

BACKGROUND

- Swing-weighting (SW) and discrete choice experiments (DCEs) are two possible techniques for developing scoring functions in the context of multiple criteria decision analysis (MCDA).
- The benefit of SW is that it uses fewer respondents (i.e., smaller sample size) but more information is obtained per respondent than DCE. Also, internal validity of SW is automatically enforced but with DCE needs to be assessed.(1)
- Although SW can obtain more information per respondent than DCE, the task of eliciting a response may be considered more complex of a task than DCE.
- A critical assessment of SW and DCE has been performed. The assessment evaluated both methods in terms of comprehension by study participants, cognitive biases, sample representativeness, ability to capture heterogeneity in preferences, reliability and validity, and robustness of the results. (1)
- If SW is as reliable an estimation method as DCE, it is worth exploring whether it is easier to use in practice than DCE.
- This study is a first step to validate the method versus DCE and a known underlying population.

OBJECTIVES

- The objective of this study is to compare estimates from both methods using a simulation study.

METHODS

Simulation Study

- A fictitious utility model representing preferences of a population among 5 attributes, each with 3 levels, was defined as the sum of marginal utilities over attributes.
- The model allowed for variability around marginal utilities between people and included a random term capturing aspects of the utility not related to the 5 attributes. The marginal utility distributions are shown in Table 1.

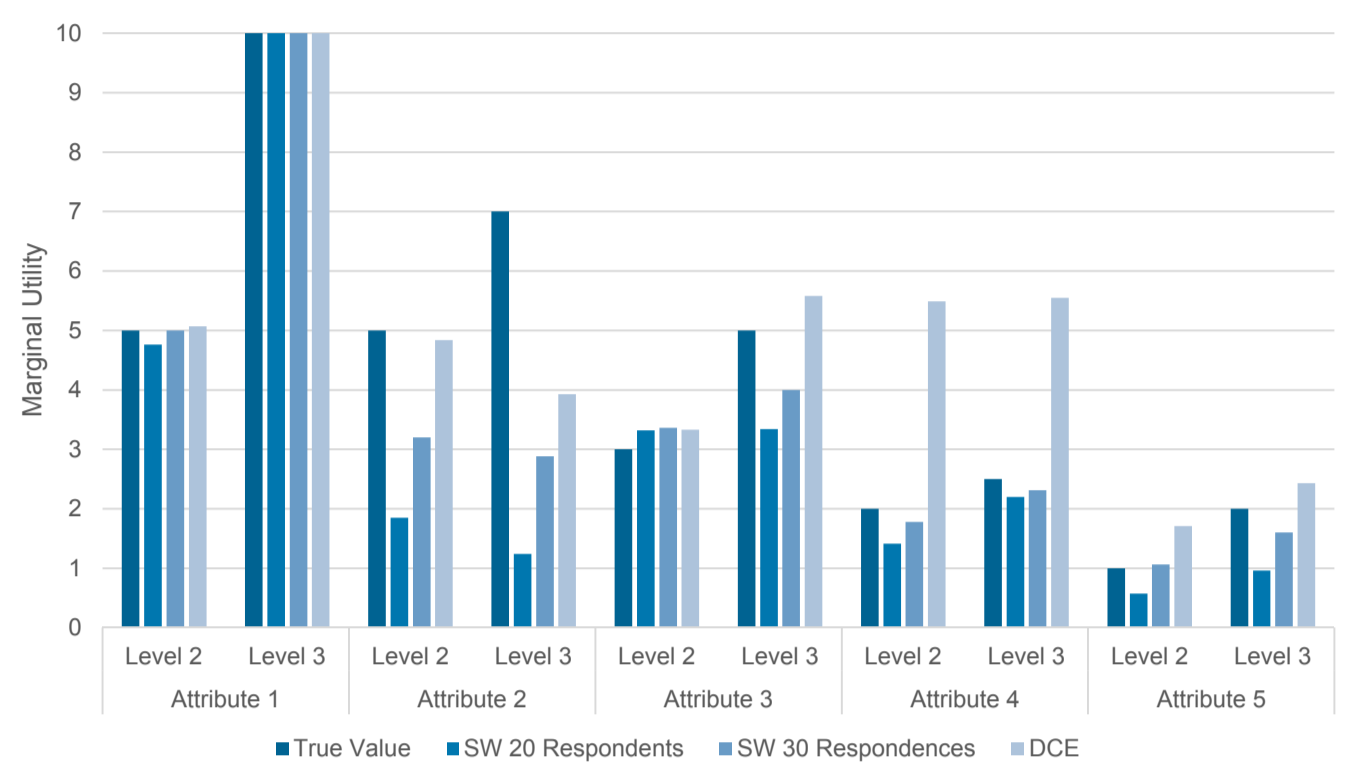
Table 1: Distribution of partial utility associated with each attribute's level			
Attribute	Utility value for each attribute level [Distributed normally with N(mean, sd)]		
	Level 1	Level 2	Level 3
Attribute 1	0	N(5, 5)	N(10, 10)
Attribute 2	0	N(5, 10)	N(7, 15)
Attribute 3	0	N(3, 2.5)	N(5, 5)
Attribute 4	0	N(2, 2.5)	N(2.5, 5)
Attribute 5	0	N(1, 0.5)	N(2, 1)

- The DCE was assumed to be conducted using an orthogonal balanced design with 18 choice tasks.
- Responses to choices were simulated randomly, based on probabilities calculated as a logit function of utilities.
- The DCE was assumed to be conducted among 200 individuals.
- The SW were assumed to be conducted by between 10 and 40 individuals.
- The marginal utilities for each attribute and level were compared between the DCE, SW, and population value after scaling so level 1 equaled zero and attribute 1, level 3 equaled 10.
- The correlation was calculated between the estimated utility scores from SW, the estimated utility scores from the DCE model and the underlying "true" utility. The correlation is calculated over every possible combination of levels.

RESULTS

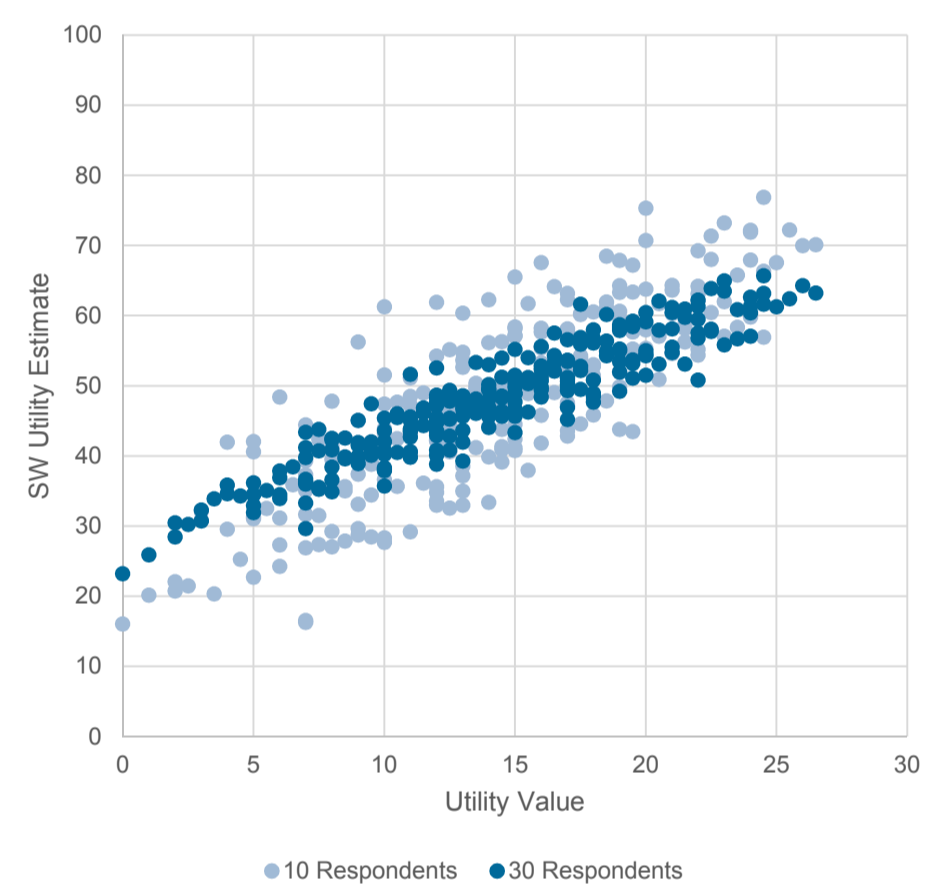
- The marginal utilities when SW is estimated using 20 or 30 respondents or DCE with 200 are shown in Figure 1. Attribute 4 is especially hard for the DCE to estimate, and Attribute 2 for SW. With more respondents both methods appear to be converging to the correct values.

Figure 1: Marginal utilities for the SW method simulating 20 and 30 respondent, the DCE simulating 200, and the true underlying values



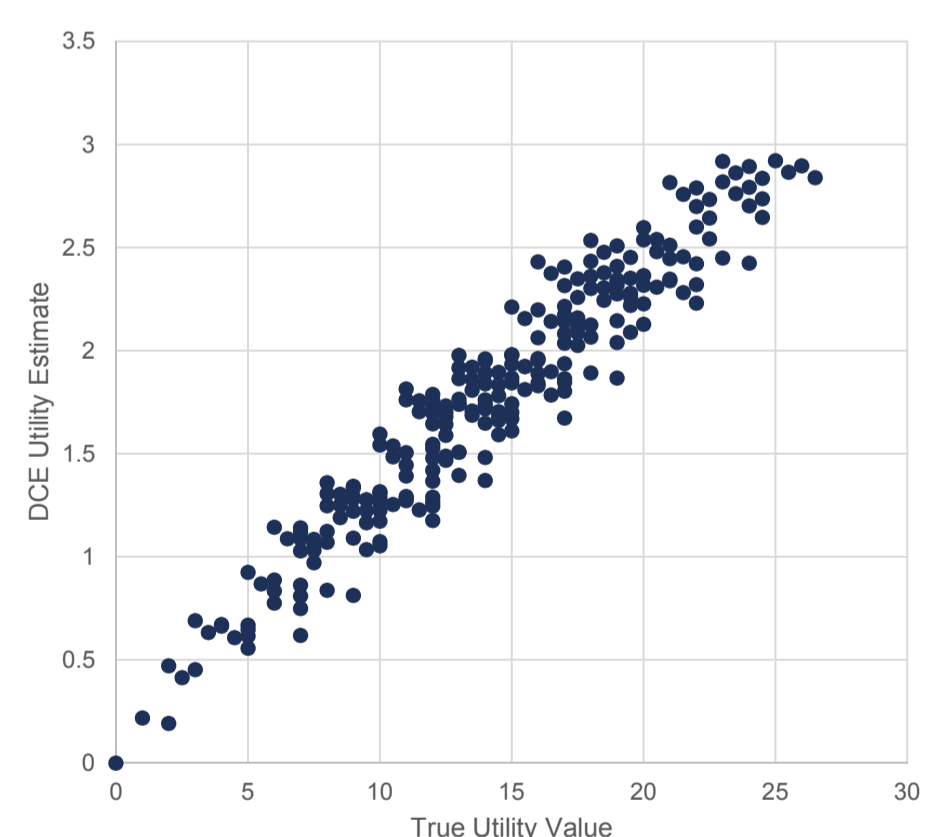
- The correlations between SW score and true utilities were 0.83, 0.84, and 0.93 with 10, 20, and 30 respondents, respectively. Figure 2 shows the relationship between the DCE utility estimates and the true values for every possible combination of attribute levels for 10 and 30 respondents.

Figure 2: Relationships between DCE utility estimates and SW utility estimates and the true value for 10 and 30 respondents.



- The correlation between DCE utility estimates and the true utilities was 0.96. Figure 3 shows the relationship between the DCE utility estimates and the true values for every possible combination of attribute levels.

Figure 3: Relationships between DCE utility estimates and the true value for 200 respondents.



LIMITATION

- This simulation study does not account for biases due to different ways of framing the questions in DCE and SW, and potential misunderstanding of the tasks by respondents.

CONCLUSIONS

SW and DCE scores both correlate relatively well with true utility. A DCE with 18 choice tasks and 200 individuals performed as well as a SW with 30 individuals when similar levels of error were assumed. This simulation framework could be applied when deciding which technique to use when planning a MCDA.

REFERENCES
1. Tervonen T, Gelhorn H, Sri Bhashyam S, et al. MCDA swing weighting and discrete choice experiments for elicitation of patient benefit-risk preferences: a critical assessment. *Pharmacoepidemiology and drug safety* 2017;26:1483-91.

