not know the addresses of shopping and leisure destinations.

The large number of POI combinations and possible post codes made it impossible to use the routing mechanism outlined above for creating the POI recommendations in real time because it would require at least hundreds to thousands of queries to the routing interfaces. To mitigate this problem, 44,761 zip code x zip code tuples were pre-calculated and stored in an indexed database table. Using this data, POI recommendations were generated by local database queries to identify likely suitable POIs for particular distance ranges. Once one of the recommended POIs is selected by the user, detailed routing data is created using the routing interfaces following the standard workflow described above.

Interactive Questionnaire

Figure 3 presents an overview of the content of the interactive questionnaire and the software systems used to feed the information into the survey. The figure shows which specific processes were used for each of the three different survey types considered: tablet, Internet or pencil and paper.
Figure 3 Interactive questionnaire including software system [8].
As shown in Figure 3, the start page provides general information about the project to the participant and transmits the time of the interview start to the database, when the start button is pushed. In the tablet-based surveys, users are asked to select their vehicle type on a screen with pictures that serve as eye catchers, so the interviewees can get used to the touch screen.

Next, all users are asked about their employment status. This information is used to determine which type of trips the user will be asked about: work or leisure and shopping. Non-working persons and 30% of working persons are asked about leisure and shopping trips. All participants provide their home address, while workers also provide their work address. Leisure and shopping trips participants also select from a set of POIs generated randomly for each participant depending on their home address.

Next, the software generates a set of alternative routes with the given starting point and destination for each user. As outlined above these choice sets are developed with the aid of timetables and route information from the internet. Participants can select from four options (3 by car, 1 by public transport). Once the stated preference experiment is finished, questions follow about road choice behavior, willingness to pay road tolls and the environment. Finally a feedback form covering usability and handling of the tablets is offered to tablet users. When the interview is finished, local and temporary GPS data are sent to the database.

While all three types of survey (tablet-based, web-based and pencil and paper) use the same basic approach, Figure 3 summarizes the differences. For example, the pencil and paper survey did not generate automatic data (e.g. timestamps) and it was also carried out in two steps with the initial information collected first, then the results were fed-into the same route generation software to develop the choice sets. These choice sets were then mailed back to the participants in the second step.

Table-based User Interface Design

An important objective of the research was to test the use of tablet-based data collection. Therefore during the questionnaire development phase theoretical considerations [1] regarding the programming of the interactive questionnaire and the operation of the user interface were undertaken.

The goal of interface design is to develop a simple-human-computer interaction. In terms of usability measures this means: friendliness and simple handling, buttons of an appropriate size, minimizing the word count, using appropriate font size. In this research these factors were considered carefully in the pre-test phase where a variety of designs were compared and evaluated.

One method for evaluating interface design is eye tracking (see Figure 4). This method collects information on eye movement, saccades and fixation times. Eye tracking takes the unconscious, rapid psychological processes of perception into consideration. In subsequent in-depth interviews, participants provide information on the aspects they remember and on their questionnaire comprehension. The eye tracking data obtained in questionnaire pre-testing was used to help design the final questionnaire in terms of horizontal and vertical field of view as well as visibility in lowered position. The aim was to optimize the visual perception and operation whilst taking into account the minimization of operating errors.

Figure 4 Verification of the design and arrangement of the questions with the help of eye tracking.
5. SURVEY METHOD COMPARISON

As outlined above, four types of data collection were used in this research: two versions of Tablet Assisted Interviews, Web Assisted Interviews and Paper and Pencil Interviews. An important objective of the research was to compare the different data collection methods. This section summarizes the main findings from the comparison and describes the advantages of tablet-based data collection for stated preference experiments.

Response Rate

Table 2 summarizes data on response rates for the survey methods used in this research. In total 2,455 people were willing to participate in TAI surveys either on-street or at their working place. However about one third of the responses were invalid since the respondents were not frequent car users. In sum the 1,610 valid TAI interviews provided 9,660 route choices for use in the research.

As control group a total of 5,000 people were randomly selected for a WASI or PAPI survey by sending an invitation card. The primary response rate was 29%, who were willing to participate. However only 24% valid web-based or paper-based interviews were returned. A total of 263 interviews were invalid due to infrequent car usage or incomplete answers. The face-to-face interview as well as the 2-level recruitment via mail took about 2 month (53 versus 60 days).

As shown in Table 2, the TAI respondents are slightly younger than the WASI respondents and by far younger than PAPI respondents. This verifies the hypothesis that elderly people are more likely to respond to written questionnaires.

Table 2: Response rate of TAI, WASI and PAPI

<table>
<thead>
<tr>
<th></th>
<th>TAI (Face to Face)</th>
<th>WASI (Recruitment via Mail)</th>
<th>PAPI (Recruitment via Mail)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Sample</td>
<td>2,455</td>
<td>5,000</td>
<td></td>
</tr>
<tr>
<td>Responses</td>
<td>2,455</td>
<td>849</td>
<td>615</td>
</tr>
<tr>
<td>Valid Interviews</td>
<td>1,610</td>
<td>768</td>
<td>434</td>
</tr>
<tr>
<td>Avg Age / Std Dev</td>
<td>40.5 y / 14.6</td>
<td>43.6 y / 12.8</td>
<td>49.9 y / 12.3</td>
</tr>
<tr>
<td>Duration</td>
<td>53 d</td>
<td>60 d</td>
<td>60 d</td>
</tr>
</tbody>
</table>

Survey Duration

Table 3 compares the amount of time it took to collect data using each of the data collection methods. As shown, the survey methods differ in the total time it took to obtain a valid interview. It took about 1½ hours per valid TAI interview (this time includes: training, support, interviewer travel time, recruiting people willing to participate and actually conducting the interview). The WAP interviews take much less time for the survey staff (16.6 minutes), because contacts can be automated via email. Standard paper questionnaires take longer (54 minutes) than WAP due to labor intensive process of surface mailing of reminders and data coding but are faster than TAP.
Table 3  Comparison of Survey Duration per Valid Questionnaire

<table>
<thead>
<tr>
<th>TAI</th>
<th>WASI</th>
<th>PAPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>91 min/pc.</td>
<td>16.6 min/pc</td>
<td>54 min/pc.</td>
</tr>
</tbody>
</table>

- Training of Interviewee
- Support
- Travel
- Recruiting Participants
- Respondents answer quicker if an interviewer is present
- Multiple Contact
- Mailing & collection of questionnaire
- Data coding

Time breakdown per questionnaire

- 24 min/pc recruiting participant
- 5 min/pc explanation of survey
- 10.4 min/pc survey
- Total 39 min/pc
- 14.4 min/pc Web interview
- No record taken

However, it is interesting to note that the average tablet interview itself was quicker than the web-based interview (10.4 min vs. 14.4 min). Apparently, respondents answer quicker if an interviewer is present. It is also remarkable that the interview duration for tablet interviews was much more predictable than for the web-assisted method based on standard deviation. The duration data is very reliable since the response times were automatically recorded by the system, which was not possible with PAPI.

Tablet Usability

In addition to comparing the qualities of the different survey methods used in this research, the respondents taking tablet-assisted interviews were asked to evaluate their experience. The results confirm that tablet computers are rated as relatively easy to use by a large share of the population.

For interviewees the tablets allowed for a user friendly presentation of the questions on a large touch-screen. The user feedback provided insight into participants' attitudes towards the tablet and how they assess its handling. The tablets' advantages in terms of user friendliness are the simple handling, the optimal visibility of the presented contents and the clean menu navigation. A very high number of participants (over 90%) believed that the tablet was easy to control, although approximately one fifth commented negatively on the tablet's weight and the reflections on its touch-screen.

Tablet-based Interviews in Stated-choice Experiments

While tablet computers provide certain advantages to users, this research showed that they are quite valuable for stated choice experiments. The particular advantage is that tablets allow standardized personalized questioning in real time. In other words, tablets can display customized stated-preference questions generated during the questionnaire using web-robots of journey planners.

The use of tablet computers also improves validity by enabling a project manager in a data center to supervise the data collection process in real time. Because data is collected and analyzed centrally in real time, it is easier to obtain a stratified random sample since demographic data (sex, age, income) can be constantly checked against overall distributions. Invalid interviews due to not meeting the selection criteria (e.g. in this research infrequent car usage) or non-comprehension of the stated choices can be eliminated during the interview. The continuous GSM based data also helps improve data quality.

While the data presented above shows that it takes longer for the interviewer to recruit participants for tablet-assisted interviews than for WASI or PAPI it can still be quicker to obtain a statistically relevant number of valid questionnaires simply by increasing the number of well trained interviewers. Another benefit of TAI is that no effort needed to be made sending out reminders by e-mail or post.

Finally, it has been found that tablets have a positive affect on motivation to participate in surveys (in one study 92% of respondents agreed to participate at least partly due to the tablet). [3]
6. CONCLUSIONS

This paper describes two key methodologies developed and used in a stated preference experiment designed to assess travel behavior under potential road pricing programs in Austria. The two methodologies were a software-based process developed to generate realistic travel choice alternative for the stated preference experiment, and a detailed analysis of the use of tablet-assisted data collection for stated preference experiments.

The innovative process for generating realistic travel choices used software to generate routes based on the Google Maps API (for road-based travel) and online public transport journey planners. The software used the respective addresses of homes, work places and leisure destinations to generate alternative sets of standardised and personalised itineraries. The data was then embedded in an interactive questionnaire for the stated preference survey.

The same software was used to generate itineraries for the three different survey methods used in the research: tablet-assisted interviews, web-assisted self interviews and pencil and paper interviews. While with TAI and WASI data was retrieved and integrated on-the-fly, the PAPI questionnaires were printed and distributed via post in a two-step process.

The use of different data collection methods made it possible to generate a representative sample of the Austrian population. While tablet-assisted face-to-face interviews appealed to the younger generation, older people were incorporated with the pen-and-paper method. Additionally, via systematic sampling members of “hard-to-reach” groups became accessible as test persons.

With regard to time expenditure, survey methods differed considerably in the total time it took to receive a valid interview. It took about 1½ hours per valid TAI interview, when training, support and travel time of interviewers, recruiting people willing to participate and conducting the actual interview is considered. WASI interviews are much faster to conduct for the survey staff (16.6 minutes), since contacts can be automated via email. Paper questionnaires took longer (54 minutes) than WAPI due to labor intensive surface mailing of reminders and data coding but were faster than TAI.

Along with relying on technical innovations, the employment of a diverse mix of methodologies forms the basis for the generation of representative samples. This approach allows for results of high quality concerning route choice behavior and rewarding projections for traffic planning.

ACKNOWLEDGEMENTS

The valuable help of Andreas Holzinger from Graz Medical University, Hci14Med, in considering elderly people in the design of the questionnaire is gratefully acknowledged.

REFERENCES