Transportation Research Part B 89 (2016) 149–167

Contents lists available at ScienceDirect

Transportation Research Part B

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

Can you ever be certain? Reducing hypothetical bias in stated choice experiments via respondent reported choice certainty

Matthew J. Beck[a,*], Simon Fifer[b], John M. Rose[c]

[a] Institute of Transport and Logistics Studies, The University of Sydney, 378 Abercrombie St, Darlington, NSW 2006, Australia
[b] Community and Patient Preference Research, Suite 112, 29 Kiora Rd, Miranda, NSW 2228, Australia
[c] Institute for Choice, University of South Australia, Level 13, 140 Arthur St, North Sydney, NSW 2060, Australia

ARTICLE INFO

Article history:
Received 11 May 2015
Revised 4 April 2016
Accepted 5 April 2016
Available online 27 April 2016

Keywords:
Hypothetical bias
Stated preference
Certainty calibration
GPS
Insurance
Revealed preference

ABSTRACT

Stated choice experiments are a preeminent method for researchers and practitioners who seek to examine the behavior of consumers. However, the extent to which these experiments can replicate real markets continues to be debated in the literature, with particular reference to the potential for biased estimates as a result of the hypothetical nature of such experiments. In this paper, a first in the transportation literature, we compare stated choice responses to revealed preference behavior and examine three methods proposed in the literature for calibrating choice experiments via reported choice certainty. In doing so we provide evidence that the incorrect calibration of responses can produce stated choice results that are more biased than doing nothing at all, however we show that by jointly estimating choice and choice certainty there is a significant reduction in hypothetical bias such that stated choice responses more directly replicate real behavior.

1. Introduction

Stated choice methods are used extensively to understand preference structures across a broad range of literatures, in particular to forecast behavioral responses to product or policy changes or to provide estimates of how respondents value both market and non-market attributes of products or services. The widespread adoption of stated choice methods in fields such as marketing, health, environment and transport are due to the advantages this approach offers with respect to reliable estimates of the relative importance of choice attributes, the ability to incorporate attributes that do not currently exist in the market, the ability to minimize confounding between estimates of effects and creation of designs that enable the efficient recovery of such effects. While these benefits exist, there have been a number of criticisms of stated choice surveys, one of which is the potential for hypothetical bias.

1.1. What is hypothetical bias?

One criticism that is often raised of stated choice experiments (and stated preference methods more widely) is that the intentions which are stated in these experiments are not the behaviors which are observed (or revealed) in actual markets (see for example Samuelson, 1955; Cummings et al., 1986; Mitchell and Carson, 1989). This discrepancy is broadly termed "hypothetical bias" in the literature, as it is argued that bias is generated by the hypothetical nature of a stated preference

* Corresponding author. Tel.: +61 2 9114 1834.
E-mail addresses: matthew.beck@sydney.edu.au (M.J. Beck), simon.fifer@cappre.com.au (S. Fifer), john.rose@unisa.edu.au (J.M. Rose).

http://dx.doi.org/10.1016/j.trb.2016.04.004
0191-2615/

experiment wherein respondents are not obliged to actually carry out the choices or behaviors they state, or are unable to fully predict their own real market behavior in a hypothetical setting like a stated choice survey. There is no widely accepted general theory of respondent behavior that explains hypothetical bias (Loomis, 2011), so in the context of this paper we use the term hypothetical bias to simply refer to discrepancies between preferences exhibited in the hypothetical choice experiment and the preferences revealed by actual behavior.

Before addressing hypothetical bias in more detail, we would like to draw the reader's attention to a wider concept. In much, if not all, of the literature on hypothetical bias there is very little reference to external validity. This connection should be made as external validity has been a long standing concept in statistics. External validity relates chiefly to generalizability; the degree to which the conclusions in a study would hold for other persons in other places at other times. For a comprehensive review of external validity in the social sciences we refer to the reader to Lucas (2003). Included in this paper is a summary of the typical approaches to which external validity is assessed (many of which are applicable to studies of stated choice). Specifically, to have external validity a study must ensure: *construct validity* (the extent to which measure accurately reflect the theoretical concepts they are intended to measure); *relevance* (the degree to which the experimental situation designed to capture the theory adheres to the scope of the theory being tested); *reproducibility* (research findings can be replicated successfully if the research is performed again); *consistency* (the extent to which observations in the study are consistent with each other and the theory being tested); and *confirmatory* (the extent to which the proposition has been supported in numerous tests in diverse settings.

There has been work in the stated preference literature that looks broadly at external validity, typically making combined use of stated and revealed preference data to assess how responses from stated preference experiments can be generalized to real market behavior (see for example Louivere, 1988; Ben-Akiva et al., 1994 and Herriges et al., 1999) using methods such as the nested-logit or error components model to ensure parameter estimates across data types are not confounded by differences in scale (Hensher and Bradley, 1993; Hensher et al., 2008). These methods (and those used in the rest of this paper) are focused on aligning stated preference data with revealed behaviors, which is only partly concerned with the concept of reproducibility; partly because all of the studies on hypothetical bias examine methods to bring the results from one study (on one sample) of stated preferences to one study (on one sample) of revealed behaviors. It is our belief that hypothetical bias, in the literature, has come to mean external validity, whereas external validity is a much more exhaustive concept that requires a number of conditions to be satisfied.

While our discussion addresses methods to reduce hypothetical bias in stated preference studies, an area to which much more attention should be given, we feel that it is important to state that reductions in hypothetical bias do not automatically equate to external validity, rather it is a narrow focus on a much wider issue.

1.2. Does hypothetical bias exist?

The exploration of this source of bias has been most extensive within the contingent valuation literature. A 1994 issue of the Journal of Economic Perspectives included a symposium on contingent valuation, with three articles discussing the role of hypothetical bias in this method (Portney, 1994; Hanemann, 1994; Diamond and Hausman, 1994). For a broader overview of the impact of hypothetical bias in contingent valuation methods, List and Gallet (2001), Little and Berrens (2004) and Murphy et al., (2005) all perform meta-analysis on a number of contingent valuation studies and conclude that hypothetical bias is a major concern, with median bias levels ranging anywhere from 25% to 300%.

Given the significant differences between the stated and revealed values in contingent valuation studies, it is unsurprising that similar exploration has begun to emerge in the stated preference literatures. Evidence of hypothetical bias can be observed in a transportation context, for example Brownstone and Small, (2005) observed that, in a toll road study in California, the revealed value of travel time savings in the morning commute is $20–$40 per hour, which is more than double the values estimated from stated preference studies of the same travel choices. This evidence is supported by Isacsson, (2007) who finds that stated preference values of time are understated for buses, and by Wardman and Shires, (2001) who find that stated valuations similarly overestimate actual values in the context of penalties for having to change trains. In contrast, Wardman and Whelan, (2001) find implausibly large stated values for new or improved trains across 45 stated preference studies. Hensher, (2010) compares values of travel time for a number of different data sets and concludes that differences are generally not statistically significant. Loomis, (2011) also found evidence that hypothetical willingness to pay values exceed actual values by a factor of two to three, though is not always present in stated choice surveys.

Outside of transportation, three of the more interesting studies that examine the phenomenon of hypothetical bias can be found. Chang et al., (2009) compare hypothetical choices of ground beef, wheat flour and dishwashing liquid to actual retail shopping behavior and find evidence that hypothetical choices are a poor predictor of changes to market share, as calculated by the mean square error around predicted and actual shares in each product category. Miller et al., (2011) examine choice of a cleaning product for high-tech equipment, using a tailored online store to gather real purchasing data, and also find the existence of hypothetical bias. Interestingly however, they conclude that stated preference experiments may still lead to the right demand curves and right pricing decisions. Hudsen et al., (2012) investigate the choice of a new product, freshwater prawns, using a mail survey and a controlled in-store experiment. Overall the authors find that hypothetical bias is not present in the choice of the new product (freshwater prawns) but that it is present in the choice of the substitute product (lobster).