Ecological Economics 131 (2017) 87–97

Contents lists available at ScienceDirect

Ecological Economics

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

Analysis

Temporal stability of stated preferences for endangered species protection from choice experiments☆

Daniel K. Lew [a,b,*], Kristy Wallmo [c]

[a] Resource Ecology and Fisheries Management Division, Alaska Fisheries Science Center, National Marine Fisheries Service, United States
[b] Department of Environmental Science and Policy, University of California, Davis, United States
[c] Office of Science and Technology, National Marine Fisheries Service, United States

ARTICLE INFO

Article history:
Received 1 December 2015
Received in revised form 3 August 2016
Accepted 7 August 2016
Available online 31 August 2016

JEL classification:
Q51
Q57

Keywords:
Temporal stability
Willingness to pay
Stated preference
Choice experiments
Generalized mixed logit models
Endangered species

ABSTRACT

Benefit transfer methods rely on past models and results, so it is important to know whether economic values are stable over time or are subject to change, either because of the reliability of the methodology or due to actual preference changes. The temporal stability of willingness to pay (WTP) has been tested extensively for contingent valuation, but rarely for stated preference choice experiments (CE). We use data from two identical CE surveys on different samples from the same population that occurred 17 months apart (Spring 2009 and Fall 2010) to estimate and compare mean WTP and preference parameters associated with threatened and endangered marine species protection. Our models account for both preference and scale heterogeneity, and the results suggest both types of heterogeneity matter. Tests of preference stability suggest stable preferences between 2009 and 2010. Furthermore, WTP values estimated from both surveys are not statistically different. This provides evidence that economic values estimated using CE methods are temporally stable.

Published by Elsevier B.V.

1. Introduction

Due to the cost and time required to rigorously conduct de novo valuation studies, policy analyses involving natural resource and environmental goods are increasingly turning to benefit transfer methods to incorporate economic benefits. Benefit transfer methods involve the use of economic value information from previously conducted studies being applied to new applications for which they were not originally intended (Navrud and Ready, 2007; Johnston and Rosenberger, 2010). Since these methods rely on past models and results, it is important to know whether economic benefit estimates are stable over time or vary because of either a lack of reliability of the methodology or due to actual preference changes.

The temporal stability of willingness to pay (WTP) has been tested extensively in the context of contingent valuation (CV) methods. For example, Skourtos et al. (2010) summarized 20 CV-based temporal stability studies, noting considerable differences in temporal period examined (2 weeks to 20 years), but generally finding support for stability over periods of <5 years. Temporal stability of economic values from stated preference choice experiments (CE) have also begun to be assessed in recent years (e.g., Liebe et al., 2012; Bliem et al., 2012), but thus far only over relatively short time periods (2 weeks to 1 year). There is mixed evidence of the stability of preferences over these time periods in these studies, though WTP estimates were generally found to be stable.

There are two principal types of studies that investigate the temporal stability of stated preferences (SP) and values. The first arises from the literature devoted to assessing the reliability of stated preference values, which arose in part due to concerns over the temporal reliability of CV values with respect to information effects and learning within a broad assessment of the CV method (Cummings et al., 1986). These studies repeat the same survey on the same respondents in a "test-retest" format to assess whether an individual's preferences change over time. In one of the earliest studies to assess the temporal reliability of SP values, Kealy et al. (1988) conducted an in-class experiment on a sample of undergraduate students where they asked the same CV questions to the same set of students at two points in time two weeks apart. They found that WTP for a chocolate bar (a private good) remained stable over the period. Subsequent test-retest CV studies in the economics literature have evaluated the stability of WTP for a range of public goods, such as acid rain reduction (Kealy

☆ This article and its findings are those of the authors and do not necessarily reflect the views of the National Marine Fisheries Service or the U.S. Department of Commerce.

* Corresponding author at: Resource Ecology and Fisheries Management Division, Alaska Fisheries Science Center, National Marine Fisheries Service, 7600 Sand Point Way NE, Seattle, WA 98115, United States.
E-mail address: Dan.Lew@noaa.gov (D.K. Lew).

http://dx.doi.org/10.1016/j.ecolecon.2016.08.009
0921-8009/Published by Elsevier B.V.

et al., 1990), water quality improvements (Cameron, 1997), lake level improvements (Loomis, 1989, 1990), reductions in health and environmental risk (Shiell and Hawe, 2006; Brouwer et al., 2008), and recreation opportunities (Teisl et al., 1995; McConnell et al., 1998). These CV studies examined time periods between one week and two years, and in most cases they found support for temporal stability. In addition, recent test-retest studies using CE methods have been conducted to evaluate temporal reliability of values for wind power generation (Liebe et al., 2012), natural areas (Schaafsma et al., 2014), and organic apples (Morkbak and Olsen, 2014) over time periods as short as two weeks and as long as a year. These test-retest CE studies generally report evidence of temporal stability of WTP, but mixed results with respect to underlying preferences.

The second type of temporal stability test involves administering a survey in two different time periods to samples from the same population but with different people; that is, "distinct" samples are used. While a primary focus of test-retest studies is typically on the correlation between an individual's SP responses collected at different points in time, the distinct samples test approach is primarily concerned with the stability of WTP and preferences for the population over time. This population-level focus is the preferred level of analysis for a test assessing the feasibility of conducting benefits transfer, rather than the individual-level analysis the test-retest approach enables. For population-level analyses, test-retest samples suffer from sample attrition, which jeopardizes the representativeness of the retest sample, potentially precluding population-level inferences.[1]

The first study to employ the distinct samples approach to evaluate temporal stability of CV values was Reiling et al. (1990), who surveyed distinct samples in two time periods using an open-ended CV question that elicited preferences for a black fly control program in Maine. They did not find any statistically significant differences from estimated WTP between the two samples; however the time period was short—one month. A number of other studies applied this approach to longer time periods, including several studies that administered surveys 5 years apart (Whitehead and Hoban, 1999; Brouwer and Bateman, 2005; Whitehead and Aiken, 2007) and one almost 20 years apart (Boman et al., 2011). As noted above, studies of these longer time periods usually did not find support for temporal stability. Only two CE studies to date test for temporal stability with distinct samples in the environmental valuation literature, one that found the WTP for river restoration in Austria over a one year period (Bliem et al., 2012) was, in general, temporally stable, and another that found preferences over a 4 month period for laundry detergents that are produced in different ways with differing ecological impacts were temporally stable (Arana and Leon, 2013).

In this paper, we conduct a distinct samples temporal stability test using data from two CE surveys administered to samples from the same population 17 months apart to estimate and compare WTP and preferences associated with providing protection to threatened and endangered (T&E) marine species. Although the literature on SP values for T&E species has grown considerably in recent years (Richardson and Loomis, 2009), this work represents the first direct evaluation of the temporal stability of individual T&E species values.[2] Moreover, it also examines stability over the longest time period over which CE values have been evaluated - 17 months in this study versus one year in both Bliem et al. (2012) and Schaafsma et al. (2014).

The CE data are analyzed using several discrete choice models, including one that allows for both scale and preference heterogeneity while also accounting for the panel nature of the data. In our application, this model, a form of the generalized multinomial logit model (Fiebig et al., 2010) and a random parameters, or mixed logit, model (Train, 2003) are used to conduct tests for preference stability and the stability of WTP values. Results suggest stability of both preference functions and WTP values across time.

The remainder of the paper is organized as follows. The next section provides additional details of distinct examples of temporal stability studies that inform our empirical approach for evaluating preference stability. This is followed by a description of the study design and implementation, as well as the associated data used in the analysis. We then describe the specific modeling approach used to analyze the data. Next, we present the model results and discuss our investigation into the influence of demographics on our results. We conclude with a discussion of the findings and suggestions for areas of further research.

2. Testing Temporal Stability with Distinct Samples

Temporal stability tests using the distinct samples approach generally involve testing for the stability of WTP (e.g., Reiling et al., 1990; Whitehead and Aiken, 2007), which involve testing the following hypothesis (Hypothesis 1, H1):

H1. $\overline{WTP}_1 = \overline{WTP}_2$,

where $\overline{WTP}_t$ is the unadjusted mean WTP estimated for the sample in time period $t \in \{1,2\}$. An inability to reject the null hypothesis suggests sample mean values are consistent in the two time periods. McConnell et al. (1998) discusses how testing the equality of WTP does not account for potential changes in preferences and argues for an additional test to assess the stability of preferences that tests (Hypothesis 2, H2):

H2. $\beta_1 = \beta_2$,

where β_t is the parameter vector for the preference function in time period t ($t = 1,2$). Most recent studies employ this test for preference function stability in addition to testing H1 (e.g., Brouwer, 2006; Downing and Ozuna, 1996; Carson et al., 1997). Note that for evaluating the temporal stability in the context of transferring unit values and preference functions in benefit transfer applications, both tests are necessary.

The manner in which preference stability is tested depends upon the type of SP question employed. Several CV studies, for instance, pool data across time periods and estimate time-dependent parameters (e.g., Downing and Ozuna, 1996; Carson et al., 1997), then test for equality of the time-dependent preference parameters. For example, Brouwer and Bateman (2005) conducted a test for stability using a referendum CV survey related to flood protection in the U.K. administered to two distinct samples 5 years apart, in 1991 and 1996. Two temporal stability tests were done: one comparing the statistical equality of unadjusted average WTP and another evaluating the equality of coefficient estimates and equality of variances of the estimated preference function. Carson et al. (1997) used two surveys done two years apart to show that the WTP for protecting Prince William Sound, Alaska, from future oil spills was temporally stable. Distinct random samples were drawn from the same population in 1991 and 1993 and given the same in-person survey (interview). To evaluate stability of the preferences, they compared estimates from several discrete choice models estimated by pooling the data for the two surveys. For each type of discrete choice model, they separately estimated the effects from (1) a dummy variable representing the data are from the 1993 survey (that is, a shifter) and (2) both a shift dummy and slope shifters.

[1] In addition, the test-retest approach potentially suffers from carryover effects, which occur when respondents base their answers in the retest on what they recall from how they answered the original test (McConnell et al., 1998). These carryover effects can confound the results.

[2] In a meta-regression of endangered species non-market values, Richardson and Loomis (2009) included both a dummy variable to represent studies conducted on or after 1995 that differ from those before then, as well as a year variable to capture time effects on species values. Their results suggest there is a difference between values from pre-1995 and post-1995 studies, with newer studies leading to larger values in some models. While this can be viewed as providing some evidence of increasing (general) species values over time, it is not a robust test given that the approach introduces considerable potential for misspecification biases related to the ability to control for the characteristics of the studies and values included in the model. The work represented in this study provides a more direct test of values for specific individual species.