Transportation Research Part A 71 (2015) 141–152

Contents lists available at ScienceDirect

Transportation Research Part A

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

The joint estimation of respondent-reported certainty and acceptability with choice

John M. Rose [a,1], Matthew J. Beck [b,*], David A. Hensher [b,2]

[a] Institute for Choice, University of South Australia, 140 Arthur St, North Sydney, NSW, Australia
[b] Institute of Transport and Logistics Studies, University of Sydney, St James Campus (C13), Sydney, NSW 2006, Australia

ARTICLE INFO

Article history:
Received 29 November 2012
Received in revised form 25 November 2014
Accepted 27 November 2014

Keywords:
Certainty calibration
Vehicle choice
Joint estimation
Heterogeneity
Choice certainty
Alternative acceptability

ABSTRACT

In the stated choice literature, increasing attention has been paid to methods that seek to close the gap between the choices from these experiments and the choices experienced in the real world. Attempts to produce model estimates that are truer to real market behaviours are especially important for transportation, where many important policy decisions rely on such experiments. A recent approach that has emerged makes use of a certainty index whereby respondents report how certain they are about each choice they make. Additional literature also posits that when making decisions, people first identify an acceptable set of alternatives (alternative acceptability) such that a consideration set if formed and it is from this reduced set that the ultimate choice is made. This paper presents two models that jointly estimates choice and choice certainty and choice and alternative acceptability. This joint estimation allows the modeller to overcome potential endogeneity that may exist between these responses. In comparing choices of differing certainty, surprisingly little difference in marginal sensitivities are found. This is not the case in the alternative acceptability models however. An important finding of this research is that what could be interpreted as preference heterogeneity may in fact be more closely linked to scale. The ramifications of these results on future research are discussed.

1. Introduction

The hypothetical nature of stated choice (SC) experiments has resulted in many researchers questioning whether such experiments induce ‘hypothetical bias’, a condition whereby respondents answering SC survey tasks respond in a manner other than how they would if faced with similar choices in real markets (see e.g., Hensher, 2010). Given the lack of an alternative methodology to collect preference data in many research contexts, in particular for markets where the object of study interest currently does not exist, a number of attempts over past decades have been made to improve the external validity of SC experiment outcomes. For example, researchers have explored a number of innovations, introducing methods such as information acceleration (e.g., Urban et al., 1996, 1997) where researchers create a choice environment that mimics better the context in which future consumption will be made, the combining of SC data with revealed preference data (e.g., Hensher et al., 1988; Kroes and Sheldon, 1988; Wardman, 1988) which is designed to augment the hypothetical SC data with real

* Corresponding author. Tel.: +61 2 9114 1834.
E-mail addresses: john.rose@unisa.edu.au (J.M. Rose), matthew.beck@sydney.edu.au (M.J. Beck), david.hensher@sydney.edu.au (D.A. Hensher).
[1] Tel.: +61 8 8302 1622.
[2] Tel.: +61 2 9114 1722.

http://dx.doi.org/10.1016/j.tra.2014.11.009
0965-8564/

world data, and attempts to make the choice questions more realistic (e.g., Collins et al., 2012) with the aim to make them mirror as closely as possible similar real market places. The design of incentive compatible experiments such that choices had consequential outcomes has also occurred (e.g. Rousseas and Hart, 1951, an approach replicated closely by subsequent authors such as Lusk et al., 2008 and Ding et al., 2005).

More recently, there has been a growing literature on the use of supplementary questions designed to elicit from the respondent either: *a priori* information on acceptable alternatives or attribute levels such as questions on attribute level thresholds (see e.g., Swait, 2001 and Cantillo and Ortúzar, 2006); or *ex post* such as questions as to what attributes where not considered or ignored during the SC survey (see e.g., Hensher et al., 2005; Rose et al., 2005). Whether captured *a priori* or *ex post*, this information is then incorporated in the econometric modelling of the SC data with the aim of better representing the respondents' true decision processes as reported to have occurred during the survey task (see e.g., Hess and Rose, 2007). Although the use of questions on attribute level thresholds and attribute level consideration are now common in SC experiments within the transportation field, relatively new supplementary questions are now entering the wider choice literature. In particular, there appears to be a growing interest in the use of certainty scales (see e.g., Norwood, 2005).

Certainty scales involve respondents being asked how certain they are that they would act in the manner they indicated in the SC task if confronted with a similar choice in a real market. Typically, certainty scales are asked after each SC task and involve respondents answering on some form of quantitative scale, such as a 1–10 point scale, how certain they are that they would actually purchase or choose the alternative they did (where e.g., 1 represents being 'very uncertain' whilst 10 represents being 'very certain'). As such, certainty scale questions directly deal with the degree of confidence an individual has in terms of whether they would actually choose the alternative they did in the SC experiment. Thus, unlike questions on attribute level thresholds and attribute consideration which are used to attempt to better understand the choice processes underlying SC experiments, certainty scales are specifically designed to mitigate hypothetical bias in such experiments (Ready et al., 2010). The rationale for certainty scales in the extant literature is that choices in which respondents state they are more certain they would make in the real world are more likely to approximate the real market behaviours.

Typically, certainty scale data has been incorporated into the modelling of discrete choice data in one of three ways. The first two methods involve the initial determination of some cut-off or threshold value of certainty. Once identified, the analyst then either removes from the analysis any choice task where the certainty rating fails to exceed this threshold value or alternatively, changes the observed choice to a status quo or no choice option (see e.g., Champ et al., 1997; Ethier et al., 2000; Champ and Bishop, 2001). The third approach involves exogenously weighting each choice task according to the reported certainty rating using the approach outlined in Lerman and Manski, 1981 (see e.g., Beck et al., 2013a on the use of exogenous weighting in the context of certainty scales). Treatment of certainty scales in any of the above manners however is problematic and open to criticism. Firstly, the choice of threshold or cut-off points is somewhat arbitrary and has no support in theory (e.g., using a 1–10 point scale, Champ et al., 1997 used a cut-off of certainty responses of 10, Ethier et al., 2000 a cut-off of seven, Champ and Bishop, 2001 a cut-off of eight, whilst Beck et al., 2013a used a cut-off of six). Secondly, it is probable that the certainty rating is in some manner linked to preference and as such there exists the potential for endogeneity in the modelling process, as the certainty rating may be correlated with the error term of the model.

To help overcome both of these criticisms, we jointly estimate certainty and choice. In doing so, we are still required however to rely on some form of certainty cut-off given the choice context explored in the current paper. This is because the choice context involves respondents choosing between three labelled vehicle type alternatives described by alternative fuel options after which they complete a 1–10 point certainty scale, in addition to which we construct a no choice option, thus providing for 40 choice alternatives in total to be modelled (i.e., four alternatives $\times$ 10 certainty points). As such, in the current study, we restrict the certainty scale to less than or equal to six and greater than six in line with Beck et al. (2013a) who analysed the same data. In this way, we retain a more manageable six choice options in the modelling (i.e., three alternatives $\times$ two certainty points)[3].

In addition to the joint estimation of certainty and choice, this paper also examines the role of alternative acceptability on choice (i.e. whether or not respondents find one or more of the alternatives in the choice sets they are presented with one that they would find acceptable in a real choice situation). Drawing on literature on choice set formation where respondents are asked which alternatives out of the set shown would be acceptable and hence considered at the time a choice is made (see e.g., Gilbride and Allenby, 2004), such questions have been used in the past to assign a zero probability at the time of modelling to alternatives that have been deemed to be out of the acceptable consideration set (see e.g., Horowitz and Louviere, 1995). Once again, such questions are subject to issues of endogeneity, as acceptability of an alternative is highly likely to be linked to preference and in particular to the error term of the modelled alternatives. Further, asking whether an alternative is acceptable or not does not preclude the fact that the alternative was actually considered when the final choice was made, and hence does not necessarily suggest that the alternative should be assigned a zero choice probability.

[3] While this decision resulted in a higher degree of manageability with respect to data manipulation and model estimation, previous work trialled multiple iterations of threshold points revealing that, for this data set, a threshold of six produced the best model fit criterion (see Beck et al., 2013a). This is similarly true for this research. We recommend that others who are interested in using choice certainty to calibrate choice experiments follow a similar iterative process in order to identify the appropriate threshold in their data. Additionally, we also recommend that those who use certainty scales examine the role of experimental features in determining choice certainty (see Hensher et al., 2012). Such richer insights into the behaviour of respondents with respect to the survey and the characteristics of the choice itself are invaluable in more fully understanding your research topic.