Sample restrictions and the elicitation of a constant WTP per QALY

Jytte Seested Nielsen[1], Dorte Gyrd-Hansen[2] and Trine Kjær[2]

[1]Newcastle University Business School, 5 Barrack Road, NE1 4SE Newcastle upon Tyne, United Kingdom

[2]DaCHE, Department of Public Health, University of Southern Denmark, J.B. Winsløws Vej 9B, DK-5000 Odense C, Denmark

Contact per e-mail: Dorte Gyrd-Hansen dgh@sdu.dk**; Trine Kjær** tkj@sdu.dk**; Jytte Seested Nielsen (corresponding author)** jytte.nielsen@ncl.ac.uk

Extended Abstract

There is an ongoing debate about which value to place on a quality adjusted life year (QALY) and appropriate ways of estimating such a threshold. One approach for establishing an individual value for a QALY is the stated preference method in which individuals are asked to value hypothetical health states both in terms of QALYs and willingness to pay (WTP). It is well established that the underlying theoretical assumptions needed to obtain a constant proportional trade-off between QALY and WTP, and thus a constant WTP per QALY (WTP-Q) estimate, are restrictive (Gyrd-Hansen 2005). Despite this, and as recently emphasized by Sund and Svensson (2018), empirical studies on the value of a QALY are important as they provide valuable input to decision makers on the range of WTP-Q estimates in a population, and improve our understanding of the factors that influence the relationship between the two metrics.

One dilemma inherent to the use of stated preference data is how to legitimately decide if some observations/respondents can be considered inconsistent and irrational and whether these should be excluded from the analysis. When it comes to establishing a WTP-Q based on individual stated preferences, validity issues are especially pertinent. In such cases, WTP as well as QALY measurement may be subject to bias. To the authors' knowledge, no study has previously attempted to systematically investigate how the use of different exclusion/restriction criteria impact on the validity of WTP-Q estimates. This study seeks to fill this gap in the

literature. We contribute by systematically analysing the impact of applying different exclusion and restriction criteria to ascertain whether these sample restrictions have a positive effect on a range of specified validity dimensions. In particular, we assess whether sample selection improves performance relating to scope sensitivity, sensitivity to health state severity and elicitation approach. We adopt a step-wise approach to excluding respondents using general accepted validity criteria including the elimination of inconsistent respondents both in terms of QALYs and WTP. Furthermore, to examine the extent to which inter-respondent data performs better, we restrict our sample to include only the first or last WTP question eliminating all intra-respondent variation.

For this purpose, we revisit the Danish data from the EuroVaQ survey comprising a 'chained' approach of deriving WTP-Q estimates (Robinson et al 2013). In each chain, respondents were first asked to complete either a standard gamble (SG) or a time-trade-off (TTO) exercise, in order to ascertain the QALY value of a given health state. Subsequently, respondents were randomized to one of two WTP questions. Either they were to imagine that they were in the aforementioned health state and asked to state their WTP for reducing the time spent in the health state, or they were to imagine that they were at risk of falling into the health state and asked to state their WTP for reducing the risk. The time and risk reduction were tailored according to the responses to the previous SG/TTO questions such that all respondents were asked to state their WTP for the same QALY gains (0.1 and 0.05 in random order). WTP was elicited using the random card sorting procedure. The set of payment cards presented to respondents aimed to keep the 'range' constant in terms of implied WTP-Q across the scenarios where 0.05 and 0.1 QALY was on offer. WTP was elicited using the random card sorting procedure.

We use random effects log-linear WTP models to analyse our data. Overall, we observe some very consistent patterns irrespective of restriction criteria. This indicates that the factors influencing WTP-Q are pertinent across different sub-samples and also present among the (few) respondents who pass the most stringent validity criteria. This seems to suggest that the observed violations, causing a non-constant WTP-Q, constitute widespread preference patterns and not merely a result of poor responses in stated preference methods.

References

Gyrd-Hansen, D. (2005). Willingness to pay for a QALY. Pharmacoeconomics, 23(5), 423-432.

Robinson, A; Gyrd-Hansen, D; Bacon, P; Baker, R; Pennington, M; Donaldson, C; EuroVaQ-team (2013). Estimating a WTP-based value of a QALY: The `chained` approach. Social Science & Medicine 92; 92-104

Sund, B., Svensson M (2018). Estimating a constant WTP for a QALY – a mission impossible? European Journal of Health Economics 19(6); 871-880