Journal of Transport Geography 17 (2009) 21–29

Contents lists available at ScienceDirect

Journal of Transport Geography

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

Identifying differences in willingness to pay due to dimensionality in stated choice experiments: a cross country analysis

John M. Rose [a,*], David A. Hensher [a], Sebastian Caussade [b], Juan de Dios Ortúzar [c], Rong-Chang Jou [d]

[a] Institute of Transport and Logistics Studies, The University of Sydney, Australia
[b] LAN Airlines
[c] Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile
[d] Department of Civil Engineering, National Chi Nan International University, Taiwan

ARTICLE INFO

Keywords:
Stated choice
Willingness to pay
Cross-country comparisons
Australia
Chile
Taiwan

ABSTRACT

This paper explores the influence of both cultural and socio-economic characteristics on the perception of complexity and cognitive load associated with stated choice (SC) experiments. Complexity is analysed in terms of five design dimensions which were systematically varied according to a macro experimental design. To study the influence of cross country differences on willingness to pay estimates, we combined datasets collected in Sydney, Santiago de Chile and Taichung city in Taiwan, all of them related to an equivalent route choice experiment. Several mixed logit models were specified and estimated; our results show that design dimensions do have an impact on the behavioural outputs of discrete choice models estimated on SC data. However, these influences seem to be data-specific, suggesting that the impact of design dimensions upon SC outcomes may be local and not necessarily transferable across different countries and cultures.

1. Introduction

Stated choice (SC) methods have become the dominant data paradigm used to study the behavioural response of agents (be they individuals, households, or other organizations), given various choice contexts. The popularity of these methods since the pioneering contributions of Louviere and Woodworth (1983) and Louviere and Hensher (1983) has resulted in numerous applications in fields as diverse as transportation, environmental science, health economics, entertainment, marketing, political science and econometrics. Independent of the specific field of application, the generation of SC experiments has evolved to become an increasingly significant but complex component of SC studies (Carlsson and Martinsson, 2002; Huber and Zwerina, 1996; Kanninen, 2002; Sándor and Wedel, 2001). Typically, SC experiments present sampled respondents with a number of hypothetical scenarios consisting of a universal but finite number of alternatives that differ on a number of attribute dimensions and require that these respondents specify their preferred alternative. These responses are then pooled before being used to estimate parameter weights for each of the design attributes (or in some cases, even attribute levels). Depending on the type of experiment conducted, researchers may obtain estimates of the direct or cross elasticities (or marginal effects) of the alternatives as well as the marginal rates of substitution respondents are willing to make in trading between two attributes (i.e., willingness to pay (WTP) measures).

Unlike most data, SC data requires that the analyst designs the experiment in advance by assigning attribute levels to the attributes that define each of the alternatives which respondents are asked to consider. Traditionally, the attribute levels are allocated to each alternative according to some generated experimental design, with the most common approach being to use a fractional factorial design to generate a series of single alternatives, which are then allocated to choice sets using randomised, cyclical, Bayesian or fold over procedures (Bliemer and Rose, 2006; Bunch et al., 1996; Huber and Zwerina, 1996; Kanninen, 2002; Sándor and Wedel, 2001, 2002, 2005).

The need for respondents to repeatedly process information on the attributes and attribute levels of alternatives within SC choice surveys has resulted in the continual calling into question of their ability to accurately undertake such tasks. Of concern is the cognitive load under which respondents are placed in answering SC choice surveys as well as the possibility of fatigue effects rendered through repeated questioning. Research efforts have tended to focus on the impact various design characteristics have upon respondent's ability to respond to choice tasks. Specific issues examining the impact upon behavioural responses have included the number of alternatives within the task (Hensher et al., 2001), the number of attributes (Pullman et al., 2000), the number of attributes and

* Corresponding author.
E-mail address: johnr@itls.usyd.edu.au (J.M. Rose).

0966-6923/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtrangeo.2008.05.001

alternatives (Arentze et al., 2003; DeShazo and Fermo, 2002), the impact of attribute level range upon response (Cooke and Mellers, 1995; Ohler et al., 2000; Verlegh et al., 2002) and the number of choice profiles shown to respondents (Brazell and Louviere, 1998). More recently, Hensher (2004, 2006a,b) and Caussade et al. (2005) examined all of the above effects simultaneously.

Of particular concern is the fact that research examining the impact of design dimensionality upon SC studies has often yielded contradictory evidence. For example, early research found that the first few choice situations of SC tasks were often used by respondents to adapt to the task and develop a decision strategy (Louviere, 1988). In this vein, Meyer (1977) demonstrated that individuals' decision calculi stabilized within three choice situations given a three-attribute choice task. A number of researchers have examined the impact that the number of profiles has on the behavioural responses of respondents completing SC tasks (Arentze et al., 2003; Brazell and Louviere, 1998; Hensher, 2004; Stopher and Hensher, 2000). In each instance, these researchers found no evidence that the number of choice tasks had more than a marginal impact on the behavioural responses of respondents. Nevertheless, Bradley and Daly (1994) found contradictory evidence to the above in the form of increased unexplained variance as the number of choice task replications was increased. A similar finding was reported by Caussade et al. (2005), although they concluded that other design dimensions had a larger impact upon error variance. On another hand, Ortúzar et al. (2000) had to block a design as they found that submitting respondents to 16 choice situations led to confusion, boredom and too many inconsistencies. As such, the issue of the number of choice profiles respondents are capable of handling remains a contentious issue within the literature and there is still the perception that respondents can handle no more than a small number of choice tasks (usually less than 16 with nine or 10 being the most common amounts within the literature).

With rare exceptions, empirical studies have used a single SC design in which the numbers of attributes, alternatives, choice sets, attribute levels and ranges are fixed across the entire design. As a consequence the opportunity to investigate the influence of design dimensionality on behavioural response has been denied. Accumulated wisdom has promoted a large number of positions on what design features are specifically challenging for respondents (e.g., the number of choice sets to evaluate); and although a number of studies have assessed the influence of subsets of design dimensions (e.g., varying the range of attribute levels), precious few studies have systematically varied all of the main dimensions of SC experiments.

Ultimately, this research stream seeks to examine what influences specific design configurations, in the aggregate, have on the behavioural outputs of discrete choice models estimated on SC data. Underlying the research is what Hensher originally termed the *Design of Designs* (DoD) SC experiment in which the 'attributes' of the design are the design dimensions themselves including the attributes of each alternative in a choice set. The design dimensions that are varied are the number of choice sets presented, the number of alternatives in each choice set, the number of attributes per alternative, the number of levels of each attribute and the range of attribute levels.

One problem in forming a clear understanding as to the exact influence different design dimensions play in obtaining SC results is that different country (and underlying socio-demographic) settings may impact the results. For example, the number of choice situations viewed by respondents in Australia may have little impact on parameter estimates but a large impact on respondents from Chile, given the socio-demographic context. In turn, these parameter differences will directly influence the WTP outcomes from the two countries. Thus, the impact of design dimensions upon SC outcomes may be local and not necessarily transferable across different countries and cultures, thus making formation of concrete conclusions difficult.

In the current empirical context, our interest is on both the influence of the design as well as the nationalities of respondents on the WTP for time savings (otherwise known as the value of travel time savings (VTTS)). In order to study these influences, we establish whether the distribution of VTTS derived from the three urban contexts in three different countries – Sydney in Australia, Taichung city in Taiwan and Santiago in Chile – are statistically different from one another and vary in terms of their sensitivity to the dimensionality of the SC experiment. In doing this, we seek to locate any possible systematic differences, especially socio-demographic, that arise in the behavioural VTTS as a result of increases (or decreases) in design complexity.

The remainder of this paper is organised as follows. In the next section, the experimental design used in the study is explained in detail. The survey task is discussed next. The samples collected in each of the three countries are discussed in Section 4 and the modelling undertaken in Section 5. Section 6 presents the model results. Finally, Section 7 provides discussion and conclusions based on the results presented in Section 6. Limitations and future research proposals are also discussed in this section.

2. The experimental design

The focus of the empirical approach is the choice amongst a set of unlabelled tolled and non-tolled routes available to a commuting car driver. A number of design dimensions were identified from the literature as possible sources of influence on choice behaviour. Table 1 shows the five design dimensions selected and their levels.

The design dimensions were manipulated according to a master plan consisting of 16 sub designs. The master plan was designed to allow for the interaction between the number of choice situations and number of alternatives as well as between the number of alternatives and number of attributes. Each of the 16 sub designs was integrated into a SC instrument with two versions (i.e., blocking of 32 rows into sets of 16). The instrument then assigned at random a design to different respondents.

For designs with the largest number of attributes, six attributes were selected, based on Hensher (2001a). These include free-flow, slowed down, and stop/start time components and travel time variability, toll and running costs. To explore the impact of having less attributes on choice, these attributes were combined in various ways. The groupings employed were:

- *designs with three attributes:* total time (free flow + slowed down + stop/start time), trip time variability, total costs (toll + running cost)
- *designs with four attributes:* free flow time, congestion time (slowed down + stop/start), trip time variability, total costs
- *designs with five attributes:* free flow time, slowed down time, stop/start time, trip time variability, total costs
- *designs with six attributes:* free flow time, slowed down time, stop/start time, trip time variability, toll cost, running cost.

Table 1
Dimensionality of the design plan

Choice set size	Number of alternatives[a]	Number of attributes	Number of attribute levels	Range of attribute levels
6	2 + 1	3	2	Narrower than base
9	3 + 1	4	3	Base
12	4 + 1	5	4	Wider than base
15	–	6	–	–

[a] Each experiment also included a reference alternative (see Section 3).