Social Science & Medicine 315 (2022) 115474

Contents lists available at ScienceDirect

Social Science & Medicine

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

Do they just know more, or do they also have different preferences? An exploratory analysis of the effects of self-reporting serious health problems on health state valuation

Jack Elliott[a,*], Aki Tsuchiya[b,c]

[a] Health Organisation, Policy and Economics, Division of Population Health, Health Services Research and Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

[b] Department of Economics, University of Sheffield, Sheffield, UK

[c] School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK

ARTICLE INFO

Keywords:
Self-reported health
DCE with Duration
EQ-5D
Health state valuation
Whose preferences
Hypothetical valuation
Experience-based valuation

ABSTRACT

Background: Health state valuation is often conducted by people valuing either only their own health state (experience-based valuation) or several stylised states (hypothetical valuation). The approach used can affect the elicited values, but it is not clear whether this is caused by different understandings of the states (an "*information mechanism*") and/or by different opinions of the states (a "*preference mechanism*"). Justifying privileging the valuations of those with the relevant health state experience solely because they are better informed is insufficient when their opinions differ. This study proposes a new framework to examine the effect of having health problems on health state valuation by distinguishing "within-dimensional" effects (which can be due to better information or due to differences in opinion) and "cross-dimensional" effects (which must be due to differences in opinion).

Methods: Secondary data from the UK that valued EQ-5D-5L using Discrete Choice Experiments with duration (DCE_{TTO}) are remodelled controlling for whether a respondent self-reports serious (viz., severe or extreme) problems in "pain or discomfort" (PD) or "anxiety or depression" (AD). The main analysis uses respondents who have serious PD or serious AD alongside matched respondents who do not, and assumes constant proportional time trade-off.

Results: Self-reporting serious PD or serious AD problems has no within-dimensional effect on health state preferences. However, self-reporting serious AD problems has negative cross-dimensional effects on the utility of having any problem in PD, which suggests that the preference mechanism is present. A similar pattern holds when all available (unmatched) data are used and when constant proportional time trade-off is relaxed.

Conclusions: Findings consistent with the preference mechanism indicate that those with serious health problems may have different opinions on the value of health states compared to the rest of the population, which has normative implications for the debate on which values to use.

1. Introduction

1.1. Background

When eliciting health state utility values for use in economic evaluation, both whose values should count (general public vs patients) and what they should be valuing (stylised health states vs own current health state) are debated (Brazier et al., 2018). These two issues are interrelated. Most valuation studies with a general public sample ask respondents to value several stylised health states (Devlin et al., 2018; Versteegh and Brouwer, 2016), but they can also be asked to value only their current health state (Burström et al., 2014). Similarly, patient samples can be asked to value their current (De Wit et al., 2000; Mann et al., 2009; Pickard et al., 2013; Ratcliffe et al., 2007) or stylised health

Abbreviations: DCE, discrete choice experiment; DCE_{TTO}, discrete choice experiment with duration; MO, mobility; SC, self-care; UA, usual activities; PD, pain or discomfort; AD, anxiety or depression; AIC, Akaike information criteria; BIC, Bayesian information criteria.

* Corresponding author. Suite 12, Floor 7, Williamson Building, Oxford Road, Manchester, M13 9PL, UK.

E-mail addresses: jack.elliott@manchester.ac.uk (J. Elliott), a.tsuchiya@sheffield.ac.uk (A. Tsuchiya).

https://doi.org/10.1016/j.socscimed.2022.115474
Received 11 May 2022; Received in revised form 25 September 2022; Accepted 21 October 2022
Available online 25 October 2022

states (Krabbe et al., 2011; Ludwig et al., 2021; Wang et al., 2014). We will primarily focus on whether value sets used for economic evaluation in health care should be informed by the general public valuing stylised health states (which we refer to as hypothetical valuation) or valuing their current health state (or, experience-based valuation).

Hypothetical valuation is the approach recommended by many national agencies in the context of priority setting within publicly financed health care systems (e.g., Australia, Canada, France, Netherlands, and England; Kennedy-Martin et al., 2020; NICE, 2022; Versteegh et al., 2016) and is usually justified based on the belief that everyone should have a say in how the health care resources they finance via taxation are distributed (Helgesson et al., 2020). However, this approach is contested (Brazier et al., 2005; Wolff et al., 2012): it is also argued that much of the general public is not able to properly comprehend living in impaired health and, as such, utility values should be obtained from those knowledgeable about the health states being valued in the form of experience-based valuation. The Dental and Pharmaceutical Benefits agency in Sweden are proponents of this view, recommending that appraisals of classification systems be made by "persons in the health condition in question" (The Dental and Pharmaceutical Benefits Agency, 2003). Some also argue for a combined approach (Versteegh and Brouwer, 2016).

The debate on who should be valuing what is particularly pertinent given that experienced-based valuation typically assigns more utility to the same health states than hypothetical valuation does (Burström et al., 2014; De Wit et al., 2000; Peeters and Stiggelbout, 2010; Ratcliffe et al., 2007). However, this pattern is not universal, and the extent and direction of differences can depend on: the health dimension and/or severity of the impairment being valued (Insinga and Fryback, 2003; Mann et al., 2009; Rand-Hendriksen et al., 2012), the population from where experienced-based samples are recruited (Heijink et al., 2017; Mann et al., 2009), and the technique used to elicit valuations (Krabbe et al., 2011).

There are several mutually non-exclusive explanations for why experience-based values differ from hypothetical values, and many of them concern the health experience of the respondent. Some explanations refer to those with certain health experiences having different perceptions and understanding of health problems that make them interpret the same health state description differently (i.e., an information-based mechanism). For example, those with depression may better understand the nuances of a depressive state description compared to those who have never lived with depression. Other explanations refer to those with experience having different opinions and evaluations of the health states even if they were to have the same information on the health state descriptor as those without any experience (i.e., a preference-based mechanism). For example, those with experience may conduct valuations according to a different internal standard of severity, they may consider contrasting effects (i.e., impaired health making other frustrations feel less detrimental), and/or they may have a vested interest in the subsequent allocation decisions (Brazier et al., 2005; Helgesson et al., 2020; Ubel et al., 2003).

Normatively speaking, it may seem uncontentious to conduct an experience-based valuation, asking people about what they know. On the other hand, a normative implication of experience-based value sets in the presence of the preference mechanism is that it amounts to privileging the preferences of some over others. For example, an experience-based value set excludes the preferences of those with no health problems. This selectiveness requires a stronger justification than would be the case if only the information mechanism is present. Indeed, if the preference mechanism is found, the use of experience-based valuation to estimate a value set would require the same stronger justification, even if the motivation for their use was only those with health experience being better informed.

1.2. Study aims

Against this background, the question of interest is: do experience-based values differ only because of *an information mechanism,* or is there a *preference mechanism* (either on its own or alongside the information mechanism)? Since these mechanisms are not observable, this question needs to be answered indirectly.

Regarding the information mechanism, it is quite plausible that those with direct experience of health problems are better informed about those specific health problems. However, there is no reason to suppose that such people are better informed about other health problems that they have no direct experience of. Given this, if the information mechanism alone is to hold, we postulate that differences in valuation will be contained within the relevant dimensions of health. For example, those with depression may value the dimensions affected by depression differently from those without any problems but not the other dimensions. However, note that this is only testable in a hypothetical valuation study because in an experience-based valuation study respondents are not asked to value health states they are not in.

However, if the discrepancy between hypothetical and experience-based valuations is due to the preference mechanism, we postulate that the discrepancy will not necessarily be restricted to the relevant dimensions of health. Compared to those with no health problems, those with depression may have different utility decrements in any dimension of health if, for example, having depression made them appreciate being pain-free more. As before, empirically testing for this requires a hypothetical valuation study.

As we see below (Section 1.3), the empirical literature has more often examined the difference in observed preferences at the level of health *states* rather than at the level of health *dimensions*. We will examine people's observed preferences for health dimensions. An effect will be termed "within-dimensional" if respondents with serious problems in a given dimension have different observed preferences from those without serious problems in the said dimension. An effect will be termed "cross-dimensional" if respondents with serious problems in a given dimension have different observed preferences in any other dimension. Observing significant within-dimensional effects cannot distinguish between the information and preference mechanisms. We postulate, however, that significant cross-dimensional effects can only be explained by the preference mechanism.

Thus, the objective of this study is to explore the possibility of the preference mechanism in health state valuation by way of cross-dimensional effects of having a serious health problem. Since most people do not have serious health problems, the ideal study would over-sample respondents with serious health problems and collect information on the nature and duration of these health problems alongside the health state valuation exercise. In the absence of such an ideal study, our analysis is carried out using a secondary health state valuation dataset with a general public sample that has only limited information on respondents' own health.

1.3. Existing empirical literature

There are two studies that explore the effect of self-reported health on the dimension level by examining whether the severity of a self-reported problem in a particular health dimension affects preferences for the same (within) dimension.

Jonker et al. (2017) use a discrete choice experiment (incorporating a duration attribute [DCE_{TTO}]) to examine, for each EQ-5D-5L dimension, how latent health state utility values are affected by respondents having a level of health lower than or equal to the health states being valued. Further detail on the DCE_{TTO} and EQ-5D-5L is provided in Section 2.1. They find that, for every dimension of health, having a lower or equal level of health to the health state being valued results in smaller utility decrements. Thus, the study captures the within-dimensional effect of self-reporting a problem in a particular EQ-5D dimension on observed preferences for the same dimension but does so *relative to* the stylised health profile. Furthermore, they do not capture the effect *across* health dimensions.

Ogorevc et al. (2019) use time trade-off to examine the average effect of belonging to a metastatic breast cancer or rheumatoid arthritis patient group on valuations for each EQ-5D-5L dimension (and overall health state values). They test how the valued health state having a worse level than the patient's self-reported level in a given dimension affects the health state value and find heterogeneous effects by dimension. Thus, like Jonker et al. (2017), the study explores the effect *within* health dimensions but not *across*. We are not aware of any studies that explicitly examine cross-dimensional effects.

2. Methods

2.1. Data

This study uses data from the "Further Exploration of Discrete Choice Experiments with Duration (DCE_{TTO}) for EQ-5D-5L Valuation" (FEDEV) project (Mulhern et al., 2017), DCE design Type Ia. DCE_{TTO} is an extension of the traditional DCE, which can be used to elicit preferences between dimensions of health on a latent scale, but by including duration of survival as a dimension, the latent preferences are anchored onto the utility-scale where 1 is equivalent to full health and 0 death (Bansback et al., 2012; Mulhern et al., 2019).

Design Type Ia was the reference DCE variant in FEDEV against which a number of methodological variants were compared. Respondents were asked to choose between two health state profiles describing the respondent living in a particular health state for one of six specified durations (6 months, 1, 2, 4, 5, and 10 years) and then dying.

Table 1
The EQ-5D-5L instrument.

	Mobility (MO)
MO1	I have *no* problems in walking about
MO2	I have *slight* problems in walking about
MO3	I have *moderate* problems in walking about
MO4	I have *severe* problems in walking about
MO5	I am *unable* to walk about
	Self-care (SC)
SC1	I have *no* problems in washing or dressing myself
SC2	I have *slight* problems in washing or dressing myself
SC3	I have *moderate* problems in washing or dressing myself
SC4	I have *severe* problems in washing or dressing myself
SC5	I am *unable* to wash or dress myself
	Usual activities (UA)
UA1	I have *no* problems doing my usual activities
UA2	I have *slight* problems doing my usual activities
UA3	I have *moderate* problems doing my usual activities
UA4	I have *severe* problems doing my usual activities
UA5	I am *unable* to do my usual activities
	Pain or discomfort (PD)
PD1	I have *no* pain or discomfort
PD2	I have *slight* pain or discomfort
PD3	I have *moderate* pain or discomfort
PD4	I have *severe* pain or discomfort
PD5	I have *extreme* pain or discomfort
	Anxiety or depression (AD)
AD1	I am *not* anxious or depressed
AD2	I am *slightly* anxious or depressed
AD3	I am *moderately* anxious or depressed
AD4	I am *severely* anxious or depressed
AD5	I am *extremely* anxious or depressed

Adapted from Herdman et al. (2011).

The health states are defined by the EQ-5D-5L instrument, which covers the dimensions of mobility (MO), self-care (SC), usual activities (UA), pain or discomfort (PD), and anxiety or depression (AD) (Herdman et al., 2011). Each dimension is subdivided into five levels of severity (Table 1).

The experimental design of the DCE_{TTO} is based on D-efficiency and is composed of 120 pairs of EQ-5D-5L health profiles across 12 blocks of ten choices. The estimation of the baseline model with 21 parameters (see Section 2.2.4 for details) does not require 120 pairs: the minimum number of pairs required is the number of parameters being estimated. The design is deliberately overgenerous for estimating the 21-parameter baseline model to allow for the exploration of quadratic duration and to relax the assumption of linear utility of duration, which would require 42 parameters to estimate.

The FEDEV sample was recruited from a commercial internet panel hosted by IPSOS Observer, with age and gender quotas representing the UK general public aged 18–65. Background characteristics included respondents' own health in EQ-5D-5L. The Type Ia dataset we use includes 802 individuals.

2.2. Analysis

2.2.1. Defining a serious health problem

To examine the effect of having a serious health problem on health state utility values, we define health problems according to respondents' self-reported health in the EQ-5D-5L dimensions for PD or AD. Health problems are deemed "serious" if the respondent self-reports either a level 4 or 5 problem. PD and AD were chosen because, as we see below (Section 3.1), more respondents reported having level 4 or 5 problems in these dimensions than in the other three dimensions of EQ-5D-5L, increasing the probability of identifying the regression parameters (see Section 2.2.5). In what follows, the serious health problem is represented by $s =$ PD, AD, while having or not having a given serious health problem is represented by $h_s = 1, 0$. These provide four subsamples.

2.2.2. Selecting the analysis sample

Those respondents who provide self-reported PD and AD data alongside the DCE_{TTO} choice tasks comprise *the valid sample*. To control for possible effects of the different background characteristics on preferences across the four subsamples (Al Shabasy et al., 2022; Pickard et al., 2013), *the analysis sample* for each s consists of those who report $h_s = 1$ alongside those who report $h_s = 0$ but otherwise have matched characteristics to those who do. To achieve this, for each s, we matched each respondent with $h_s = 1$ to the most similar respondent, or "nearest neighbour", with $h_s = 0$. We used the Mahalanobis distance as our measure of similarity (Mahalanobis, 1936) in terms of age, gender, whether married, and whether completed a higher degree. Where there was a tie for the most similar individual, multiple nearest neighbours were selected. While this matching approach reduces the risk that the analysis is confounded by the different socioeconomic covariates depending on h_s, it does not control for the possibility that those with $h_s = 1$ in a given s are more likely to also have problems in other dimensions. This matched sample is the main focus of our analysis, but we also present results of the same analyses using the larger valid sample in the supplementary appendices.

2.2.3. Null hypotheses

There are two independent null hypotheses. The first concerns within-dimensional effects: that self-reporting $h_s = 1$ in either s has no effect on people's observed preferences for the same dimension. Rejecting this null hypothesis does not distinguish between the information mechanism, where respondents who self-report $h_s = 1$ in a given s have a different (and most likely better) understanding of this dimension than those who do not; and the preference mechanism, where people have different opinions of health states depending on their experience of health. The second and more important null hypothesis concerns cross-dimensional effects: that self-reporting $h_s = 1$ in a given s has no effect on people's observed preferences for the other dimensions. Rejecting this null hypothesis on cross-dimensional effects will be interpreted as evidence of the preference mechanism.

2.2.4. Baseline model (model 0)

The DCE_{TTO} data are modelled using a conditional logit regression (Bansback et al., 2012). The utility μ for each individual i, for profile j, is defined as:

$$\mu_{0ij} = \beta_{01} t_{ij} + \beta_{02}^{'} \boldsymbol{x}_{ij} t_{ij} + \varepsilon_{0ij} \quad (0)$$

where t_{ij} is the number of life years in health state profile j, and $\boldsymbol{x}_{ij}$ is the set of dummy variables representing the EQ-5D-5L dimensions in profile j (with no problems as the reference category). The estimate of β_{01} is the latent preference associated with an individual living for an additional life year in full health and is expected to be positive. The set of $\beta_{02}^{'}$ estimates are associated with living for one year with the set of health problems indicated by $\boldsymbol{x}_j$ and is expected to be non-positive. The specification assumes linear utility in time and constant proportional time trade-off and has 21 parameters to estimate. The error term ε_{0ij} is a random term, which is assumed to be Type 1 extreme value distributed. We clustered standard errors at the individual level to account for potential serial correlation between respondent choices.

These coefficients are on a latent scale and are not meaningful on their own. We anchor them on a meaningful scale with 1 for full health and 0 for being dead to facilitate the comparison of the coefficients across models. The 20 $\widehat{\beta}_{02}^{'}$ coefficients of Model 0 divided with $\widehat{\beta}_{01}$ gives the anchored parameters for $\boldsymbol{x}_j$, made up of four parameters for each of the EQ-5D dimensions, with level 1 (no problems) as the baseline (Bansback et al., 2012).

2.2.5. The effect of respondent health

The effect of self-reporting $h_s = 1$ in a given s can be discerned by adapting Model 0 to include interactions with h_s. This approach allows for direct formal post-estimation tests to compare across the subsamples. Note that anchored coefficients are comparable across respondents who may have different latent scales (Vass et al., 2018) because they are standardised onto a common scale with 1 for full health and 0 for being dead.

Using Model 1, we formally test the effect of whether respondent i has a serious health problem by adding interactions with h_{si}:

$$\mu_{1ij} = \beta_{11} t_{ij} + \beta_{12}^{'} \boldsymbol{x}_{ij} t_{ij} + \beta_{13} t_{ij} h_{si} + \beta_{14}^{'} \boldsymbol{x}_{ij} t_{ij} h_{si} + \varepsilon_{1ij} \quad (1)$$

Model 1 for s = PD and Model 1 for s = AD each has 42 parameters to estimate.

The use of the interaction term for self-reported serious health problems means that Model 1 has two sets of coefficients, and these are anchored differently. Where $h_s = 0$, the anchored average utility decrements are obtained in a similar process to Model 0, by dividing the $\widehat{\beta}_{12}^{'}$ coefficients with $\widehat{\beta}_{11}$. Where $h_s = 1$, they are obtained by dividing $[\widehat{\beta}_{12}^{'} + \widehat{\beta}_{14}^{'}]$ with $[\widehat{\beta}_{11} + \widehat{\beta}_{13}]$. Thus, Model 1 for s = PD has 40 anchored parameters: 20 for $\boldsymbol{x}_j$ given $h_{si} = 0$ and 20 for $\boldsymbol{x}_j$ given $h_{si} = 1$. Similarly, Model 1 for s = AD has another two sets of 20 anchored parameters. For illustration, the predicted value of the worst state (55555) is reported for each model by h.

Using the anchored parameters from Model 1, we test the null hypothesis that having $h_s = 1$ has no effect on the utility decrements associated with living for one year with health problems:

$$H_0: \left[\frac{\widehat{\beta}_{12}^{'} + \widehat{\beta}_{14}^{'}}{\widehat{\beta}_{11} + \widehat{\beta}_{13}} - \frac{\widehat{\beta}_{12}^{'}}{\widehat{\beta}_{11}}\right] \boldsymbol{x}_j = 0. \quad \text{(Test 1)}$$

The test is carried out by taking the difference in average utility decrements across respondent groups with $h_s = 1$ and $h_s = 0$. This involves 20 test statistics for each EQ-5D-5L dimension in both the PD and AD models.

We will also compare Akaike and Bayesian information criteria (AIC and BIC, respectively).

2.2.6. Additional analyses

There are three additional analyses, the details of which are reported in Supplementary Appendices 1-3, respectively.

First, as we will describe below, the four subsamples are small, and the observations are distributed unevenly across the 12 survey blocks. This could feasibly affect whether the choice design is able to identify the parameters of interest. To test this, we simulate choice data for each of the four subsamples, representing the relevant number of respondents per block, and estimate Model 0.

Second, Test 1 above is based on absolute differences observed on the anchored parameters by h_s. However, the distinction between within-dimension effects and cross-dimension effects may be confounded if the range of the value sets differ (in other words, if the predicted values for state 55555 are different) depending on h_s. For example, if $h_s = 1$ in a given s made the range of health state values narrower, this would reduce the anchored absolute decrements for all other dimensions, which would appear as if there was a cross-dimensional effect even if their relative importance remained constant. In order to explore this, the anchored parameters are normalised as relative attribute importance scores (Mott et al., 2021) and their ratios by h_s are compared.

Third, Model 0 (and by extension, Model 1) assumes that respondents trade-off between quality and quantity of life at a constant rate regardless of duration, and therefore the anchored coefficients are independent of duration. However, given that constant proportional time trade-off has previously been rejected in EQ-5D-5L DCE_{TTO} studies (Jakubczyk et al., 2018), we explore alternative specifications that relax this assumption to check our main findings are not affected.

Table 2
Number of respondents with and without serious PD and serious AD health problems for the valid sample and the analysis samples.

Analysis sample for PD			
		$\boldsymbol{s}$ **= AD**	
$\boldsymbol{s}$ **= PD**	$h_s = 0$	$h_s = 1$	Total
$h_s = 0$	88 (91.7%)	8 (8.3%)	96 (100.0%)
$h_s = 1$	38 (71.7%)	15 (28.3%)	53 (100.0%)
Total	126 (84.6%)	23 (15.4%)	149 (100.0%)
Analysis sample for AD			
		$\boldsymbol{s}$ **= PD**	
$\boldsymbol{s}$ **= AD**	$h_s = 0$	$h_s = 1$	Total
$h_s = 0$	123 (94.6%)	7 (5.4%)	130 (100.0%)
$h_s = 1$	44 (74.6%)	15 (25.4%)	59 (100.0%)
Total	167 (88.4%)	22 (11.6%)	189 (100.0%)
Valid sample			
		$\boldsymbol{s}$ **= AD**	
$\boldsymbol{s}$ **= PD**	$h_s = 0$	$h_s = 1$	Total
$h_s = 0$	670 (93.8%)	44 (6.2%)	714 (100.0%)
$h_s = 1$	38 (71.7%)	15 (28.3%)	53 (100.0%)
Total	708 (92.3%)	59 (7.7%)	767 (100.0%)

Notes: $\boldsymbol{s}$ **= AD:** serious anxiety or depression; $\boldsymbol{s}$ **= PD:** serious pain or discomfort; $\boldsymbol{h_s}$ **= 0:** DCE respondents without serious health problem, s; $\boldsymbol{h_s}$ **= 1:** DCE respondents with serious health problem, s; **Total**: DCE respondents with and without serious health problem, s. The **percentages** represent the relative frequency of each cell within its row. The **analysis sample for** $\boldsymbol{s}$ consists of those respondents who have $\boldsymbol{h_s}$ **= 1** and a matched sample of those who have $\boldsymbol{h_s}$ **= 0.** The **valid sample** consists of all respondents who have provided self-reported PD and AD alongside the DCE data.

3. Results

3.1. Descriptive statistics

Of the 802 respondents in the dataset, there is complete self-reported PD and AD data alongside their DCE choice data for 767 respondents, forming our valid sample. As Table 2 illustrates, the analysis sample for PD consisted of 53 respondents who self-reported $h_{PD} = 1$ and 96 matched respondents who did not, totalling 149. The respective amounts for the AD analysis sample were 59, 130, and 189. Table A1 (Supplementary Appendix 1) shows that across the 12 survey blocks, the distribution of individuals with $h_s = 1$ in the analysis sample is imbalanced with some blocks having only one or two respondents.

Fig. 1 shows the distribution of self-reported EQ-5D-5L dimensions for the full sample (n = 802), the valid sample (n = 767), and the two analysis samples (n = 149 for PD; n = 189 for AD).

The number of respondents reporting severe or extreme health problems in the full sample in each dimension was higher for PD (53 [6.9%]) and AD (59 [7.7%]) than in MO (33 [4.3%]), SC (7 [0.9%]) or UA (38 [4.9%]). In the valid sample for each of the EQ-5D dimensions, the majority of respondents self-reported no problems. However, because small proportions of individuals with $h_{PD} = 0$ and $h_{AD} = 0$ were matched and included, a relatively smaller proportion of individuals self-reported no problems in the analysis samples, particularly for the PD and AD dimensions.

Table 3 describes the characteristics of the valid and analysis samples by h_s. We also provide the p-values for a two-sided exact Fisher test examining whether each sociodemographic characteristic is independent of whether the respondent reports $h_s = 1$.

As intended, the two groups with $h_s = 1$ and $h_s = 0$ become demographically more similar to each other in the analysis samples compared to the valid sample. For both s = PD and s = AD, age, attainment of a degree, and carer illness experience are no longer associated with h_s after matching. For s = PD, marital status is no longer associated with h_s after matching.

3.2. The effect of respondent health

Derived from Model 0 and Model 1 in the analysis samples, Table 4 presents a selection of results using the anchored results (Table A7; Supplementary Appendix 4 reports the unanchored coefficients for Model 0 and Model 1) as well as AIC and BIC for Model 0 and Model 1. Columns 4 and 8 present the results of Test 1 derived from Model 1: the effect of having $h_s = 1$ on the average utility decrements. Columns 2, 3, 6, and 7 present the inputs to Test 1: the average anchored utility decrements for $h_s = 1$ (columns 2 and 6) and for $h_s = 0$ (columns 3 and 7). Columns 1 and 5 give the anchored utility decrements for Model 0 for comparison. The predicted utility value for the worst health state (55555) is also presented in the bottom row of the table for all individuals (columns 1 and 5), those with $h_s = 1$ (columns 2 and 6), and those $h_s = 0$ (columns 3 and 7).

As an example, Test 1 indicates that for those with $h_{PD} = 1$, the utility of MO2 is lower by 0.116 on the anchored scale compared to those with $h_{PD} = 0$. This corresponds to the difference between −0.107 (the utility associated with living for a year with MO2 for respondents with $h_{PD} = 1$) and 0.009 (the utility associated with living for a year with MO2 for respondents with $h_{PD} = 1$). As can be seen, regarding $h_{PD} = 1$, the only significant effect is cross-dimensional and can be found on the anchored coefficient for SC2. The positive sign suggests that respondents self-reporting $h_{PD} = 1$ value an additional year in health states with SC2 higher by 0.209 than $h_{PD} = 0$.

For the effect of $h_{AD} = 1$, there are negative cross-dimensional effects suggesting that respondents self-reporting $h_{AD} = 1$ value an additional year in health states with MO2, PD2, PD3, PD4, or PD5 lower by 0.291, 0.388, 0.291, 0.343, and 0.256, respectively, compared to those who $h_{AD} = 1$. There is also a positive cross-dimensional effect of AD on SC2, suggesting that those with $h_{AD} = 1$ value living a year in SC2 more than those with $h_{AD} = 0$ by 0.227.

For both s = PD and s = AD, those with $h_s = 1$ assign higher utility to state 55555. This difference indicates those with $h_s = 1$ have a narrower range of disutility, lending weight to the further analysis using relative

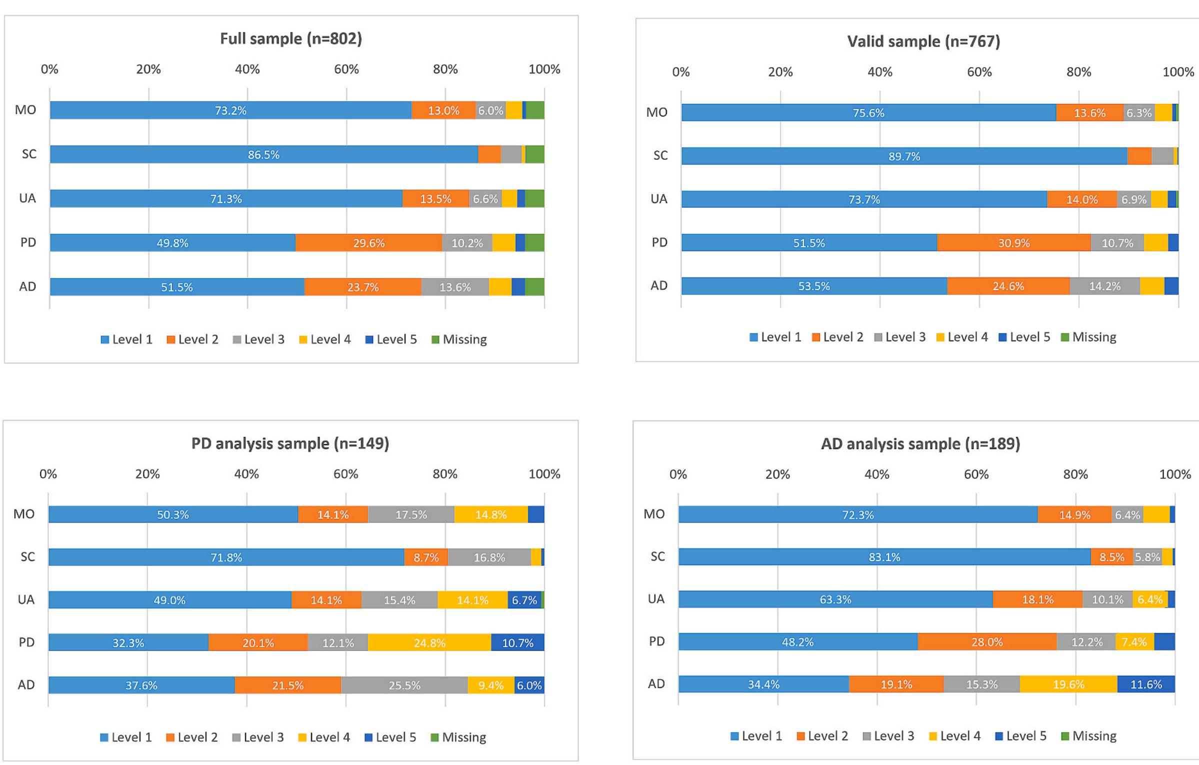

Fig. 1. Self-reported EQ-5D-5L by dimension for full sample, valid sample, and analysis samples. Notes: **MO**: Mobility; **SC:** Self care; **UA:** Usual activities; **PD:** Pain or discomfort; **AD**: Anxiety ordepression.

Table 3
Summary statistics for valid and analysis samples.

	$s = $ PD						$s = $ AD					
	Valid sample			Analysis sample			Valid sample			Analysis sample		
	$h_s = 0$	$h_s = 1$	P-value: assoc. with h_s	$h_s = 0$	$h_s = 1$	P-value: assoc. with h_s	$h_s = 0$	$h_s = 1$	P-value: assoc. with h_s	$h_s = 0$	$h_s = 1$	P-value: assoc. with h_s
Characteristics												
Age			0.004			0.761			0.001			0.695
18-25 [count]	111 (15.5%)	3 (5.7%)		7 (7.3%)	3 (5.7%)		98 (13.8%)	16 (27.1%)		35 (26.9%)	16 (27.1%)	
26-35 [count]	171 (23.9%)	7 (13.2%)		18 (18.8%)	7 (13.2%)		164 (23.2%)	14 (23.7%)		37 (28.5%)	14 (23.7%)	
36-45 [count]	163 (22.8%)	9 (17.0%)		16 (16.7%)	9 (17.0%)		162 (22.9%)	10 (16.9%)		21 (16.2%)	10 (16.9%)	
46-55 [count]	130 (18.2%)	19 (35.8%)		25 (26.0%)	19 (35.8%)		133 (18.8%)	16 (27.1%)		25 (19.2%)	16 (27.1%)	
56-65 [count]	139 (19.5%)	15 (28.3%)		30 (31.3%)	15 (28.3%)		151 (21.3%)	3 (5.1%)		12 (9.2%)	3 (5.1%)	
Age [range, years]	18–65	20–65	–	20–65	20–65	–	18–65	18–65	–	18–65	18–65	–
Employment status			0.000			0.000			0.000			0.000
In work [count]	484 (67.9%)	13 (24.5%)		65 (67.7%)	13 (24.5%)		473 (66.9%)	24 (40.7%)		84 (64.6%)	24 (40.7%)	
Not in work [count]	200 (28.1%)	39 (73.6%)		29 (30.2%)	39 (73.6%)		207 (29.3%)	32 (54.2%)		32 (24.6%)	32 (54.2%)	
Student [count]	29 (4.1%)	1 (1.9%)		2 (2.1%)	1 (1.9%)		27 (3.8%)	3 (5.1%)		14 (10.8%)	3 (5.1%)	
Illness experience												
Self [count]	167 (23.4%)	48 (92.3%)	0.000	30 (31.3%)	48 (92.3%)	0.000	173 (24.5%)	42 (71.2%)	0.000	30 (23.1%)	42 (71.2%)	0.000
Family [count]	437 (61.5%)	45 (84.9%)	0.001	64 (67.4%)	45 (84.9%)	0.021	437 (62.0%)	45 (76.3%)	0.034	81 (62.3%)	45 (76.3%)	0.068
From caring [count]	143 (20.2%)	19 (35.8%)	0.014	29 (30.2%)	19 (35.8%)	0.583	146 (20.8%)	16 (27.1%)	0.250	23 (18.0%)	16 (27.1%)	0.177
Male [count]	250 (49.0%)	23 (43.4%)	0.478	15 (28.3%)	23 (43.4%)	1.000	344 (48.6%)	29 (49.2%)	1.000	63 (48.5%)	29 (49.2%)	1.000
Married [count]	421 (59.0%)	27 (50.9%)	0.312	53 (55.2%)	27 (50.9%)	0.732	425 (60.0%)	23 (39.0%)	0.002	59 (45.4%)	23 (39.0%)	0.433
Has degree [count]	415 (58.5%)	23 (44.2%)	0.058	50 (52.6%)	23 (44.2%)	0.389	415 (58.9%)	23 (39.7%)	0.005	55 (42.6%)	23 (39.7%)	0.750
Children [count]	161 (22.6%)	8 (15.1%)	0.233	8 (8.3%)	8 (15.1%)	0.269	156 (22.1%)	13 (22.0%)	1.000	35 (26.9%)	13 (22.0%)	0.589
EQ-5D 11111 [count]	248 (34.9%)	0 (0.0%)	0.000	24 (25.3%)	0 (0.0%)	0.000	248 (35.2%)	0 (0.0%)	0.000	42 (32.6%)	0 (0.0%)	0.000
Total	714	53	–	96	53	–	708	59	–	130	59	–

Notes: $s = $ **AD:** serious anxiety or depression; $s = $ **PD:** serious pain or discomfort; $h_s = 0$**:** DCE respondents without serious health problem, s; $h_s = 1$**:** DCE respondents with serious health problem, s; **Total**: DCE respondents with and without serious health problem, s. The **analysis sample for** s consists of those respondents who have $h_s = 1$ and a matched sample of those who have $h_s = 0$. The **valid sample** consists of all respondents who have provided self-reported PD and AD alongside the DCE data. Each **percentage** represents the frequency relative to the total number of respondents that provided information on the given characteristic. The **P-values** are from a two-sided exact Fisher test, testing the null hypothesis that the respective characteristics are independent of whether the respondent has a serious health problem.

Table 4
Anchored effects of having serious PD or serious AD on health state preference (analysis samples).

	$s = $ PD				$s = $ AD			
	1) Model 0	**2) Model 1, $h_s = 1$**	**3) Model 1, $h_s = 0$**	**4) Test 1: effect of h_s**	**5) Model 0**	**6) Model 1, $h_s = 1$**	**7) Model 1, $h_s = 0$**	**8) Test 1: effect of h_s**
MO2	−0.025	−0.107	0.009	−0.116	−0.014	−0.235*	0.056	−0.291*
	(−0.154, 0.103)	(−0.301, 0.088)	(−0.140, 0.158)	(−0.361, 0.129)	(−0.175, 0.147)	(−0.471, 0.001)	(−0.130, 0.242)	(−0.592, 0.010)
MO3	−0.101	−0.124	−0.109	−0.015	0.023	0.027	0.026	0.001
	(−0.238, 0.037)	(−0.333, 0.085)	(−0.292, 0.074)	(−0.293, 0.262)	(−0.159, 0.205)	(−0.263, 0.318)	(−0.183, 0.235)	(−0.357, 0.359)
MO4	−0.341***	−0.321***	−0.356***	0.035	−0.276***	−0.339***	−0.256***	−0.083
	(−0.436, −0.246)	(−0.465, −0.177)	(−0.484, −0.229)	(−0.157, 0.228)	(−0.388, −0.165)	(−0.510, −0.169)	(−0.387, −0.125)	(−0.298, 0.132)
MO5	−0.297***	−0.314***	−0.296***	−0.019	−0.253***	−0.217**	−0.243***	0.025
	(−0.406, −0.188)	(−0.487, −0.142)	(−0.434, −0.157)	(−0.240, 0.202)	(−0.365, −0.141)	(−0.421, −0.013)	(−0.370, −0.115)	(−0.215, 0.266)
SC2	−0.044	0.090	−0.119	0.209*	−0.076	0.094	−0.132*	0.227*
	(−0.158, 0.069)	(−0.080, 0.260)	(−0.265, 0.027)	(−0.015, 0.433)	(−0.209, 0.057)	(−0.108, 0.296)	(−0.289, 0.024)	(−0.029, 0.482)
SC3	−0.134**	−0.119	−0.142*	0.023	−0.112*	−0.121	−0.087	−0.034
	(−0.257, −0.010)	(−0.300, 0.063)	(−0.295, 0.012)	(−0.215, 0.260)	(−0.244, 0.020)	(−0.328, 0.086)	(−0.221, 0.046)	(−0.280, 0.212)
SC4	−0.235***	−0.125	−0.321***	0.196	−0.246***	−0.177	−0.279***	0.102
	(−0.378, −0.092)	(−0.318, 0.068)	(−0.491, −0.150)	(−0.062, 0.453)	(−0.391, −0.101)	(−0.395, 0.041)	(−0.438, −0.120)	(−0.168, 0.371)
SC5	−0.376***	−0.333***	−0.396***	0.063	−0.395***	−0.312**	−0.400***	0.088
	(−0.491, −0.260)	(−0.515, −0.152)	(−0.539, −0.254)	(−0.168, 0.294)	(−0.548, −0.241)	(−0.601, −0.023)	(−0.566, −0.234)	(−0.246, 0.421)
UA2	−0.138**	−0.051	−0.202**	0.151	−0.012	−0.033	−0.014	−0.019
	(−0.250, −0.026)	(−0.202, 0.101)	(−0.360, −0.044)	(−0.067, 0.370)	(−0.144, 0.120)	(−0.266, 0.200)	(−0.171, 0.142)	(−0.300, 0.262)
UA3	−0.195***	−0.176**	−0.215**	0.040	−0.077	0.072	−0.140*	0.212
	(−0.328, −0.062)	(−0.342, −0.009)	(−0.404, −0.026)	(−0.212, 0.291)	(−0.215, 0.061)	(−0.201, 0.346)	(−0.288, 0.008)	(−0.099, 0.524)
UA4	−0.267***	−0.214**	−0.314***	0.099	−0.149**	−0.074	−0.186**	0.112
	(−0.401, −0.133)	(−0.387, −0.042)	(−0.498, −0.130)	(−0.153, 0.352)	(−0.278, −0.020)	(−0.295, 0.147)	(−0.337, −0.035)	(−0.156, 0.380)
UA5	−0.380***	−0.275***	−0.428***	0.152	−0.311***	−0.247**	−0.337***	0.090
	(−0.510, −0.250)	(−0.454, −0.097)	(−0.604, −0.252)	(−0.098, 0.403)	(−0.442, −0.179)	(−0.486, −0.007)	(−0.492, −0.182)	(−0.195, 0.376)
PD2	−0.066	−0.038	−0.036	−0.002	−0.086	−0.389***	−0.001	−0.388**
	(−0.212, 0.080)	(−0.261, 0.185)	(−0.214, 0.142)	(−0.287, 0.283)	(−0.247, 0.075)	(−0.661, −0.116)	(−0.189, 0.187)	(−0.719, −0.057)
PD3	−0.077	−0.073	−0.057	−0.016	−0.089	−0.319***	−0.028	−0.291*
	(−0.238, 0.084)	(−0.262, 0.115)	(−0.282, 0.167)	(−0.310, 0.278)	(−0.255, 0.078)	(−0.542, −0.097)	(−0.231, 0.175)	(−0.593, 0.010)
PD4	−0.246***	−0.124	−0.288***	0.165	−0.288***	−0.561***	−0.218***	−0.343**
	(−0.380, −0.111)	(−0.336, 0.089)	(−0.452, −0.125)	(−0.103, 0.433)	(−0.424, −0.151)	(−0.850, −0.271)	(−0.368, −0.068)	(−0.669, −0.017)
PD5	−0.338***	−0.228**	−0.378***	0.150	−0.441***	−0.640***	−0.384***	−0.256*
	(−0.471, −0.206)	(−0.412, −0.044)	(−0.544, −0.212)	(−0.098, 0.398)	(−0.590, −0.292)	(−0.871, −0.409)	(−0.557, −0.211)	(−0.545, 0.033)
AD2	−0.078	−0.154	−0.060	−0.094	−0.086	−0.048	−0.109	0.061
	(−0.201, 0.045)	(−0.368, 0.059)	(−0.218, 0.097)	(−0.359, 0.171)	(−0.216, 0.044)	(−0.271, 0.175)	(−0.252, 0.033)	(−0.203, 0.326)
AD3	−0.051	−0.084	−0.055	−0.028	−0.139*	−0.080	−0.180**	0.099
	(−0.195, 0.093)	(−0.346, 0.178)	(−0.221, 0.110)	(−0.338, 0.281)	(−0.294, 0.017)	(−0.369, 0.209)	(−0.331, −0.028)	(−0.227, 0.426)
AD4	−0.283***	−0.369***	−0.238***	−0.131	−0.269***	−0.098	−0.317***	0.218
	(−0.406, −0.161)	(−0.557, −0.182)	(−0.400, −0.075)	(−0.380, 0.117)	(−0.403, −0.135)	(−0.345, 0.149)	(−0.457, −0.176)	(−0.066, 0.503)
AD5	−0.366***	−0.393***	−0.352***	−0.041	−0.421***	−0.275**	−0.474***	0.199
	(−0.519, −0.212)	(−0.629, −0.157)	(−0.535, −0.169)	(−0.340, 0.258)	(−0.576, −0.266)	(−0.541, −0.008)	(−0.644, −0.303)	(−0.117, 0.515)
Obs	1490	1490	1490	1490	1890	1890	1890	1890
AIC	1815.466	1830.625		–	2379.021	2393.34		–
BIC	1941.459	2082.611		–	2510.008	2655.314		–
55555	−0.757	−0.544	−0.850	–	−0.820	−0.691	−0.838	–
	(−1.023, −0.490)	(−0.890, −0.198)	(−1.191, −0.508)	–	(−1.138, −0.503)	(−1.184, −0.198)	(−1.213, −0.462)	–

Notes: MO: Mobility; SC: Self-care; UA: Usual activities; PD: Pain or discomfort; AD: Anxiety or depression; $\boldsymbol{s} = $ **AD:** serious anxiety or depression; $\boldsymbol{s} = $ **PD:** serious pain or discomfort; $\boldsymbol{h_s} = \mathbf{0}$: DCE respondents without serious health problem, s; $\boldsymbol{h_s} = \mathbf{1}$: DCE respondents with serious health problem, s; **Obs**: the number of DCE choice tasks across respondents; **AIC:** Akaike information criterion; **BIC**: Bayesian information criterion; **55555**: predicted utility value of the worst health state. 95% confidence intervals in parentheses. Stars indicate significance as follows: ***p < 0.01, **p < 0.05, *p < 0.1. Columns 4) and 8) represent the difference in respondents' average utility decrements where $h_s = 1$ $\left(\frac{\widehat{\beta}'_{12}\boldsymbol{x}_{ij} + \widehat{\beta}'_{14}\boldsymbol{x}_{ij}}{\widehat{\beta}_{11} + \widehat{\beta}_{13}}\right)$ and $h_s = 0$ $\left(\frac{\widehat{\beta}'_{12}\boldsymbol{x}_{ij}}{\widehat{\beta}_{11}}\right)$.

attribute importance scores.

For both $s = \text{PD}$ and $s = \text{AD}$, Akaike information criterion (AIC) and Bayesian information criterion (BIC) agree that Model 1 with health problem interactions is not an improvement over Model 0 in terms of model fit. Given that Model 1 does not improve the overall model fit, our results should be treated with caution.

Table A8 (Supplementary Appendix 5) is for the non-matched, valid sample and presents the anchored results of Test 1. We find that the negative cross-dimensional effects of $h_{AD} = 1$ for every PD severity level are robust. We also find that having $h_{PD} = 1$ has a positive within-dimensional effect on PD5. Those with $h_{PD} = 1$ assign more utility to state 55555. Aside from the BIC on the PD interaction, AIC and BIC both agree that Model 1 is not an improvement over Model 0 in terms of model fit.

3.3. Additional analyses

First, Supplementary Appendix 1 outlines the results of the simulations. It finds that the choice design with 120 choice sets across 12 blocks is sufficient to identify the 21 parameters of Model 0 even when the sample is as small as the analysis sample and is distributed unevenly across the blocks. While we do not consider this dataset to be large enough to base definitive claims about all cross-dimensional relationships or their magnitude, the pattern of the effect of having h_{AD} on preferences for PD severity levels that we observe is more likely to be an actual signal rather than chance random noise.

Second, Table A2 (Supplementary Appendix 2) reports the relative attribute importance scores of the worst utility decrement of each dimension, normalised using SC5 (which had the most stable anchored parameter across all subsamples), by subsample, and the ratios between them by h. The results are noisy, and it is difficult to draw any definitive conclusions, but overall, do not contradict our main findings. Further detail is provided in Supplementary Appendix 2.

Third, Table A3, A4, A5, and A6 (Supplementary Appendix 3) report the effect of $h_s = 1$ after relaxing the constant proportional time trade-off assumption. The main results above are unaffected by the limited range of non-constant specifications we examined. Further detail is provided in Supplementary Appendix 3.

4. Discussion

4.1. Main findings

The paper proposes two mechanisms to understand the role of experience of serious health problems on health state valuation. One of the two mechanisms concern information (i.e., how the health state in question is perceived or understood), while the other concerns preferences (i.e., how the health state in question is liked or appreciated). The presence of the preference mechanism has implications for the normative debate on what valuations should be used to estimate value sets for health care priority setting, and we propose to detect it by examining the effect of having a serious problem in one dimension of health on another.

This is the first empirical study to examine such a cross-dimensional effect in health state valuation. The few previous studies that have looked at the effect of impaired health on valuations by health state dimension have done so either without making the explicit distinction between cross-dimensional and within-dimensional effects (Ogorevc et al., 2019) or examined the effect for the *same* dimension (Jonker et al., 2017; Ogorevc et al., 2019).

The main analysis found no evidence of within-dimensional effects associated with self-reporting serious PD and limited evidence of cross-dimensional effects associated with self-reporting serious PD. It also found that self-reporting serious AD has no within-dimensional effect but has a negative cross-dimensional effect so that those reporting serious AD assign less utility to any PD problem. Furthermore, the magnitude of these cross-dimensional effects (0.256–0.388 across the four PD levels) is substantial. For example, these are larger than the largest anchored gap between adjacent levels in Model 0, which is 0.199, between PD levels 3 and 4. A similar pattern is observed consistently in the non-matched valid sample.

Three sets of further analyses were conducted: a simulation study to explore the small and imbalanced sample across the 12 blocks of choice tasks and the ability of the design to identify the parameters of interest; a comparison of the ratios of relative attribute importance scores by whether respondents self-report a serious health problem to examine the possible confounding from the narrower range of health state values elicited from those with serious problems; and a set of alternative models to test the effect of relaxing the assumption of constant proportional time trade-off. The main finding that having serious AD has cross-dimensional effects on the preference for PD was not contradicted by any of these.

We postulate that this effect is caused by the preference mechanism: those with serious AD have different opinions on PD severity levels from the rest of the population. The negative cross-dimensional effect could feasibly be capturing that a higher proportion of individuals who report serious AD, relative to those who do not, also report serious PD (see Table 2). However, if this were driving the effect, we would expect it to be positive because the within-dimensional effect of reporting serious PD, where it exists (Test 1, valid sample; Test 2, Model 2.1 [see Supplementary Appendix 3]), is positive. Given this, the negative cross-dimensional effect clearly cannot be attributed to the information mechanism alone. An ideal analysis might control for possible comorbidity, for example, by matching the analysis sample not only by socioeconomic covariates but also by serious health problems in the other dimensions, but this would require a substantially larger dataset.

While we do not aim to defend a position in the normative debate over hypothetical valuation and experience-based valuation, the findings have implications for that debate. If our results are found to hold more widely, then justification is needed for the use of experience-based preferences that privilege the opinions of those with serious health problems over that of the rest of the population. It is not sufficient to argue that respondents with direct experience of serious health problems better understanding such states, even where this was the only motivation.

Since cross-dimensional effects, by definition, do not affect experience-based value sets, this finding may not appear directly relevant to that discussion. However, what matters here is not the presence of cross-dimensional effects but the presence of the preference mechanism, which could also be within-dimensional. Furthermore, if people have different opinions of health states depending on their health experience, this implies that their indifference curves over the dimensions of health and survival intersect. For example, experience-based valuation could allow a logically worse state (valued by those in that state) to be given a better experienced-utility value than a logically better state (valued by those in that state). This is fundamentally different from the issue of preference heterogeneity in hypothetical valuation, where the allocation of health states to value is randomised across respondents.

4.2. Limitations

This study repurposes data that were not collected with this analysis in mind. Most respondents do not have serious health problems, and Model 1 does not necessarily improve the goodness of fit. In this respect, our results, whilst robust, should be regarded as producing exploratory evidence of the preference mechanism. Further examination of the topic requires prospective research that oversamples individuals with severe health problems.

We are also unable to interpret our estimates as being necessarily caused by respondents having a serious health problem. Other characteristics associated with health problems could instead be driving the preference. While we take measures to account for this by restricting respondents who do not report serious health problems to those that are demographically similar to those who do (based on age, gender, education, and marital status), we cannot be sure that some confounding does not persist.

Respondents reported their own health state "today" by considering their health on the day they completed the EQ-5D-5L questionnaire. While this enabled us to examine cross-dimensional effects, it also makes it difficult to identify those experiencing acute compared to long-term health problems. The preference mechanism might be more likely to develop over time. If so, our study is likely to underestimate the cross-dimensional effect of having long-term serious problems with PD or AD. Similarly, we are unable to capture whether individuals not self-reporting a particular health problem "today" have had prior exposure to a problem and therefore have preferences that are closer to those who report the problem than those that do not. If so, we could again be underestimating the extent of the preference mechanism.

Whilst studying the effect of self-reported EQ-5D-5L problems allows us to examine the effect of current personal health experience on health state valuation, current experience is not the only way to become knowledgeable about a health state. For example, individuals who have past health experience, anticipate future health experience, or know somebody with health experience may also be better informed about a given health state than most of the general public (Cubi-Molla et al., 2018). Our study does not seek to determine how a wider definition of health experience affects health state valuation.

We do not attempt to speculate why the preference mechanism is best observed for having serious AD on the PD levels of severity. Qualitative methods could provide insight into why cross-dimensional effects are observed for some EQ-5D dimensions and not others, but this is beyond the scope of our study.

4.3. Conclusion

This paper proposes a new framework to analyse the effect of respondents' health on health state valuations, with implications for the normative debate on what health states should be used in the estimation of a value set to inform publicly funded health care resource allocation. Normatively speaking, we acknowledge that it is relatively uncontentious to use experience-based, non-representative samples to estimate the value set if the objective is to ascertain the sample is well informed, but only provided that opinions on health state do not depend on health experience. Privileging the opinions of select groups is more contentious and requires a stronger justification. In analysing the impact of serious problems on health state valuation, we propose a distinction between within-dimensional effects and cross-dimensional effects because the latter indicates that the preference mechanism is present. And finally, we present evidence of cross-dimensional effects of self-reporting serious AD on preferences for PD severity levels. The findings, if established in larger datasets with more information on respondent's health problems, mean that the use of experience-based value sets cannot be justified simply because those with serious health problems are better informed about living with those problems: they have different opinions of health states from those without the same health experience.

Credit author statement

Jack Elliott: Conceptualisation, Methodology, Formal Analysis, Software, Writing - Original Draft, Writing - Review & Editing, Visualisation. **Aki Tsuchiya:** Conceptualisation, Methodology, Writing - Original Draft, Writing - Review & Editing, Visualisation.

Data availability

The authors do not have permission to share data.

Acknowledgements

We would like to thank David Mott, Juan Manuel Ramos-Goñi, and Phil Shackley for their suggestions, and the FEDEV project for the use of the data. Our special thanks go to Arne Risa Hole. JE's time was funded by the National Institute for Health and Care Research Pre-doctoral Research Methods Fellowship programme (RM–FI–2017-09-011). The views and opinions expressed in this paper are those of the author(s) and do not necessarily reflect those of the National Institute for Health and Care Research.

Appendices. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2022.115474.

References

Al Shabasy, S., Al Sayah, F., Abbassi, M., Farid, S., 2022. Determinants of health preferences using data from the Egyptian EQ-5D-5L valuation study. Patient - Patient-Cent. Outc. Res. https://doi.org/10.1007/s40271-022-00572-0.
Bansback, N., Brazier, J., Tsuchiya, A., Anis, A., 2012. Using a discrete choice experiment to estimate health state utility values. J. Health Econ. 31 (1), 306–318. https://doi.org/10.1016/j.jhealeco.2011.11.004.
Brazier, J., Akehurst, R., Brennan, A., Dolan, P., Claxton, K., McCabe, C., Sculpher, M., Tsuchyia, A., 2005. Should patients have a greater role in valuing health states? Appl. Health Econ. Health Pol. 8.
Brazier, J., Rowen, D., Karimi, M., Peasgood, T., Tsuchiya, A., Ratcliffe, J., 2018. Experience-based utility and own health state valuation for a health state classification system: why and how to do it. Eur. J. Health Econ. 19 (6), 881–891.
Burström, K., Sun, S., Gerdtham, U.-G., Henriksson, M., Johannesson, M., Levin, L.-Å., Zethraeus, N., 2014. Swedish experience-based value sets for EQ-5D health states. Qual. Life Res. 23 (2), 431–442.
Cubi-Molla, P., Shah, K., Burström, K., 2018. Experience-based values: a framework for classifying different types of experience in health valuation research. Patient-Patient-Cent. Outc. Res. 11 (3), 253–270.
De Wit, G.A., Busschbach, J.J., De Charro, F.T., 2000. Sensitivity and perspective in the valuation of health status: whose values count? Health Econ. 9 (2), 109–126.
Devlin, N.J., Shah, K.K., Feng, Y., Mulhern, B., van Hout, B., 2018. Valuing health-related quality of life: an EQ-5 D-5 L value set for E ngland. Health Econ. 27 (1), 7–22.
Heijink, R., Reitmeir, P., Leidl, R., 2017. International comparison of experience-based health state values at the population level. Health Qual. Life Outcome 15 (1), 1–13.
Helgesson, G., Ernstsson, O., Åström, M., Burström, K., 2020. Whom should we ask? A systematic literature review of the arguments regarding the most accurate source of information for valuation of health states. Qual. Life Res. 29 (6), 1465–1482.
Herdman, M., Gudex, C., Lloyd, A., Janssen, M.F., Kind, P., Parkin, D., Bonsel, G., Badia, X., 2011. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual. Life Res. 20 (10), 1727–1736.
Insinga, R.P., Fryback, D.G., 2003. Understanding differences between self-ratings and population ratings for health in the EuroQOL. Qual. Life Res. 12 (6), 611–619.
Jakubczyk, M., Craig, B.M., Barra, M., Groothuis-Oudshoorn, C.G., Hartman, J.D., Huynh, E., Ramos-Goñi, J.M., Stolk, E.A., Rand, K., 2018. Choice defines value: a predictive modeling competition in health preference research. Value Health 21 (2), 229–238.
Jonker, M.F., Attema, A.E., Donkers, B., Stolk, E.A., Versteegh, M.M., 2017. Are health state valuations from the general public biased? A test of health state reference dependency using self-assessed health and an efficient discrete choice experiment. Health Econ. 26 (12), 1534–1547. https://doi.org/10.1002/hec.3445.
Kennedy-Martin, M., Slaap, B., Herdman, M., van Reenen, M., Kennedy-Martin, T., Greiner, W., Busschbach, J., Boye, K.S., 2020. Which multi-attribute utility instruments are recommended for use in cost-utility analysis? A review of national health technology assessment (HTA) guidelines. Eur. J. Health Econ. 21 (8), 1245–1257.
Krabbe, P.F., Tromp, N., Ruers, T.J., van Riel, P.L., 2011. Are patients' judgments of health status really different from the general population? Health Qual. Life Outcome 9 (1), 31. https://doi.org/10.1186/1477-7525-9-31.
Ludwig, K., Ramos-Goñi, J.M., Oppe, M., Kreimeier, S., Greiner, W., 2021. To what extent do patient preferences differ from general population preferences? Value Health 24 (9), 1343–1349.
Mahalanobis, P.C., 1936. On the Generalized Distance in Statistics.
Mann, R., Brazier, J., Tsuchiya, A., 2009. A comparison of patient and general population weightings of EQ-5D dimensions. Health Econ. 18 (3), 363–372.
Mott, D.J., Shah, K.K., Ramos-Goñi, J.M., Devlin, N.J., Rivero-Arias, O., 2021. Valuing EQ-5D-Y-3L health states using a discrete choice experiment: do adult and adolescent preferences differ? Med. Decis. Making 41 (5), 584–596.
Mulhern, B., Bansback, N., Hole, A.R., Tsuchiya, A., 2017. Using discrete choice experiments with duration to model EQ-5D-5L health state preferences: testing experimental design strategies. Med. Decis. Making 37 (3), 285–297. https://doi.org/10.1177/0272989X16670616.

Mulhern, B., Norman, R., Street, D.J., Viney, R., 2019. One method, many methodological choices: a structured review of discrete-choice experiments for health state valuation. Pharmacoeconomics 37 (1), 29–43.

NICE, 2022. Health Technology Evaluations: the Manual. NICE, p. 181. In: https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741.

Ogorevc, M., Murovec, N., Fernandez, N.B., Rupel, V.P., 2019. Questioning the differences between general public vs. Patient based preferences towards EQ-5D-5L defined hypothetical health states. Health Pol. 123 (2), 166–172. https://doi.org/10.1016/j.healthpol.2017.03.011.

Peeters, Y., Stiggelbout, A.M., 2010. Health state valuations of patients and the general public analytically compared: a meta-analytical comparison of patient and population health state utilities. Value Health 13 (2), 306–309. https://doi.org/10.1111/j.1524-4733.2009.00610.x.

Pickard, A.S., Tawk, R., Shaw, J.W., 2013. The effect of chronic conditions on stated preferences for health. Eur. J. Health Econ. 14 (4), 697–702. https://doi.org/10.1007/s10198-012-0421-8.

Rand-Hendriksen, K., Augestad, L.A., Kristiansen, I.S., Stavem, K., 2012. Comparison of hypothetical and experienced EQ-5D valuations: relative weights of the five dimensions. Qual. Life Res. 21 (6), 1005–1012.

Ratcliffe, J., Brazier, J., Palfreyman, S., Michaels, J., 2007. A comparison of patient and population values for health states in varicose veins patients. Health Econ. 16 (4), 395–405.

The Dental and Pharmaceutical Benefits Agency, 2003. General Guidelines for Economic Evaluations from the Pharmaceutical Benefits Board (LFNAR 2003:2). https://www.tlv.se/download/18.2e53241415e842ce95514e9/1510316396792/Guidelines-for-economic-evaluations-LFNAR-2003-2.pdf.

Ubel, P.A., Loewenstein, G., Jepson, C., 2003. Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public. Qual. Life Res.: Int. J. Qual. Life Aspects Treat. Care Rehabil. 12 (6), 599–607. https://doi.org/10.1023/a:1025119931010.

Vass, C.M., Wright, S., Burton, M., Payne, K., 2018. Scale heterogeneity in healthcare discrete choice experiments: a primer. Patient-Patient-Cent. Outc. Res. 11 (2), 167–173.

Versteegh, M., Brouwer, W., 2016. Patient and general public preferences for health states: A call to reconsider current guidelines. Social Science and Medicine 165, 66–74. Springer.

Versteegh, M.M., Vermeulen, K.M., Evers, S.M., De Wit, G.A., Prenger, R., Stolk, E.A., 2016. Dutch tariff for the five-level version of EQ-5D. Value Health 19 (4), 343–352.

Wang, P., Tai, E.S., Thumboo, J., Vrijhoef, H.J.M., Luo, N., 2014. Does diabetes have an impact on health-state utility? A study of asians in Singapore. Patient - Patient-Cent. Outc. Res. 7 (3), 329–337. https://doi.org/10.1007/s40271-014-0059-y.

Wolff, J., Edwards, S., Richmond, S., Orr, S., Rees, G., 2012. Evaluating interventions in health: a reconciliatory approach. Bioethics 26 (9), 455–463.