VALUE IN HEALTH ▮ (2018) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

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

# Assessing Rationality in Discrete-Choice Experiments in Health: An Investigation into the Use of Dominance Tests

*Tommi Tervonen, PhD*[1,*], *Tabea Schmidt-Ott, MSc*[1,2], *Kevin Marsh, PhD*[1], *John F.P. Bridges, PhD*[3], *Matthew Quaife, PhD*[2], *Ellen Janssen, PhD*[3]

[1]*Evidera, London, UK;* [2]*Department of Global Health and Development, London School of Hygiene and Tropical Medicine, London, UK;* [3]*Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA*

ABSTRACT

**Background:** Dominance tests are often applied to test for the rationality in the choice behavior of participants in discrete-choice experiments (DCEs). **Objectives:** To examine how dominance tests have been implemented in recent DCE applications in health and discuss their theoretical and empirical interpretation. **Methods:** Health-related DCEs published in 2015 were reviewed for the inclusion of tests on choice behavior. For studies that implemented a dominance test, information on application and interpretation of the test was extracted. Authors were contacted for test choice sets and observed proportions of subjects who chose the dominated option. Coefficients corresponding to the choice set were extracted to estimate the expected probability of choosing the dominated option with a logistic model and compared with the observed proportion. The theoretical range of expected probabilities of possible dominance tests was calculated. **Results:** Of 112 health-related DCEs, 49% included at least one test for choice behavior; 28 studies (25%) included a dominance test. The proportion of subjects in each study who chose the dominated option ranged from 0% to 21%. In 46% of the studies, the dominance test led to the exclusion of participants. In the 15 choice sets that were analyzed, 2 had larger proportions of participants choosing the dominated option than expected ($P < 0.05$). **Conclusions:** Although dominance tests are frequently applied in DCEs, there is no consensus on how to account for them in data analysis and interpretation. Comparison of expected and observed proportions of participants failing the test might be indicative of DCE quality.

***Keywords:*** discrete-choice experiment, internal validity, preference, random utility theory.



## Introduction

Discrete-choice experiments (DCEs) are used in health care decision making to elicit the stated preferences of stakeholders (e.g., patients and clinicians) on the attributes of treatments and other health care goods and services [1–3]. The use of DCEs in health-related topics has steadily increased over the years [1]. Their importance in health policy decision making is highlighted by the recent regulatory interest in their usefulness for capturing patient preferences on treatment benefit-risk trade-offs [4,5].

DCEs are founded on Lancaster's theory of consumer behavior [6] and commonly analyze choice data with McFadden's random utility model (RUM) [7]. The theory of consumer behavior assumes that the participants are utility-maximizing agents and are willing to trade off between attributes in the choice experiment. Accordingly, stated preferences captured by DCEs should conform with axioms of rational choice, such as completeness, transitivity, and monotonicity [8]. Different ways to identify DCE subjects whose choice behavior violates common rationality axioms have been proposed in the literature [2]. The most frequently applied test in DCEs has been the dominance test [1] where, given researchers' a priori assumptions on attribute level ordering, one of the choice alternatives is clearly superior. Participants who choose the dominated alternative are considered to have failed the test. These participants may not have understood the choice task, may not have paid sufficient attention to it, or may have been exhibiting nonrational choice behavior.

In many studies, participants who fail the dominance test are excluded from final data analysis [9–12]. Nevertheless, this interpretation of the dominance test is challenged by the fact that a certain proportion of participants are expected to "fail" the dominance test because of the probabilistic property of the RUM [13]. That is, the RUM includes an error component that can be interpreted in various ways that can account for the seemingly irrational behavior. First, this might result from measurement error—the subjects understand the task and attend to the questions appropriately, but make a mistake and answer the dominated choice question incorrectly. Given this, their responses to other questions may still reflect their true, rational preferences. Second, this might result from

*Address correspondence to: Tommi Tervonen, Evidera, 1 Butterwick, Metro Building, 6th Floor, London W6 8DL, UK.
E-mail: tommi.tervonen@evidera.com.

1098-3015$36.00 – see front matter 

https://doi.org/10.1016/j.jval.2018.04.1822

unobserved variables—the participants might infer additional information on attribute value beyond what is presented in the DCE. For instance, they may infer a higher quality of care from the cost of care. This leads to choice behavior that is seemingly not in line with the researchers' a priori expectations of what is rational behavior, but when the additional information is taken into account, the participants are actually choosing according to the axioms of rationality [14], and thus most of their preference data are likely to be useful for the analysis.

This article aims at understanding how dominance tests can and should be used for supporting choice validity assessment in DCEs, with a focus on the health domain. Past reviews of DCEs in health have surveyed the frequency of use of dominance and other rationality tests [1–3]. Our contribution surveys dominance tests in the literature in more depth, to assess how authors have used dominance tests and to analyze whether the test results have been interpreted in line with what the RUM predicts.

## Methods

### Data Source

DCE studies published in 2015 were identified from a systematic review on how qualitative methods have been used to support health DCEs [15]. The year 2015 was chosen because it reflected recent use of dominance tests and was expected to cover a wide spectrum of current applications. The articles were reviewed for general study characteristics including sample type, area of application (as defined by de Bekker-Grob et al. [2]), form of application (categories of self-completion and personal interview), funding of the study, context of country (defined by the World Bank classification of high-, middle-, or low-income countries [16]), the application of choice behavior tests (defined by a review of methodological studies [17]), and whether a pilot study has been conducted.

From the 126 DCE articles published in 2015 [15], 112 elicited preferences and reported empirical data. Eleven studies focused on the development of DCEs rather than on the elicited preferences, and so these studies were excluded. Another two studies were excluded because their preference elicitation method did not include choices between multiple profiles. Two studies presented the same DCE, and one of them was excluded to avoid double counting. Table 1 presents an overview of the extracted study characteristics.

For the 28 studies reporting the use of the dominance test, data on application and interpretation of the test were extracted, including the levels of the attributes used for the design of the dominant and dominated choice alternatives, the observed proportion of respondents choosing the dominated option, the authors' reasoning for the use of dominance test, how authors dealt with participants who "failed" the dominance test, and the conclusions that the researchers drew from the test. Because this information was not provided in the articles, all authors were contacted with a request for this information. Fourteen (50%) authors responded and provided the requested data. Because one of the studies included 2 separate DCEs, the analysis was conducted on 15 DCEs.

### Calculations

The expected probability of passing the dominance test was computed for the studies for which data were available using a logit model. This model assumes that the individual respondent's ($n$) utility ($u$) toward a specific choice option ($j$) is a function of the measured utility ($v$) of the presented attribute levels ($X_{nj}$) and an unexplainable error component ($\varepsilon_{nj}$):

$$u_{nj} = v_{ij}(X_{nj}\beta) + \varepsilon_{nj}, \tag{1}$$

where $\varepsilon_{nj}$ are independently and identically distributed (following a Gumbel distribution with a location $\eta$ and a scale $\mu > 0$). Then, in a two-choice alternative setting, the probability that choice alternative $j$ is chosen over choice alternative $i$ is

$$P_{ij} = \frac{\exp(X_{nj}'\beta)}{\exp\left(X_{nj}'\beta\right) + \exp(X_{ni}'\beta)} \tag{2}$$

The expected probability of choosing the dominated alternative ($p_e$) is estimated using Equation 2, with the attribute levels set to reflect those of the choices in the dominance test. The theoretical range of expected probabilities of possible dominance tests is $p_e min - p_e max$. The $p_e max$ is derived from the hypothetical dominance test with the largest expected probability of choosing the dominated alternative, that is, with the smallest possible utility difference between the dominant and dominated alternatives given the estimated $\beta$ and levels $X_{nj}$. Similarly, $p_e min$ is derived from the hypothetical dominance test with the smallest expected probability of choosing the dominated alternative (see Fig. 1). Both $p_e min$ and $p_e max$ are normally less than 50% and the range they span necessarily includes $p_e$.

The observed number of participants choosing a dominated option was divided by the total number of participants to calculate the observed proportion of participants choosing the dominated option. A z test was conducted to assess whether the observed proportion ($p_o$) was equal to the calculated expected probability of choosing the dominated option ($p_e$). A test statistic with a $P$ value of 0.05 or less was considered significant. All calculations were done using Microsoft Excel 2013. The extractions and calculations were quality-controlled by a second member of the study team.

## Results

The dominance test was the most frequently reported test of choice behavior, applied in 28 studies (25%) (Table 2). The studies with dominance tests were comparable with the studies without dominance tests in terms of "area of application," "country of application," "application form," and whether a pilot study had been conducted (Table 1). Nevertheless, the sample type varied between the studies; studies that used the dominance test were significantly more frequently undertaken with patients (Pearson $\chi^2 = 7.13$; $P < 0.01$).

In the 28 studies with a dominance test, the study authors used a range of terminology to refer to it: a control measure or a consistency, validity, internal validity, logic, or rationality test. The most frequently expressed reason for including a test was concern about the participants' understanding of the choice task (eight studies, 29%); the second reason was concern about participants lacking attention (four studies, 14%), and the third reason was to test for rationality (three studies, 11%). In 46% of the studies, the dominance test led to the exclusion of participants, on the basis of either a single dominance test or a combination of tests. Ten studies (36%) tested the effects of participant exclusion on the model, and in all instances, the model was not sensitive to participant exclusion. The dominance tests in 3 of the 28 studies resulted from the generation of the choice sets as part of the experimental design, and therefore were not intentionally designed [18–20].

In the 14 studies that provided additional data, the proportion of subjects choosing the dominated option ($p_o$) ranged from 0% to 21% (Fig. 2). Most of the studies had low $p_o$, with only a single