WORKING
PAPER

How do the public value different outcomes of social care?

Estimation of preference weights for ASCOT

PETER BURGE, DIMITRIS POTOGLOU,
CHONG WOO KIM, STEPHANE HESS

WR-744-ONS

June 2010

Prepared for PSSRU and Office for National Statistics

EUROPE

Preface

This working paper, prepared for and funded by the UK Office for National Statistics, presents the results of a preference study to elicit preference weights to support the development of the Adult Social Care Outcomes Toolkit (ASCOT).

The specific objectives of this working paper are to:

- Establish preference weights for the Adult Social Care Outcomes Toolkit (ASCOT) measures of social care outcomes, and
- Examine whether the use of 4-level domains improves the sensitivity of the measure at lower levels of need than 3-level domains.

The report should be of interest to those interested in developing outcome measures for public service interventions, particularly those within social care services. The direct audience is the team developing the ASCOT at the Personal Social Services Research Unit (PSSRU) and the Office for National Statistics (ONS), however, aspects of this work will be of interest to the wider academic and policy analysis community.

The research reported within this document was undertaken by RAND Europe; however, the study also drew on expertise within PSSRU, ONS, and the data collection supporting this research was undertaken by Accent.

RAND Europe is an independent not-for-profit policy research organization that serves the public interest by improving policymaking and informing public debate. Clients are European governments, institutions, and firms with a need for rigorous, impartial, multidisciplinary analysis of the hardest problems they face. This report has been peer-reviewed in accordance with RAND's quality assurance standards (see http://www.rand.org/about/standards/).

For more information about RAND Europe or this document, please contact:

Peter Burge
Associate Director of Modelling

RAND Europe
Westbrook Centre, Milton Road
Cambridge, CB4 1YG
United Kingdom

Tel: +44 1223 353329
e-mail: burge@rand.org

Contents

Table of Figures

Table of Tables

Acknowledgments

This report is part of a stream of work in the development of outcome measures of social care. It forms part of the work undertaken by PSSRU in developing the Adult Social Services Outcome Toolkit (ASCOT). In addition, the study builds on the feasibility study undertaken as part of the development of the Outcomes of Social Care for Adults (OSCA) measure, funded by the Department of Health HTA Programme.

The assistance of Ann Netten, Julien Forder and Juliette Malley at PSSRU is gratefully acknowledged, as is the contribution of Beryl Wall and Rob Sheldon from Accent, who collected the data. Finally we thank Terry Flynn for his guidance on the development of the best-worst scaling exercises, Prof Andrew Daly for his guidance on the specification of the models to incorporate the panel nature of the data, and Charlene Rohr for her QA review of this Working Paper.

CHAPTER 1 Survey design

1.1 Introduction

This study is part of the Measuring Outcomes for Public Service Users (MOPSU) project funded by the UK Treasury under the Invest to Save programme and led by the Office for National Statistics (ONS).

This preference study has been undertaken as part of the personal social services element of the project, which is led by the Personal Social Services Research Unit (PSSRU), with the objective that the measures of social care outcomes will reflect the relative importance of the domains (e.g. food and nutrition, accommodation) and levels of these domains, rather than an assumption that all domains, and improvements between levels within those domains, are of equal importance.

The specific objectives of the present study are to:

- Establish preference weights for the Adult Social Care Outcomes Toolkit (ASCOT) measures of social care outcomes, and
- Examine whether the use of 4-level domains improves the sensitivity of the measure at lower levels of need than 3-level domains.

1.2 Design of best-worst scaling experiments

This study looked at two different instruments for classifying quality of life outcomes: an earlier version of the measure used in a study of low-level services (LLS) (Caeils et al 2010), which had nine 3-level domains, and the Outcomes of Social Care for Adults (OSCA) measure (currently under development) with eight 4-level domains[1]. These are summarised in Table 1. The values that individuals place on each of the domain levels were explored using best-worst scaling experiments. A sample of the general public was asked to participate in surveys which contained these experiments, half of which were asked to

[1] One of the objectives of the MOPSU studies conducted by PSSRU was to test the domains which had been identified in previous work as encompassing social care related quality of life (SCRQOL). The domain of 'anxiety' identified in the LLS study was dropped from subsequent work as potentially either duplicating concerns in other domains (e.g. safety) or being beyond the remit of social care interventions. For example anxiety may be due to mental health problems and thus more relevant to health outcome measures, which are excluded from this measure, or, at a less severe level, related to personality type.

consider the LLS measure, and half the OSCA measure. The wording of each domain for the 3-level and 4-level versions is given in Table 33 in Appendix A.

The best-worst scaling tasks in both experiments were developed using fractional-factorial orthogonal design matrices. These designs allow estimation of the main effects of all nine LLS domains with three levels and all eight OSCA domains with four levels. In particular, the full plan for the LLS design consisted of 3^9 possible scenarios, which were reduced to 27 scenarios using an orthogonal fractional factorial design. The full factorial for the OSCA design consisted of 4^8 possible scenarios, which again were reduced to 32 scenarios using an orthogonal fractional factorial design (see Table 1).

Table 1: Definition of domains in the experimental design

Domain		LLS	OSCA
1	Accommodation, cleanliness and comfort The service user feels their home environment, including all the rooms, is clean and comfortable	3	4
2	Safety The service user feels safe and secure. This means being free from fear of abuse, falling or other physical harm and fear of being attacked or robbed	3	4
3	Food and nutrition The service user feels he/she has a nutritious, varied and culturally appropriate diet with enough food and drink he/she enjoys at regular and timely intervals	3	4
4	Personal cleanliness and comfort The service user feels he/she is personally clean and comfortable and looks presentable or, at best, is dressed and groomed in a way that reflects his/her personal preferences	3	4
5	Control over daily life The service user can choose what to do and when to do it, having control over his/her daily life and activities	3	4
6	Social participation and involvement The service user is content with their social situation, where social situation is taken to mean the sustenance of meaningful relationships with friends, family and feeling involved or part of a community should this be important to the service user	3	4
7	Dignity The psychological impact of support and care on the service user's personal sense of significance	3	4
8	Occupation and employment The service user is sufficiently occupied in a range of meaningful activities whether it be formal employment, unpaid work, caring for others or leisure activities	3	4
9	Anxiety The service user does not feel distressed or uneasy about something in their personal lives that is happening or might happen in the future	3	-
Size of full-factorial design matrix		3^9 (19,683)	4^8 (65,536)
Size of fractional factorial design matrix		27	32

Each design matrix in the LLS and OSCA experiment was further blocked into three and four segments, respectively, using a procedure which sought to minimise the correlation of the levels being presented for the domains with the block (i.e. avoiding all respondents in one version of the survey always seeing a given domain at the same level). Each respondent received nine best-worst tasks if they participated in the LLS experiment and eight tasks if participated in the OSCA experiment. An additional consideration in defining the blocks was the minimisation of "easy-options" or straightforward choices in each task. To achieve this, the domain levels were recoded in the design matrices through an iterative procedure to ensure that each scenario had more than one domain at the end-point levels (i.e., level 1 or 4), thereby generating situations where explicit trade-offs between domains were required.

1.3 Format of the best-worst scaling experiments

The best-worst scaling task asked respondents to provide their best, worst, second-best and second-worst choice of domain levels. The method for doing this for those respondents participating in the LLS experiment is shown in Figure 1, where respondents were first asked to choose their best domain from the list of nine domains available. The chosen domain was then removed from the list, and respondents were asked to choose the worst domain in the same task in which the remaining eight domains were available. The chosen domain was again excluded from the list and respondents were asked to choose the second-best domain out of seven domains and finally, the second worst domain out of six domains that were left available.

There are a number of benefits from asking respondents about their second-best and second-worst choices. First, it is a plausible way of reducing the effects of lexicographic and non-trading behaviour in a best-worst task. If a respondent has a strong preference regarding a specific domain then there is the risk that they may always choose that domain as being best (or worst). Asking for the second-best and second-worst, it is therefore possible to gain additional information on the importance of the other issues in the absence of that domain (whilst also recognising the importance of the dominant domain within the analysis). Secondly, it is possible to draw more information per choice task as we are able to obtain a partially complete ranking of domains. By the end of a single best-worst task, we had elicited the ranking of four out of nine domains in the LLS experiment and the ranking of four out of eight domains in the OSCA experiment.

	Aspect of life
1	My home is less clean and comfortable than I want
2	I feel as safe as I want
3	I don't always eat the right meals I want, and I think there is a risk to my health
4	I feel much less clean than I want, with poor personal hygiene
5	Sometimes I don't feel I have as much control over my daily life as I want
6	Sometimes I feel my social situation and relationships are not as good as I want
7	I would be treated by other people with the dignity and respect that I want
8	I don't do any of the activities I want to do
9	I sometimes feel worried and concerned

(a) Which of these nine aspects would rate as being the best?

	Aspect of life
1	My home is less clean and comfortable than I want
3	I don't always eat the right meals I want, and I think there is a risk to my health
4	I feel much less clean than I want, with poor personal hygiene
5	Sometimes I don't feel I have as much control over my daily life as I want
6	Sometimes I feel my social situation and relationships are not as good as I want
7	I would be treated by other people with the dignity and respect that I want
8	I don't do any of the activities I want to do
9	I sometimes feel worried and concerned

(b) Which of the remaining eight aspects would rate as being the worst?

	Aspect of life
1	My home is less clean and comfortable than I want
3	I don't always eat the right meals I want, and I think there is a risk to my health
4	I feel much less clean than I want, with poor personal hygiene
5	Sometimes I don't feel I have as much control over my daily life as I want
6	Sometimes I feel my social situation and relationships are not as good as I want
8	I don't do any of the activities I want to do
9	I sometimes feel worried and concerned

(c) Which of the remaining seven aspects would now rate as being the next best?

	Aspect of life
1	My home is less clean and comfortable than I want
4	I feel much less clean than I want, with poor personal hygiene
5	Sometimes I don't feel I have as much control over my daily life as I want
6	Sometimes I feel my social situation and relationships are not as good as I want
8	I don't do any of the activities I want to do
9	I sometimes feel worried and concerned

(d) Which of the remaining six aspects would now rate as being the next worst?

Figure 1: An illustration of the sequence in the best-worst scaling task (LLS experiment)

Each respondent received randomised versions of the best-worst tasks in which the domain-levels were in a different order. This procedure helped to control for possible ordering bias. The *a priori* expectation was that respondents would be more likely to choose domains that appeared in the first or last two places of a best-worst task. Therefore, randomising the order of domains across respondents ensures that each domain had an equal probability of being in any position within a best-worst task. The intention was that the randomisation would occur between individuals, but not within individuals.

Due to an error in the survey development and miscommunication between the market research agency and their interviewers, 37% of the sample for both the LLS and OSCA based surveys were given a version of the survey where the order of the domains in were also randomised within the best-worst tasks for an individual. This is referred as "excessive

randomisation" in Table 2. In those cases, the best-worst experiment ran the risk of being confusing for respondents.

Table 2: Participants in the excessive and intended version of randomisation in best-worst tasks

	LLS 3-level version		OSCA 4-level version	
	Intended randomisation	Excessive randomisation	Intended randomisation	Excessive Randomisation
Number of respondents	305	182	323	190
% of sample	63%	37%	63%	37%

To obtain a better understanding of the issue, the diagnostic questions collected within the survey were examined to determine whether respondents who participated in the version with excessive randomisation were able to understand the choice task or whether this additional randomisation led to greater confusion amongst the respondents. Overall, the diagnostics summarised in Table 3 suggested that there was more confusion amongst respondents that were given the best-worst tasks in which the order of the domains kept changing in each task. However, the increase in confusion was not large, and even after those that were identified as exhibiting lower levels of understanding were excluded from the dataset there was still a significant volume of data for developing the models.

Table 3: Level of respondents' understanding in the excessive and intended version of randomisation in best-worst tasks

Diagnostic (used as exclusion criteria in model)	LLS 3-level version		OSCA 4-level version	
	Intended randomisation	Excessive randomisation	Intended randomisation	Excessive randomisation
In the choices, the respondent did not understand the description	2.6%	2.7%	2.5%	4.7%
The respondent was not able to answer the choices	1.0%	1.6%	1.9%	3.2%
The interviewer thinks the respondent didn't understand what he was asked to do	1.3%	7.1%	0.6%	7.4%
The interviewer thinks the respondent gave little or no consideration in responding	1.3%	2.7%	1.9%	3.2%

In addition, two sets of models were developed for both the LLS and the OSCA experiment data to test whether the responses with correct and excessive randomisation were significantly different from each other. The models in both the LLS and the OSCA experiment showed that the data from respondents that saw best-worst tasks with excessive randomisation contained more noise, but the point estimates of the domain level coefficients were not on the whole significantly different (once the scale difference between the datasets was taken into account). The impact of the excessive randomisation in the best-worst task was therefore found to be relatively small, and it is was therefore possible to specify the models in a way to capture the difference in the noise between the two questionnaire versions.

CHAPTER 2 Sample characteristics

2.1 Sample composition

The survey was undertaken with 1,000 respondents located in Birmingham, Newcastle, London and the South East of England. The surveys were undertaken using face-to-face interviews in respondents' homes, with the interviewer using a computerised version of the questionnaire on a laptop computer. The market research agency was commissioned to collect a sample that was nationally representative, with half of the respondents seeing a best-worst scaling experiment based on the LLS definition of the domains, and half seeing a best-worst scaling experiment based on the OSCA definition of the domains.

The background information of the participants is shown in Table 4 to Table 15, where they are compared between experiments and with general population figures. From these tables we can observe that the samples of respondents facing the three and four level versions of the domains were broadly consistent in all of these observable dimensions.

Five hundred and twenty six (526) participants were females and 474 were male. The distribution of males and females is consistent across the two experiments and representative of the national population, as shown in Table 4 below.

Table 4: Gender distribution

	LLS experiment		OSCA experiment		Total		Population
	Frequency	%	Frequency	%	Frequency	%	%
Male	233	47.8	241	47.0	474	47.4	48.6
Female	254	52.2	272	53.0	526	52.6	51.3
Total	**487**	**100**	**513**	**100**	**1,000**	**100**	**100**

(population: ONS mid 2008 estimate, Table 4 [England])

Table 5 presents the age distribution of the sample. Older people were sufficiently represented in the sample with the total percentage of individuals aged 65 or older being 18% in the complete sample, again this distribution is consistent across the experiments, with 18% in the LLS experiment and 19% in OSCA experiment. The age distribution is broadly representative of the national population, although the sample under-represents those aged 80 years or older.

Table 5: Age distribution

	LLS experiment		OSCA experiment		Total		Population
	Frequency	%	Frequency	%	Frequency	%	%
18-24 years	44	9.0	57	11.5	101	10.1	12.1
25-30 years	61	12.5	61	11.9	122	12.2	10.1
31-39 years	71	14.6	70	13.6	141	14.1	15.8
40-49 years	88	18.1	98	19.1	186	18.6	18.8
50-64 years	136	27.9	132	25.7	268	26.8	22.8
65-79 years	73	15.0	84	16.4	157	15.7	14.7
80 years or older	14	2.9	11	2.1	25	2.5	5.8
Total	**487**	**100**	**513**	**100**	**1,000**	**100**	**100**

(population: ONS mid 2008 estimate, Table 4 [England])

Table 6 shows the education level of participants. A notable proportion of respondents in the sample have GCSE qualifications (29%) whereas 26% of respondents have no formal qualification. The distribution in the population is slightly different, but if the two groups are considered together, the proportion with GCSE grade or lower is similar.

Table 6: Level of education

	LLS experiment		OSCA experiment		Total		Population
	Frequency	%	Frequency	%	Frequency	%	%
No formal qualification	132	27.1	123	24.0	255	25.5	18.9
GCSE / O level	140	28.7	152	29.6	292	29.2	34.4
A levels or equivalent	60	12.3	52	10.1	112	11.2	12.4
Professional qualification below degree level	66	13.6	93	18.1	159	15.9	8.7
Degree level qualification or equivalent	60	12.3	65	12.7	125	12.5	7.8
Higher degree	18	3.7	16	3.1	34	3.4	3.5
Other	11	2.3	12	2.3	23	2.3	14.4
Total	**487**	**100**	**513**	**100**	**1,000**	**100**	**100**

(population: 2001 Census, National report for England and Wales, Table S115 [England and Wales])

As shown in Table 7, only thirteen (13) out of 1,000 respondents belonged to the highest category of social grade (A). More than half of the participants belonged to band C (50%). The percentage of respondents being in social grade C was higher in the sample for the OSCA experiment (52%) than the LLS experiment (47%), but broadly the distribution was consistent across the two experiments. The sample over-represents those in band C2, and under-represents those in bands D and E.

Table 7: Social grade

	LLS experiment		OSCA experiment		Total		Population
	Frequency	%	Frequency	%	Frequency	%	%
A	9	1.8	4	0.8	13	1.3	22.0
B	113	23.2	111	21.6	224	22.4	
C1	121	24.8	141	27.5	262	26.2	29.7
C2	109	22.4	126	24.6	235	23.5	15.1
D	66	13.6	68	13.3	134	13.4	17.2
E	69	14.1	61	11.9	130	13.0	16.1
Not stated			2	0.4	2.0	0.2	-
Total	**487**	**100**	**513**	**100**	**1,000**	**100**	**100**

(population: 2001 Census, National report for England and Wales, Table S066 [England and Wales])

Table 8 shows that approximately half of the participants were married.

Table 8: Marital status

	LLS experiment		OSCA experiment		Total		Population
	Frequency	%	Frequency	%	Frequency	%	%
Married	250	51.3	248	48.3	498	49.8	51.6
Living together	59	12.1	64	12.5	123	12.3	9.9
Single	95	19.5	101	19.7	196	19.6	23.1
Widowed	36	7.4	45	8.8	81	8.1	7.6
Divorced	36	7.4	42	8.2	78	7.8	5.9
Separated	10	2.1	11	2.1	21	2.1	1.9
Refused/DK	1	0.2	2	0.4	3	0.3	-
Total	**487**	**100**	**513**	**100**	**1,000**	**100**	**100**

(population: 2001 Census, Key Statistics for Health Areas, Table KS03 [England])

As shown in Table 9, only 12% of the participants worked part-time, 39% were working full-time, approximately 24% were retired and about 5% of participants were unable to work because of medical reasons. It is notable that the sample under-represents those in full time employment, but over-represents those that are retired.

Table 9: Employment status

	LLS experiment		OSCA experiment		Total		Population
	Frequency	%	Frequency	%	Frequency	%	%
Working full time (30+ hours a week)	187	38.4	206	40.2	393	39.3	49.1
Working part time (less than 30 hours a week)	58	11.9	65	12.7	123	12.3	11.8
Full-time student	18	3.7	16	3.1	34	3.4	7.3
Part-time student	5	1.0	5	1.0	10	1	
Not working – looking for work	22	4.5	23	4.5	45	4.5	3.4
Not working – not looking for work	15	3.1	10	1.9	25	2.5	
Not working – unable for medical reasons	27	5.5	23	4.5	50	5	5.3
Retired	116	23.8	125	24.4	241	24.1	13.5
Looking after home	37	7.6	37	7.2	74	7.4	6.5
Other	2	0.4	3	0.6	5	0.5	3.1
Total	**487**	**100**	**513**	**100**	**1,000**	**100**	**100**

(population: 2001 Census, Key Statistics for Health Areas, Table KS09 [England])

The majority of participants were white, including white British (79%), white Irish (1%) and any other white background (2%). Overall, 10% of respondents were of Asian background including Indian (3.5%) and Pakistani (3.6%). The sample over-represents the non-white (although this is useful in providing more data within this segment for the purposes of estimating the models).

Table 10: Ethnicity

	LLS experiment		OSCA experiment		Total		Population
	Frequency	%	Frequency	%	Frequency	%	%
White	394	80.9	419	81.7	813	81.3	90.9
Asian	48	9.9	51	9.9	99	9.9	5.0
Black	38	7.8	32	6.2	70	7	2.3
Any other background	7	1.4	11	2.1	18	1.8	1.8
Total	**487**	**100**	**513**	**100**	**1,000**	**100**	**100**

(population: 2001 Census, Key Statistics for Health Areas, Table KS06 [England])

Table 11 shows the number of children under the age of 16 who live in the household. In 64.3% of the households there are no children under the age of 16.

Table 11: Number of children

	LLS experiment		OSCA experiment		Total		Population
	Frequency	%	Frequency	%	Frequency	%	%
None	316	64.9	327	63.7	643	64.3	56.6
1	75	15.4	81	15.8	156	15.6	18.5
2	64	13.1	77	15.0	141	14.1	17.2
3	25	5.1	18	3.5	43	4.3	
4	5	1.0	8	1.6	13	1.3	7.7
5 or more	2	0.4	2	0.4	4	0.4	
Total	**487**	**100**	**513**	**100**	**1,000**	**100**	**100**

(population: 2001 Census, National report for England and Wales, Table S007 [England and Wales])

Table 12 shows the annual household income distribution of respondents. A significant proportion of respondents did not know (22%) or refused to report (15%) their income. Of those who reported their income, a relatively high proportion of respondents (61%) reported having income lower than £25,000. In contrast, we see that 50% of the population have an income less that £26,000. This suggests that our sample contains a higher proportion of lower income respondents than may be expected. However, it should also be noted that we do not know the composition of those that refused or reported that they did not know their income, and it may be that lower income respondents are more likely to provide such responses.

Table 12: Annual household income

	LLS experiment		OSCA experiment		Total		Population
	Frequency	%	Frequency	%	Frequency	%	%
Under £4,500 per year	8	1.6	4	0.8	12	1.2	
£4,500-£6,499	19	3.9	25	4.9	44	4.4	
£6,500-£7,499	19	3.9	14	2.7	33	3.3	
£7,500-£9,499	20	4.1	22	4.3	42	4.2	28.0[2]
£9,500-£11,499	25	5.1	30	5.8	55	5.5	
£11,500-£13,499	14	2.9	23	4.5	37	3.7	
£13,500-£15,499	16	3.3	17	3.3	33	3.3	
£15,500-£17,499	19	3.9	29	5.7	48	4.8	22.0[3]
£17,500-£24,999	35	7.2	45	8.8	80	8.0	
£25,000-£29,999	36	7.4	29	5.7	65	6.5	
£30,000-£39,999	42	8.6	33	6.4	75	7.5	31.0[4]
£40,000-£49,999	22	4.5	28	5.5	50	5.0	
£50,000-£74,999	18	3.7	16	3.1	34	3.4	
£75,000-£99,999	4	0.8	6	1.2	10	1.0	19.0[5]
£100,000+	3	0.6	5	1.0	8	0.8	
Refused	113	23.2	111	21.6	224	22.4	-
DK	74	15.2	76	14.8	150	15.0	-
Total	**487**	**100**	**513**	**100**	**1,000**	**100**	**100**

(population: 2007/08 Family Resource Survey, Table 3.6 [England])

As shown in Table 13, the majority of respondents lived in a large city or the suburbs of a large city (55%) and 45% lived in a small town or a village. We would not expect the sample to be representative within these classifications as it was recruited from four distinct geographic areas.

Table 13: Area of residence

	LLS experiment		OSCA experiment		Total		Population
	Frequency	%	Frequency	%	Frequency	%	%
A London Borough	109	22.4	125	24.4	234	23.4	14.7
Another city / large town	78	16.0	75	14.6	153	15.3	66.3
Suburbs of a city / large town	81	16.6	82	16.0	163	16.3	
A small town	107	22.0	119	23.2	226	22.6	9.1
A rural area or village	112	23.0	112	21.8	224	22.4	9.8
Total	**487**	**100**	**513**	**100**	**1,000**	**100**	**100**

(population: 2001 Census, Key Statistics, Table KS20 by rural and urban classification [England])

Table 14 shows how respondents rated their health at the time of the interview. Approximately, 73% of respondents rated their health being good or better. Only 5 respondents claimed to have very bad health. We have fewer respondents with poor health than was reported within the 2001 census.

[2] Under £15,000
[3] Actual range is £15,000 - £26,000
[4] Actual range is £26,000 - £52,000
[5] £52,000 or more

Table 14: Self-reported health status

	LLS experiment		OSCA experiment		Total		Population
	Frequency	%	Frequency	%	Frequency	%	%
Very good	132	27.1	151	29.4	283	28.3	68.8
Good	219	45.0	227	44.2	446	44.6	
Fair	110	22.6	107	20.9	217	21.7	22.2
Bad	22	4.5	27	5.3	49	4.9	9.0
Very bad	4	0.8	1	0.2	5	0.5	
Total	**487**	**100**	**513**	**100**	**1,000**	**100**	**100**

(population: 2001 Census, Key Statistics for Health Areas, Table KS08 [England])

Finally, Table 15 shows that 30% of respondents or a member of their household received some type of benefit.

Table 15: Respondent or household member who received benefits

	LLS experiment		OSCA experiment		Total		Population
	Frequency	%	Frequency	%	Frequency	%	%
Income support	44	9.0	39	7.6	83	8.3	16.0
Working families' tax credit	36	7.4	38	7.4	74	7.4	
Income-based Jobseeker's Allowance	13	2.7	16	3.1	29	2.9	3.0
Disabled person's tax credit	12	2.5	9	1.8	21	2.1	15.0
Housing benefit	61	12.5	62	12.1	123	12.3	20.0
Council tax benefit	79	16.2	81	15.8	160	16.0	16.0
Pension credit	28	5.7	31	6.0	59	5.9	29.0
Do not receive any benefit	335	68.8	362	70.6	697	69.7	32.0

(population: 2000/01 Regional Trends 37, Table KS8.8 [England])

These tables show that the sample for the two experiments are well balanced, and are broadly representative of the national population. There are some dimensions where the sample is known to under- or over-represent some groups, but these differences can be addressed through weighting of the model results.

2.2 Respondents' self classification of quality of life

At the start of the survey, respondents were introduced to each of the social care domains and asked to report the level at which they would assess their current quality of life within each domain.

It was intended that those respondents that would be given the LLS three level measure in the best-worst scaling exercise would be asked to classify themselves on the three level instrument, and the respondents that would be given the OSCA four level measure in the best-worst scaling exercise would be asked to classify themselves on the four level instrument. However, the routing on these questions was incorrect for some respondents, and those respondents that were presented the incorrect survey version with excessive randomisation within the best-worst scaling exercise were also routed to the four level version of the self-classification questions regardless of whether they were shown the three or four level version of the domains in the best-worst scaling exercise.

The following tables present the distribution of respondents across each of the domains. In interpreting these tables it is important to recall that the survey sample is the general population rather than current users of social services

Table 16: Self classification on Accommodation Cleanliness and Comfort domain

Accommodation Cleanliness and Comfort			
OSCA measure (696 respondents)		**LLS measure (304 respondents)**	
1. My home is as clean and comfortable as I want	73.9%		
2. My home is adequately clean and comfortable	23.6%	1. My home is as clean and comfortable as I want	87.2%
3. My home is less than adequately clean or comfortable	2.6%	2. My home is less clean and comfortable than I want	11.5%
4. My home is not at all clean or comfortable	0.0%	3. My home is not at all as clean or comfortable as I want	1.3%

Table 17: Self classification on Safety domain

Safety			
OSCA measure (696 respondents)		**LLS measure (304 respondents)**	
1. I feel as safe as I want	64.2%		
2. Generally I feel adequately safe, but not as safe as I would like	30.2%	1. I feel as safe as I want	65.8%
3. I feel less than adequately safe	5.2%	2. Sometimes I do not feel as safe as I want	30.3%
4. I don't feel at all safe	0.4%	3. I never feel as safe as I want	3.9%

Table 18: Self classification on Food and Nutrition domain

Food and Nutrition			
OSCA measure (696 respondents)		**LLS measure (304 respondents)**	
1. I get all the food and drink I like when I want	77.0%		
2. I get food and drink adequate for my needs	20.5%	1. I eat the meals I like when I want	78.6%
3. I don't get all the food and drink I need, but I don't think there is a risk to my health	2.3%	2. I don't always eat the right meals I want, but I don't think there is a risk to my health	17.8%
4. I don't get all the food and drink I need, and I think there is a risk to my health	0.1%	3. I don't always eat the right meals I want, and I think there is a risk to my health	3.6%

Table 19: Self classification on Personal Care domain

Personal Care			
OSCA measure (696 respondents)		**LLS measure (304 respondents)**	
1. I feel clean and am able to present myself the way I like	78.6%		
2. I feel adequately clean and presentable	20.0%	1. I feel clean and wear what I want	92.4%
3. I feel less than adequately clean or presentable	1.4%	2. I sometimes feel less clean than I want or sometimes can't wear what I want	7.2%
4. I don't feel at all clean or presentable	0.0%	3. I feel much less clean than I want, with poor personal hygiene	0.3%

Table 20: Self classification on Control over Daily Life domain

Control over Daily Life			
OSCA measure (696 respondents)		**LLS measure (304 respondents)**	
1. I have as much control over my daily life as I want	58.0%		
2. I have adequate control over my daily life	33.5%	1. I have as much control over my daily life as I want	69.4%
3. I have some control over my daily life but not enough	7.8%	2. Sometimes I don't feel I have as much control over my daily life as I want	29.3%
4. I have no control over my daily life	0.7%	3. I have no control over my daily life	1.3%

Table 21: Self classification on Social Participate and Involvement domain

Social Participation & Involvement			
OSCA measure (696 respondents)		**LLS measure (304 respondents)**	
1. I have as much social contact as I want with people I like	59.8%		
2. I have adequate social contact with people	31.5%	1. My social situation and relationships are as good as I want	69.4%
3. I have some social contact with people, but not enough	7.2%	2. Sometimes I feel my social situation and relationships are not as good as I want	26.0%
4. I have little social contact with people and feel socially isolated	1.6%	3. I feel socially isolated and often feel lonely	4.6%

Table 22: Self classification on Occupation and Employment domain

Occupation & Employment			
OSCA measure (696 respondents)		**LLS measure (304 respondents)**	
1. I'm able to spend my time as I want, doing things I value or enjoy	43.7%		
2. I'm able do enough of the things I value or enjoy with my time	39.1%	1. I do the activities I want to do	49.7%
3. I do some of the things I value or enjoy with my time but not enough	16.7%	2. I do some of the activities I want to do	45.7%
4. I don't do anything I value or enjoy with my time	0.6%	3. I don't do any of the activities I want to do	4.6%

CHAPTER 3 Exploratory analysis of best-worst experiments

3.1 Choice patterns

Before embarking on the model development, respondents' choices were first examined to see the general choices patterns. Figure 2 shows the choice distribution of the five domain levels that were most frequently chosen as Best, Worst, Second Best and Second Worst across the LLS and the OSCA experiments. The distribution of all domain levels across best, worst, second best and second worst choices for both experiments can be found in Table 34 of the Appendix.

The domain *Control over daily life* at level 1 (i.e., *I have as much control over my daily life as I want*) was the most frequently chosen best domain level in both experiments. In the LLS experiment, the other four most frequently chosen best domains were: *Anxiety, Personal Safety, Social Participation and Employment* all at level 1 (i.e., the highest level of quality of life with regard to these domains). On the other hand, the OSCA experiment had a different distribution of choices as the five most frequently chosen preferences were: *Occupation and Employment* at levels 1 and 2 and *Personal Cleanliness and Comfort* at level 1.

With regard to the worst domain-levels, the domain *Control over daily life* at its lowest levels (i.e., level 3 in the LLS experiment, level 4 in the OSCA experiment) was the most frequently chosen domain level as being the worst. *Social participation*, *Dignity, Safety, and C*all at level 3 (i.e. the lowest level of quality of life relative to these domains in the LLS experiment) were the other four most frequently chosen domain levels. In the OSCA experiment, following *Control over daily life*, the most frequently chosen domain levels were *Safety, Personal Cleanliness and Comfort, Occupation and Employment*, and *Food and Nutrition* all at level 4 (i.e. the lowest level of quality of life relative to these domains in the OSCA experiment).

The five most frequently chosen second-best domain levels in the LLS experiment were *Occupation and Employment, Accommodation, Cleanliness and Comfort, Personal Safety, Control over daily life* and *Social Participation* all at level 1, whereas *Social Participation* (at level 1)*, Employment (*at level 2), *Personal Cleanliness and Comfort* (at level 1)*, Control over daily life (*at level *2)* and *Control over daily life* (at level 1) were the top five domain levels in the OSCA experiment. Finally, the five domain levels most frequently chosen as being

second worst in the LLS experiment were *Control over daily life, Social Participation, Anxiety, Safety and Personal Cleanliness and Comfort* all at the their lowest levels of quality of life relative to these domains (i.e., level 3 in this experiment). In the OSCA experiment, the five most frequently domain levels chosen as being second worst were: *Occupation and Employment*, *Personal Cleanliness and Comfort*, *Social Participation*, *Personal Safety* and *Control over daily life* all being at the lowest level of each domain in terms of quality of life (i.e., level 4).

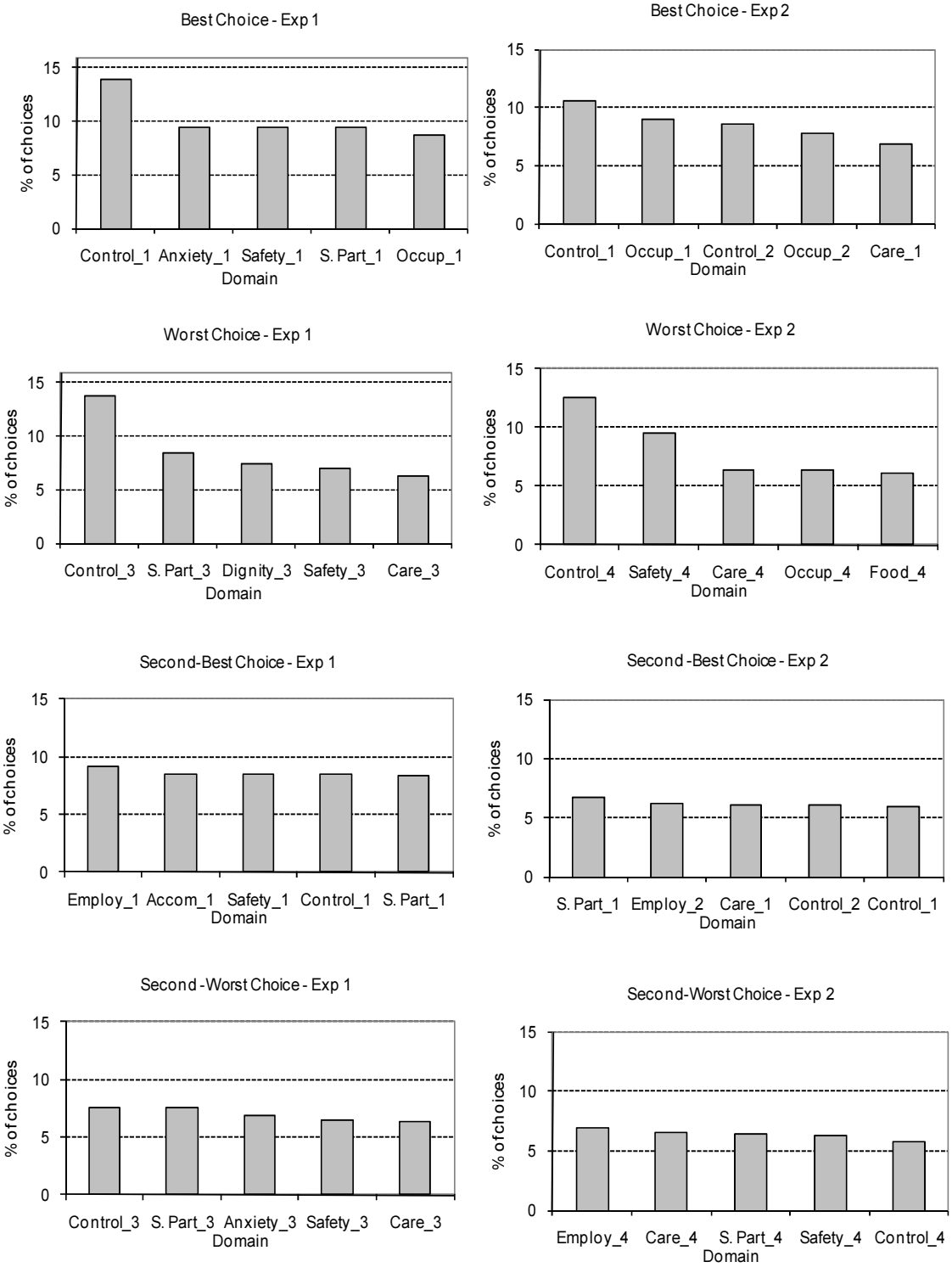

Figure 2: Percentage frequencies of best, worst, second best and second worst domains across the two experiments

(Note: Exp1 = LLS Experiment, Exp2 = OSCA Experiment)

3.2 Diagnostic questions

The survey included a number of diagnostic questions in order to determine whether respondents were able to answer the best-worst experiments, specifically asking:

- Whether respondents were able to put themselves into the imaginary position of requiring help with looking after themselves (see, Q27CHK1 in the questionnaire),
- Understood the descriptions in the choices (Q29CHK2),
- Felt that they were able to answer the choices (Q34CHK7).

Interviewers also provided an assessment of the respondent's ability to answer the questions, indicating whether the respondent:

- Understood what he/she was asked to do in the questions (Q46)
- Put thought in responding (Q47)
- Maintained concentration during the survey (fatigue) (Q48).

Table 23 shows the number of participants that provided negative responses in the diagnostic questions along with interviewer's observations about respondents' ability to get involved with the survey.

Table 23: Summary of responses in the diagnostic questions

Diagnostic questions		LLS experiment		OSCA experiment	
		Number of respondents	%	Number of respondents	%
Q27CHK1	Could not put themselves into an imaginary position	52	10.7	67	13.1
Q29CHK2	Could not understand the descriptions in the choices	13	2.7	17	3.3
Q34CHK7	Felt that they unable to answer the choices	8	1.6	16	3.1
Q46	1. Did not understand at all	7	1.4	10	1.9
	2. Did not understand very much	14	2.9	13	2.5
	3. Understood a little	55	11.3	63	12.3
	4. Understood a great deal	155	31.8	151	29.4
	5. Understood completely	255	52.4	275	53.6
	No response was provided	1	0.2	1	0.2
Q47	1. Gave the questions no consideration	6	1.2	4	0.8
	2. Gave the questions little consideration	19	3.9	28	5.5
	3. Gave the questions some consideration	69	14.2	88	17.2
	4. Gave the questions careful consideration	165	33.9	155	30.2
	5. Gave the questions very careful consideration	227	46.6	237	46.2
	No response was provided	1	0.2	1	0.2
Q48	1. Lost concentration in the later stages	6	1.2	9	1.8
	2. Lessened concentration in the later stages	45	9.2	37	7.2
	3. Maintained concentration with a deal of effort throughout the survey	57	11.7	71	13.8
	4. Maintained concentration with some effort throughout the survey	135	27.7	130	25.3
	5. Easily maintained concentration throughout the	243	49.9	265	51.7

survey No response was provided	1	0.2	1	0.2

The criteria that we set for including an observation in the models were the following:

- Respondents should have been able to:
  - Understand the descriptions in the choices (exclude if Q29CHK2 = 2)
  - Feel that they are able to answer the choices (exclude if Q34CHK7 = 2)
- Interviewers should have indicated that the respondent:
  - Understood the tasks at least a little (exclude if Q46 < 3)
  - Gave the questions some, careful or very careful consideration (exclude if Q47 < 3),
  - Did not lose concentration in the later stages if the survey (exclude if Q48 < 2).

Further tests showed that data obtained from respondents who could not put themselves in an imaginary situation (Q27CHK1 in the questionnaire) had higher variability (i.e., noise) than respondents who could put themselves in an imaginary situation. However, this finding had an insignificant impact on the weights estimated for all domain levels, and it was decided therefore not to exclude respondents who stated that they could not put themselves in the imaginary position that they would require help with looking after themselves as a result of illness or an accident.

Approximately 10% of respondents were excluded from the model analysis, as shown in Table 24. The total number of observations excluded under each criterion is shown in Table 24 where each criterion is applied sequentially.

Table 24: Number of respondents excluded from the discrete choice analysis[6]

Question		LLS experiment	OSCA experiment
Q29CHK2	Could not understand the descriptions in the choices	13	17
Q34CHK7	Felt unable to answer the choices	6	12
Q46 < 3	Did not understand very much or at all	17	16
Q47 < 3	Gave the questions little or no consideration	9	12
Q48[7] < 2	Lost concentration in the later stages	1	2
	Missing information on the order of domains as appeared in a best-worst task	1	1
Total number of observations excluded (percent of the total sample)		**47** (9.6%)	**60** (11.7%)
Total number of respondents included in the analysis		**440**	**453**

[6] The numbers reported in the table represent individuals who have been excluded when each of the above listed condition is applied sequentially. Therefore, numbers reported for each condition may not be in entire agreement with those listed in Table 23.

CHAPTER 4 Modelling the best-worst choice data

The discrete choice analysis of the best-worst data focused on the following themes:

- Estimation of domain weights of the 3- (Exp1) and 4-level (Exp2) versions of the best-worst experiment after controlling for the order of the domains in a given best-worst scaling task
- Estimation of domain weights in the 3- and 4-level versions after controlling for the order of the domains and the socio-economic and demographic characteristics of participants
- Weighting the results of the second set of models to account for the distribution of the relevant socio-economic and demographic characteristics within the population rather than within the sample
- Comparison of the 3-level model with the 4-level model when the latter has been collapsed into a 3-level version by merging levels 1 and 2 in each domain.

The discrete choice models were initially developed under the assumption that the best, worst, second best and second worst choices made by the same respondents were independent from each other (although from a diminishing set of choices). We are aware that this is a simplification and that there is the potential for correlation within the responses from a given individual, both within a given best-worst scenario (where they provide four related responses) and between best-worst scenarios (where each respondent was asked to consider either 8 or 9 different scenarios). These correlations are accounted for in the final models.

The steps involved in setting up a discrete choice model with best-worst scaling data are discussed in Appendix B.

The number of observations available for modelling in the LLS and OSCA data are shown in Table 25.

Table 25: Number of observations available for modelling

Experiment	Number of choices per best-worst task (i.e., best, worst, second best, second worst)	Number of best-worst tasks per respondent	Number of respondents	Number of observations available for modelling
LLS	4	9	440	15840
OSCA	4	8	453	16308

A multinomial logit (MNL) model[8] without explicit consideration of the correlation between the observations was used to develop the initial model specification and to explore heterogeneity in preferences, on the basis that this model structure is considerably simpler to set up and has very short run times, allowing a large number of different model specifications to be tested before moving to a more complex model that explicit accounts for correlation among repeated observations.

Each domain that could have been chosen within the best-worst scenario was specified as a separate alternative within the choice model and was given a utility function that took account of the level at which that domain was presented to the respondent in the scenario in question. The structure of this model is shown in Figure 3.

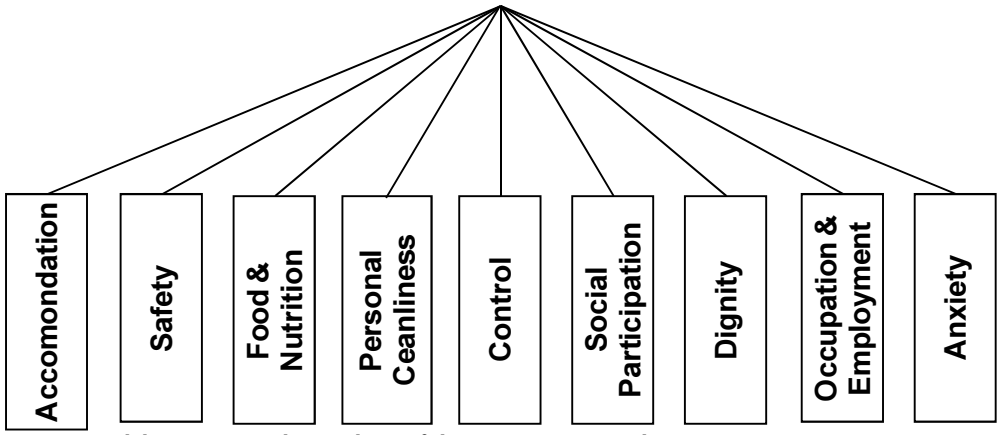

Figure 3: Model structure in the analysis of the LLS experiment data

The MNL model was specified such that the utility function for a given domain was defined as:

- a linear-additive function of the products between the weights (coefficients) to be estimated and the dummy coding of the corresponding domain-levels (only one of which will ever take a value of 1 for any given choice), along with
- a number of dummy-coded variables to control for the position of a domain level in the best-worst task when that domain level was chosen as being best, worst, second best or second worst,.

A generalised example of the utility function specification for the *Accommodation, Cleanliness and Comfort* domain in the LLS experiment is shown in Box 1.

[8] See Train, K.E. (2003), Discrete choice methods with simulation, Cambridge University Press, Cambridge, MA. for further details

Box 1: Generalised form of the utility function for domain Accommodation, Cleanliness and Comfort in the LLS experiment

U_i (Accommodation) =

$+ \beta_{accom1} * (1, \text{if Accommodation Level}^9 = 1)_i * (1, \text{if choice scenario = Best or Second-Best})_i$

$- \beta_{accom1} * (1, \text{if Accommodation Level = 1})_i * (1, \text{if choice scenario = Worst or Second-Worst})_i$

$+ \beta_{accom2} * (1, \text{if Accommodation Level = 2})_i * (1, \text{if choice scenario = Best or Second-Best})_i$

$- \beta_{accom2} * (1, \text{if Accommodation Level = 2})_i * (1, \text{if choice scenario = Worst or Second-Worst})_i$

$+ \beta_{accom3} * (1, \text{if Accommodation Level = 3})_i * (1, \text{if choice scenario = Best or Second-Best})_i$

$- \beta_{accom3} * (1, \text{if Accommodation Level = 3})_i * (1, \text{if choice scenario = Worst or Second-Worst})_i$

$+ \beta_{Top\ (best)} * (1, \text{if Accommodation appeared in the first row})_i * (1, \text{if choice scenario = Best or Second-Best})_i$

$+ \beta_{Second\ Top\ (best)} * (1, \text{if Accommodation appeared in the second row})_i * (1, \text{if choice scenario = Best or Second-Best})_i$

$+ \beta_{Second\ Bottom\ (best)} * (1, \text{if Accommodation appeared in the eighth row})_i * (1, \text{if choice scenario = Best or Second-Best})_i$

$+ \beta_{Bottom\ (best)} * (1, \text{if Accommodation appeared in the ninth row})_i * (1, \text{if choice scenario = Best or Second-Best})_i$

$+ \beta_{Top\ (worst)} * (1, \text{if Accommodation appeared in the first row})_i * (1, \text{if choice scenario = Best or Second-Worst})_i$

$+ \beta_{Second\ Top\ (worst)} * (1, \text{if Accommodation appeared in the second row})_i * (1, \text{if choice scenario = Best or Second-Worst})_i$

$+ \beta_{Second\ Bottom\ (worst)} * (1, \text{if Accommodation appeared in the eighth row})_i * (1, \text{if choice scenario = Best or Second-Worst})_i$

$+ \beta_{Bottom\ (worst)} * (1, \text{if Accommodation appeared in the ninth row})_i * (1, \text{if choice scenario = Best or Second-Worst})_i$

$+ \varepsilon_i$

where ε_i is the error term that captures the unobserved heterogeneity due to differences among observations.

One of the main concepts in the field of random utility modelling is the notion of scale of the utility. In particular, and looking at the case of an MNL model, the probability of choosing alternative i out of J alternatives is given by:

$$P(i)=\exp(\mu \cdot V_i)/[\Sigma_j \exp(\mu \cdot V_j)]$$

where: V_i gives the modelled utility of alternative i (e.g. $\beta' x_i$), and
μ is the scale, which is inversely proportional to the variance of the error terms, where the variance of the Gumbel distribution is given by $\pi^2/(6 \cdot \mu^2)$ (Ben-Akiva and Lerman, 1985, pp 104-105).

As the scale increases (variance decreases), respondents become more sensitive to the explanatory variables and the choice process becomes more deterministic. In practice, an increase in scale equates to higher marginal utility coefficients. It is not possible to identify the scale separately from the coefficients, and therefore in estimation the scale is arbitrarily set to 1. So, an increase in scale can equally well be achieved by an increase in μ or in β. However, when dealing with multiple data sources, or multiple types of responses, it is possible to estimate separate scales for different segments of the estimation data (Ben-Akiva and Morikawa, 1990). In the present context, on the basis of past experience we expect that the scale is highest (lowest error variance) for the first choice (i.e. best level) and is lower for the remaining three decisions in each best-worst scaling exercise. Early results showed this to be true, but with little or no difference between the scale for best and second best choice, and between the worst and second worst choice, meaning that we can set μ_{best} to 1 (required for identification reasons) and estimate μ_{worst}, where the former is associated with the best and second best choices, and the latter with the worst and second worst choices.

[9] The "zero-otherwise" statement was specified in all "if-statements" of the utility functions but has been omitted in above boxes to save space in the document.

Initial results also showed that the model scale for those respondents who received the excessively randomised version of the survey was lower (higher variance in the error term), and therefore we incorporated a separate scale, $\mu_{excessive}$, for those respondents who received the version of the survey with excessive randomisation.

The question now arises as to how to deal with these two types of scale differences, which can occur jointly. A possibility would be to use four separate scale parameters, namely $\mu_{best\text{-}standard}$ (set to 1), $\mu_{best\text{-}excessive}$, $\mu_{worst\text{-}standard}$, and $\mu_{worst\text{-}excessive}$. This specification allows for a differential scale impact of the two types of randomisation on best and worst choices, and early results showed that this degree of flexibility was not required. Instead, we have estimated two scale effects: worst (relative to best) and excessive (relative to standard) and applied these jointly, as necessary[10]. Here, with the expectation that the scale for worst and second worst choices is lower than for best and second best choices, and that the scale for observations from the excessively randomised survey is similarly lower, we would have that $\mu_{worst} < 1$ and $\mu_{excessive} < 1$, which would mean even lower scale for worst and second worst choices in the excessively randomised survey.

The next issue is the repeated choice nature of the data. In the most basic specification, the repeated choice nature of the data is not taken into account, and each observation is treated independently. This assumption is clearly incorrect and three possible cases of correlation arise. In possibly the most obvious extension, the model recognises the fact that the set of 36 (4 choices x 9 choice scenarios) or 32 (4 choices x 8 choice scenarios) for a given individual are in fact correlated. An extension of this comes when we additionally recognise the special relationship between the four choices made during each best-worst scaling exercise, meaning that the choices for a given individual are explicitly recognised as a set of eight or nine sequences of four choices. A third approach, which is a simplification of the second case, which was useful for testing purposes as detailed later on, ignores the relationship between the eight or nine sets of best-worst scaling exercises, and only accounts for the relationship between the four choices within each best-worst task.

Current there is no clear guidance amongst choice modellers on how the repeated choice nature of panel data should be taken into account. A number of applications make use of a post-estimation correction approaches, generally leading to an upwards correction in the standard errors. However, these methods, especially resampling approaches such as Jackknife or Bootstrap, are somewhat arbitrary and sensitive to assumptions made during specification. Additionally, in the case where individual model runs already take several hours, as is the case here due to the multiplicative scale approach, resampling approaches can be computationally very expensive. Recently, there has been a trend to make use of random coefficients models, primarily Mixed Logit, where the panel nature of the data is taken into account by allowing for variation in coefficients across respondents while keeping the coefficients constant across choices for the same respondent. In many studies, this one included, taste heterogeneity is however not the main topic of investigation, and indeed, analysts may only be interested in point estimates for average values of the coefficients, which causes complications when relying on random coefficients models.

[10] That is, the scale for a given observation was estimated as $\mu=(1\cdot\delta_{best} + \mu_{worst}\cdot\delta_{worst})\cdot(1\cdot\delta_{standard} + \mu_{excessive}\cdot\delta_{excessive})$, where δ_{best} was set to 1 for best and second best choices, δ_{worst} was set to 1 for worst and second worst choices, and $\delta_{standard}$ and $\delta_{excessive}$ were set to 1 for the standard and excessively randomised surveys respectively.

Several studies have attempted to avoid this issue by relying on an error components approach to accommodate correlation across choices without introducing a representation of taste heterogeneity, but issues with specification and identification arise, alongside a very significant rise in estimation cost.

In the present context, an error components approach would be possible, with integration[11] being carried out at two different levels, namely at the level of an individual and at the level of individual best-worst scaling exercises, i.e. having:

$$L=\prod_n \int \prod_s \int \prod_t P_{nst},$$

where: n is the index over respondents,
s is the index of best-worst scaling exercises, and
t is the index for individual choices within an exercise.

The fact that integration is carried out in two places allows the two types of correlation to be accommodated, but also leads to an almost insurmountable rise in estimation cost, especially when taking into account the fact that a very large number of error components (and hence dimensions of the integral) would be required given the high number of alternatives (individual attribute-level combinations). We therefore reject the error-component approach in this context.

The solution put forward here is to limit ourselves to the outer panel, i.e. treating the 32 or 36 choices for one respondent as a block, but without special treatment for the four choices within a given best-worst scaling experiment. This assumption is however only justified if we can first establish that this outer panel is the main (and ideally only) reason for correlation across choices.

We make the common assumption that the estimates from the naïve model are consistent and that only the standard errors need to be corrected. The correction approach used to accommodate the impact of the panel nature of the data on the standard errors is the sandwich estimator (see e.g. Train, 2003, page 205), which allows for a correction of the standard errors due to some kinds of misspecification. In particular, the sandwich estimator is defined as:

$$S = (-H)^{-1} B (-H)^{-1}$$

where: H is the Hessian, i.e. the matrix of second derivatives, and
B is the BHHH matrix, which is given by the covariance matrix of the first derivatives, calculated over the observations on which the likelihood is calculated.

This differs from the classical covariance matrix, which is given by $(-H)^{-1}$, in the case of misspecification, i.e. when $-H \neq B$. If the model is specified correctly, the two approaches give the same covariance matrix. As Train (2003, page 205) states, 'this [covariance] matrix obtained is called "the robust covariance matrix" since it is valid whether or not the model is correctly specified'.

[11] The choice probabilities are no longer closed-form expressions under this model formulation, so must be estimated through simulation.

Misspecification can obviously be caused by a multitude of factors, linked to the specific assumptions made in model specification. As a first step, a base model was thus estimated with no recognition of the panel nature of the data, and the classical covariance matrix was compared to the robust covariance matrix. Here, the standard errors with the robust specification were slightly higher, but on average only by around 6%, indicating some minor misspecification.

As mentioned above, the specific interest in using the sandwich estimator in the present study was to address the potential downwards bias in the standard errors caused by not accounting for the repeated choice nature of the data. We now return to the above statement for S, which is a function of H and B. When estimating a model with panel data, it is possible to base B on sequences of choices instead of individual choices, leading to different results for S, where this is not the case for the classical covariance matrix which is identical whether individual choices or sequences of choices are used.

Three different models were estimated in the exploratory work for this study. In the first model, we made use of individual choice probabilities in the calculation of S. This was followed by a model in which we made use of the full set of 32 or 36 choices for each respondent, which was observed to lead to a noticeable (further) upwards correction of the standard errors, by on average of 27% across all parameters (comparing the robust standard errors for the base model with those for the model taking into account the panel nature of the data). In contrast, when working with sets of four choices, i.e. looking at individual best-worst scaling exercises, we observed an upwards correction of only 3%. This suggests very little or no effects for the small panel, especially when considering that recognising only the small panel potentially captures some of the effects of the large panel (by still grouping together some of the choices for a given respondent), meaning that the effects for the small panel are possibly even smaller than observed. On this basis, the assumption to focus solely on the large panel seems justified, while conceding that a further very small upwards correction in the standard errors would be obtained by accommodating the small panel.

The preliminary MNL models, with scaling in only single dimensions, were estimated and developed using ALOGIT (2005), and the scaled MNL models with the robust standard errors estimated using the sandwich estimator (reported throughout this report) were estimated using BIOGEME (Bierlaire, 2003).

4.1 Models using domain-level and position information

Table 26 and Table 27 present the domain-level weights estimated for the LLS and the OSCA experiments. The goodness-of-fit measures as indicated by the values of Rho^2 showed that that both models performed relatively well with values between 0.197 in the LLS model and 0.221 in OSCA model. Moreover, the estimated domain-level weights in both models were statistically significant with relatively high t-ratios and the expected signs. The ranking of the estimated weights within each domain was also in agreement with *a priori* expectations since weights at the top levels (representing the better quality of life levels) were higher than weights in the middle-levels followed by the third (and fourth in the OSCA experiment) level of each domain. Finally, the estimated (mean values) weights within a particular domain were significantly different from each other at 95% confidence interval in both models.

It should be noted that all domain levels are estimated relative to a single domain level in the experiment. For these models we have selected the domain level with the lowest utility to be the base level, i.e. the responses of the respondents show that *"I have no control over my daily life"* was viewed as being the worst domain level of all those presented and all other domain levels are valued positively when compared to this.

Table 26: Estimated parameters in LLS model (3-level domains)

Domain Level	Parameter value	Robust t-ratio
Accommodation cleanliness and comfort		
1. My home is as clean and comfortable as I want	4.38	24.94
2. My home is less clean and comfortable than I want	2.47	17.09
3. My home is not at all as clean or comfortable as I want	1.76	13.55
Safety		
1. I feel as safe as I want	4.71	24.8
2. Sometimes I do not feel as safe as I want	1.71	13.8
3. I never feel as safe as I want	1.14	9.49
Food and nutrition		
1. I eat the meals I like when I want	4.16	24.65
2. I don't always eat the right meals I want, but I don't think there is a risk to my health	2.59	18.29
3. I don't always eat the right meals I want, and I think there is a risk to my health	1.96	14.96
Personal cleanliness and comfort		
1. I feel clean and wear what I want	4.54	24.88
2. I sometimes feel less clean than I want or sometimes can't wear what I want	1.87	15.04
3. I feel much less clean than I want, with poor personal hygiene	1.09	9.14
Control over daily life		
1. I have as much control over my daily life as I want	5.18	26.12
2. Sometimes I don't feel I have as much control over my daily life as I want	1.50	11.42
3. I have no control over my daily life	0.00	n/a
Social participation and involvement		
1. My social situation and relationships are as good as I want	4.67	25.14
2. Sometimes I feel my social situation and relationships are not as good as I want	2.36	17.64
3. I feel socially isolated and often feel lonely	0.76	6.80
Dignity		
1. I would be treated by other people with the dignity and respect that I want	4.25	23.32
2. Sometimes I would not be treated by other people with the dignity and respect that I want	1.63	11.70
3. I would never be treated with the dignity and respect that I want	1.18	8.77
Occupation and employment		
1. I do the activities I want to do	4.50	24.39
2. I do some of the activities I want to do	3.95	23.28
3. I don't do any of the activities I want to do	1.69	13.38
Anxiety		
1. I feel free from worry and concerns on a day-to-day basis	4.69	24.65
2. I sometimes feel worried and concerned	1.88	14.09
3. I feel very worried and concerned on a daily basis	1.24	10.43
Domain position (in "best" or "second best" choices)		
Top B (1 for domain that appeared at the top, 0 otherwise)	0.26	4.89
Second Top B (1 for domain that appeared second from the top, 0 otherwise)	0.21	4.15
Second Bottom B (1 for domain that appeared second from the bottom, 0 otherwise)	*0.00*	*-0.01*
Bottom B (1 for domain that appeared at the bottom, 0 otherwise)	-0.16	-2.94
Domain position (in "worst" or "second worst" choices)		
Top W (1 for domain that appeared at the top, 0 otherwise)	0.14	2.01
Second Top W (1 for domain that appeared second from the top, 0 otherwise)	*0.06*	*0.81*
Second Bottom W (1 for domain that appeared second from the bottom, 0 otherwise)	0.17	2.65
Bottom W (1 for domain that appeared at the bottom, 0 otherwise)	0.22	2.79
Scale parameters		
μ_{worst}: scale parameter if data relates to a "worst" or "second-worst" choice	0.75	8.52
$\mu_{excessive}$: scale parameter if data relates to case with "excessive" randomisation	0.75	5.39
Model diagnostics		
Number of observations (Hessian estimated across all choices)		15840
Number of individuals (BHHH matrix estimated on sequence of choices from each individual)		440
D.O.F.		36
Final log likelihood		-25402.20
$Rho^2(0)$		0.200
Adjusted $Rho^2(0)$		0.198

Table 27: Estimated parameters in OSCA model (4-level domains)

Domain Level	Parameter value	Robust t-ratio
Accommodation cleanliness and comfort		
1. My home is as clean and comfortable as I want	4.14	22.64
2. My home is adequately clean and comfortable	3.90	22.54
3. My home is less than adequately clean or comfortable	2.05	15.14
4. My home is not at all clean or comfortable	1.59	13.94
Safety		
1. I feel as safe as I want	4.49	22.80
2. Generally I feel adequately safe, but not as safe as I would like	2.46	17.05
3. I feel less than adequately safe	1.58	13.42
4. I don't feel at all safe	0.56	5.15
Food and nutrition		
1. I get all the food and drink I like when I want	4.15	22.18
2. I get food and drink adequate for my needs	3.96	22.33
3. I don't get all the food and drink I need, but I don't think there is a risk to my health	1.87	14.75
4. I don't get all the food and drink I need, and I think there is a risk to my health	1.20	10.80
Personal cleanliness and comfort		
1. I feel clean and am able to present myself the way I like	4.62	22.68
2. I feel adequately clean and presentable	3.93	21.63
3. I feel less than adequately clean or presentable	1.31	11.16
4. I don't feel at all clean or presentable	1.05	9.38
Control over daily life		
1. I have as much control over my daily life as I want	5.11	23.09
2. I have adequate control over my daily life	4.71	22.78
3. I have some control over my daily life but not enough	2.91	17.18
4. I have no control over my daily life	0.00	n/a
Social participation and involvement		
1. I have as much social contact as I want with people I like	4.42	22.17
2. I have adequate social contact with people	3.86	21.70
3. I have some social contact with people, but not enough	2.67	17.11
4. I have little social contact with people and feel socially isolated	1.32	11.75
Dignity		
1. The way I'm helped and treated makes me think and feel better about myself	4.16	20.80
2. The way I'm helped and treated does not affect the way I think or feel about myself	3.16	18.95
3. The way I'm helped and treated sometimes undermines the way I think and feel about myself	1.74	13.57
4. The way I'm helped and treated completely undermines the way I think and feel about myself	1.44	11.82
Occupation and employment		
1. I'm able to spend my time as I want, doing things I value or enjoy	4.88	22.50
2. I'm able do enough of the things I value or enjoy with my time	4.67	21.94
3. I do some of the things I value or enjoy with my time but not enough	2.80	17.63
4. I don't do anything I value or enjoy with my time	0.97	9.65
Domain position (in "best" or "second best" choices)		
Top B (1 for domain that appeared at the top, 0 otherwise)	0.30	5.37
Second Top B (1 for domain that appeared second from the top, 0 otherwise)	0.16	3.21
Second Bottom B (1 for domain that appeared second from the bottom, 0 otherwise)	-0.11	-2.05
Bottom B (1 for domain that appeared at the bottom, 0 otherwise)	*0.01*	*0.23*
Domain position (in "worst" or "second worst" choices)		
Top W (1 for domain that appeared at the top, 0 otherwise)	*-0.04*	*-0.60*
Second Top W (1 for domain that appeared second from the top, 0 otherwise)	*-0.06*	*-1.03*
Second Bottom W (1 for domain that appeared second from the bottom, 0 otherwise)	*0.03*	*0.56*
Bottom W (1 for domain that appeared at the bottom, 0 otherwise)	0.12	1.96
Scale parameters		
μ_{worst}: scale parameter if data relates to a "worst" or "second-worst" choice	0.85	5.00
$\mu_{excessive}$: scale parameter if data relates to case with "excessive" randomisation	0.79	4.21
Model diagnostics		
Number of observations (Hessian estimated across all choices)		14496
Number of individuals (BHHH matrix estimated on sequence of choices from each individual)		453
D.O.F.		41
Final log likelihood		-20975.22
$Rho^2(0)$		0.221
Adjusted $Rho^2(0)$		0.219

4.1.1 Discussion of model based on OSCA measure

It is informative to review the rank-order of the estimated weights as this reveals the relative importance of each of the domain levels. When considering the LLS measure with the 3-level version of domains and as shown in Figure 4, the highest weight-value was estimated for the domain *Control over daily life* at its top level ("I have as much control over my daily life as I want"*)*. Following the control domain, respondents' then prefer the top levels of *Safety, Social Participation, Anxiety, Personal Cleanliness and Comfort, Occupation and Employment, Accommodation, Cleanliness and Comfort, Dignity* and *Food and Nutrition*. The aforementioned ranking of domain levels means that if respondents were presented with a scenario in which all domains were at their top levels then they were more likely to prefer to have the best level of *Control over daily life* over *Food and Nutrition* or obtain the best level of *Safety* over *Personal, Cleanliness and Comfort*.

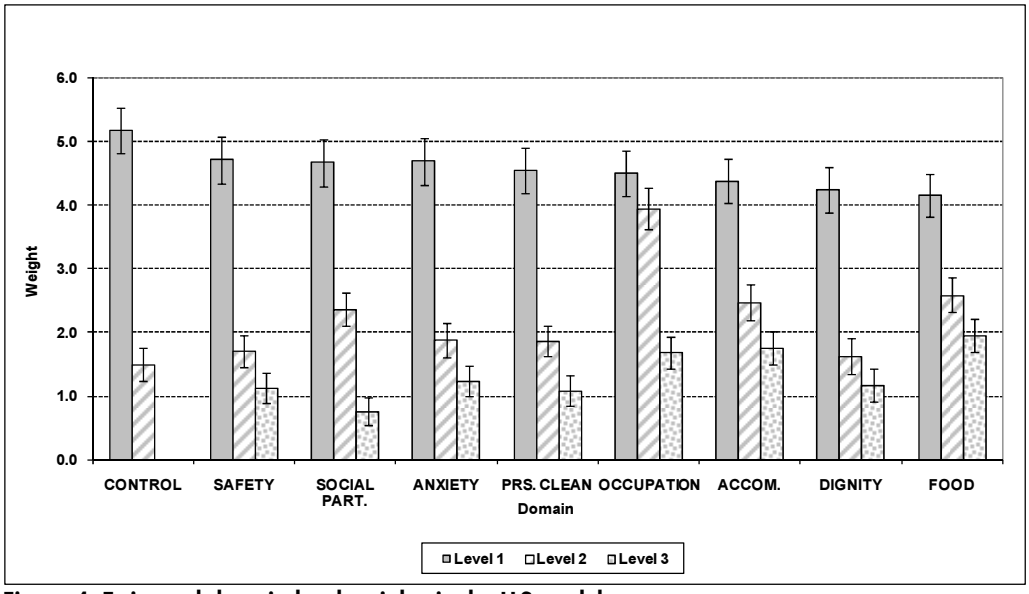

Figure 4: Estimated domain-level weights in the LLS model

The bars presented on the domain level weights are the 95% confidence intervals of the estimates.

When comparing the domain-weights in the middle-level of domains, respondents placed the highest valuation on the domain *Occupation and Employment* and this is closer in magnitude to the top levels of the rest of the domains than the other middle-level values. This is followed by *Food and Nutrition, Accommodation, Cleanliness and Comfort, Social participation, Personal Cleanliness and Comfort, Anxiety, Personal Safety, Dignity* and *Control over daily life.* Finally, comparing the low levels (e.g. level 3) of domains, participants were more likely to accept a situation where *Food and Nutrition* were at its lowest (worst) level than having *no control over their daily life*. The latter domain-level was the least preferred situation over all domains presented in LLS experiment. The descending order between *Food and Nutrition* and *Control over daily life* domains at the lowest levels included: *Accommodation, Personal Cleanliness and Comfort*, *Occupation and Employment*, *Anxiety*, *Dignity, Personal Cleanliness and Comfort, Safety and Social Participation.* It is noteworthy that the highest level of the *Control over daily life* domain was viewed as the best of all the

domain levels, whilst the lowest value of the same domain was viewed as the worst of all the domain levels.

As mentioned previously, the experiments were constructed in such a way that the order of the domains within the best-worst choice task were randomised between respondents (and in some cases within respondents), in order to control for the impact that the position of the domain may have had on the likelihood of it being selected as the best or worst option. This approach ensured that if the position of the domain was to induce any bias, all of the domains would suffer this bias equally and there would be no systematic over or underestimation of the value of any given domain. In addition, by recording the position of each domain in any given choice it was possible to include this information within the parameterisation of the model and quantify the increase (or decrease) in likelihood of a respondent choosing a given domain due to its position independent of the estimation of the value placed on the level at which it was presented. The final models contain eight position dummy-coded variables reflecting the position of a domain-level when that was chosen, contingent upon whether the respondent was asked to choose a "best" or "worst" option. In the LLS model, we see that for the "best" choices, the variables *Top*, *Second Top* and *Bottom* are statistically significant, which implies that participants were more likely to choose a domain-level that appeared first or second in the domain list and less likely to choose the bottom option when asked to indicate the "best" or "second best option". Similarly, we see for the "worst" choices, the variables *Top*, *Second Bottom* and *Bottom* are statistically significant, which implies that participants were more likely to choose a domain-level that appeared first or in the bottom two places in the domain list when asked to indicate the "best" or "second best option". The order implied by the coefficient values is intuitive; however, the overall effect of these variables as seen from the values of their coefficients was considerably smaller than the main effect of the domain-level weights, and the significance of these coefficients is also significantly lower.

Finally, the scale parameters μ_{worst}, $\mu_{excessive}$ were statistically significant (with respect to 1.0). The scale parameter μ_{worst} suggests that the variance (error) between the "best" (including best and second-best) and "worst" (including worst and second best) responses was statistically significant, with the data from the "worst" responses generating a less deterministic choice process than the "best" responses. Similarly, the scale parameter $\mu_{excessive}$ suggests that the data from those facing excessive randomisation have a less deterministic choice process than those facing the intended randomisation of the domain order, which was held constant within a given respondent.

4.1.2 Discussion of model based on OSCA measure

The 4-level-domain experiment was based around the OSCA measure. Figure 5 shows that the highest weight-value was estimated for the top-level of the domain *Control over daily life*, which is consistent with the LLS model. The rest of the domain ranking in the OSCA model appeared slightly different than the LLS model. Specifically, the next highest weight was placed on the top levels of the domains *Occupation and Employment* followed by *Personal Cleanliness and Comfort, Safety, Social Participation, Accommodation, Cleanliness and Comfort, Food and Nutrition* and *Dignity.*

If respondents were presented with a scenario with all domains at the second-top level then they would be more likely to prefer having more *Control over daily life* and how they use

their time ('I'm able do enough of the things I value or enjoy with my time') followed by *Food and Nutrition, Personal Cleanliness and Comfort, Accommodation, Social participation, Dignity* and finally, *Safety*. Looking at the ranking for domain weights at the third level, respondents placed highest values on *Control over daily life* and *Occupation and Employment* followed by *Social Participation*, *Accommodation*, *Food and Nutrition*, *Dignity*, *Safety* and *Personal Cleanliness and Comfort*. Finally, if faced with scenarios where the domains are at their lowest possible levels, respondents would be more likely to accept situations that involved the lowest levels of *Accommodation* and *Dignity* over *Social Participation*, *Food and Nutrition*, *Personal Cleanliness* and *Comfort*, *Occupation and Employment*, and *Safety* whereas a situation that involved the lowest level of control ("I have no control over my daily life") was the defined as the most negative of all the domain levels presented.

It is notable that for the majority of domains the top two levels are valued quite similarly, suggesting that respondents only place a small amount of additional value on moving from all needs being met to a more aspirational quality of life. In the case of *Control over daily life*, *Accommodation* and *Food and Nutrition* the top two levels are not statistically different from each other, as shown by the 95% confidence intervals in Figure 5. The one significant exception is the safety domain, in which there is a much larger step in utility between the second and top level – corresponding in a move from "Generally I feel adequately safe, but not as safe as I would like" to "I feel as safe as I want".

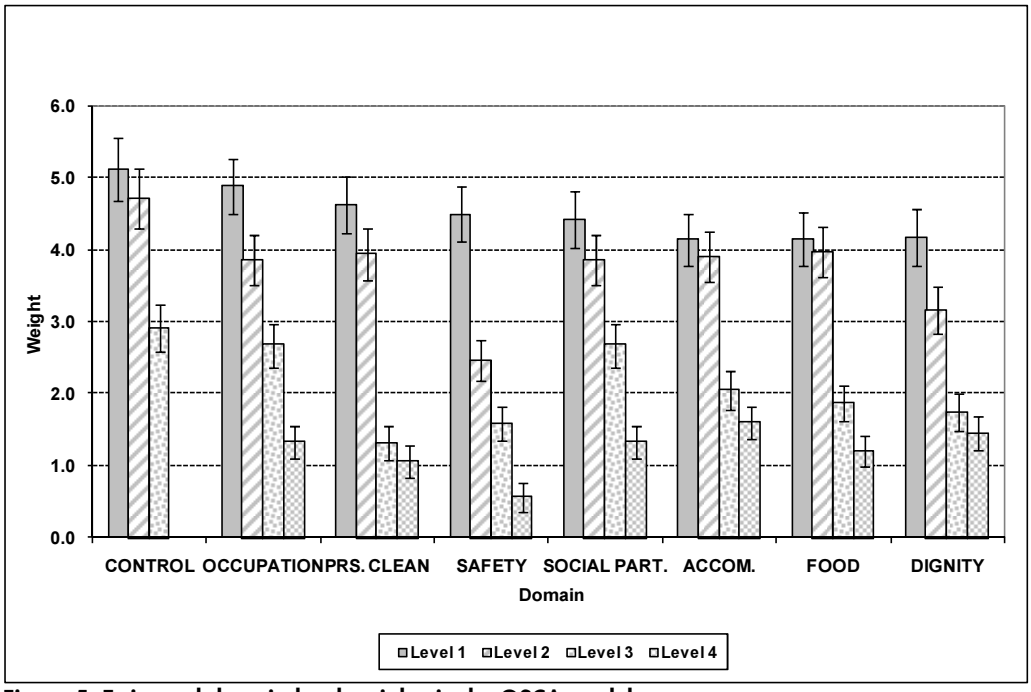

Figure 5: Estimated domain-level weights in the OSCA model

As with the LLS model, it was possible within the OSCA model to control for the effect of the position of each domain in the best-worst choices in the model. The model presented in Table 27 indicates that when asked to choose a "best" option, participants were more likely to choose a domain-level that appeared at the top or second from top in the list and were less likely to choose the domain that was presented second from the bottom. Within the "worst" choices respondents were more likely to choose the domain presented last in

the list. However, as in the LLS model, these effects are relatively small compared to the values placed on the domain levels.

The scale parameters μ_{worst}, $\mu_{excessive}$ were statistically significant (with respect to 1.0). The scale parameter μ_{worst} suggests that the data from the "worst" responses generated a less deterministic choice process than the "best" responses. Similarly, the scale parameter $\mu_{excessive}$ suggests that the data from those facing excessive randomisation have a less deterministic choice process than those facing the intended randomisation of the domain order, which was held constant within a given respondent. These effects are captured and controlled for in the model estimation to prevent these scale differences biasing the estimates of the domain weights.

Overall, comparing the patterns in "best" and "worst" preferences for the highest and lowest levels of the domains across both experiments it is shown that *Control over daily life*, *Safety, Personal Cleanliness, Safety and Comfort* were at the top of respondents' priorities. On the other hand, respondents placed lower values on the described levels of *Food and Nutrition* and *Accommodation and Cleanliness* in both experiments.

4.2 Models using domain-level, position information and individuals' characteristics

The best-worst models were further developed to capture observable heterogeneity in the preferences of the sample as a result of socio-economic and demographic characteristics of the individuals. In developing these models the following characteristics of the respondents were examined:

- Gender;
- Age;
- Level of education;
- Marital status;
- Number of children;
- Employment status;
- Household income;
- Type of area;
- Ethnicity;
- Personal experience of social care (self); and
- Personal experience of social care (relatives).

A series of tests were run to compare the predicted probabilities of choosing each domain level against the observed frequencies of these being chosen within the data across each of these different respondent characteristics. Where these tests identified that the model led to a significant discrepancy between the observed and predicted choices, the specification of the utility functions were developed to take explicit account of the apparent differences in preferences. The taste heterogeneity was introduced in to the model in two different ways:

- Additional dummy-coded variables on the utility of a given domain, representing a systematic difference in how identified groups valued the given domain relative to other domains (i.e. an effect that applies across all levels of that domain); and
- Additional covariates on an individual domain level, capturing situations where identified groups placed a significantly higher (or lower) value on one particular domain level.

The models following this further model development are presented in Table 28 and Table 29 for the 3- and 4-level of domains, respectively.

The goodness-of-fit of the models was higher in both models indicating an (as expected) improvement over the generic models that included domain-level and position information in Section 4.1. Moreover, log-likelihood ratio tests[12] showed that the difference in the log-likelihood function between the generic (domain, and position variables only) and enhanced models (domain, position and socio-economic and demographic variables)[13] were statistically significant for the models of both the LLS 3-level and OSCA 4-level versions of the domains.

Table 28: Estimated parameters in the full-version LLS model (3-level domains)

Domain Level	Parameter value	Robust t-ratio
Accommodation cleanliness and comfort		
1. My home is as clean and comfortable as I want	4.11	24.89
2. My home is less clean and comfortable than I want	2.12	14.97
3. My home is not at all as clean or comfortable as I want * married	1.48	9.23
3. My home is not at all as clean or comfortable as I want * non-married	1.19	7.81
Safety		
1. I feel as safe as I want * annual household income ≥ £25,000	4.70	24.35
1. I feel as safe as I want * annual household income < £25,000 or unknown	4.14	23.49
2. Sometimes I do not feel as safe as I want	1.34	11.43
3. I never feel as safe as I want * full-time worker	0.57	3.48
3. I never feel as safe as I want * non full-time worker	0.92	6.61
All levels. Individual lives in a London borough or another city / large town or suburbs of a city / large town	0.31	3.81
Food and nutrition		
1. I eat the meals I like when I want	4.08	24.86
2. I don't always eat the right meals I want, but I don't think there is a risk to my health	2.27	15.69
3. I don't always eat the right meals I want, and I think there is a risk to my health	1.49	10.45
All levels. Female	-0.43	-5.11
Personal cleanliness and comfort		
1. I feel clean and wear what I want	4.42	24.99
2. I sometimes feel less clean than I want or sometimes can't wear what I want	1.39	10.98
3. I feel much less clean than I want, with poor personal hygiene	0.49	3.75
All levels. Male	-0.33	-3.41
Control over daily life		
1. I have as much control over my daily life as I want * annual household income ≥ £25,000	4.73	23.23
1. I have as much control over my daily life as I want * annual household income < £25,000 or unknown	4.21	21.84
2. Sometimes I don't feel I have as much control over my daily life as I want	1.34	11.70
3. I have no control over my daily life	0.00	n/a
All levels. Age ≥ 50 year of age	0.57	6.57
All levels. Full time or part time worker	0.38	4.74
All levels. Been in a situation where someone close has not able to care for themselves	0.31	3.69
Social participation and involvement		
1. My social situation and relationships are as good as I want * full-time worker	4.62	24.85

[12] Log-likelihood ratio test $=\chi^2(d.f._{enhanced} - d.f._{plain}) = -2 * (LL_{enhanced} - LL_{plain})$

[13] 3-level model $\chi^2(17)=571>33.41$ (critical value at 99% confidence level); 4-level model $\chi^2(12)=342.5>26.22$ (critical value at 99% confidence level)

1. My social situation and relationships are as good as I want * non full-time worker	4.27	23.32
2. Sometimes I feel my social situation and relationships are not as good as I want	1.99	15.40
3. I feel socially isolated and often feel lonely	0.26	2.21
Dignity		
1. I would be treated by other people with the dignity and respect that I want	3.71	21.80
2. Sometimes I would not be treated by other people with the dignity and respect that I want	1.33	10.77
3. I would never be treated with the dignity and respect that I want	0.93	7.43
All levels. Married	0.26	2.31
All levels. White British	0.33	3.29
All levels. Single	-0.43	-2.54
Occupation and employment		
1. I do the activities I want to do	4.09	23.70
2. I do some of the activities I want to do	3.54	22.66
3. I don't do any of the activities I want to do	1.32	10.97
All levels. Male	0.34	4.34
Anxiety		
1. I feel free from worry and concerns on a day-to-day basis * (non White British who live in London borough or another city / large town or suburbs of a city / large town)	3.78	16.83
1. I feel free from worry and concerns on a day-to-day basis * (White British, or non White British who in a small town, a rural area or village)	4.38	23.46
2. I sometimes feel worried and concerned	1.52	12.24
3. I feel very worried and concerned on a daily basis	0.88	7.48
All levels. Married	0.30	3.60
Domain position (in "best" or "second best" choices)		
Top B (1 for domain that appeared at the top, 0 otherwise)	0.26	4.93
Second Top B (1 for domain that appeared second from the top, 0 otherwise)	0.20	3.92
Second Bottom B (1 for domain that appeared second from the bottom, 0 otherwise)	*0.00*	*-0.02*
Bottom B (1 for domain that appeared at the bottom, 0 otherwise)	-0.17	-3.02
Domain position (in "worst" or "second worst" choices)		
Top W (1 for domain that appeared at the top, 0 otherwise)	*0.14*	*1.93*
Second Top W (1 for domain that appeared second from the top, 0 otherwise)	*0.04*	*0.53*
Second Bottom W (1 for domain that appeared second from the bottom, 0 otherwise)	0.18	2.66
Bottom W (1 for domain that appeared at the bottom, 0 otherwise)	0.24	2.95
Scale parameters		
μ_{worst}: scale parameter if data relates to a "worst" or "second-worst" choice	0.72	10.18
$\mu_{excessive}$: scale parameter if data relates to case with "excessive" randomisation	0.76	5.22
Model diagnostics		
Number of observations (Hessian estimated across all choices)		15840
Number of individuals (BHHH matrix estimated on sequence of choices from each individual)		440
D.O.F.		53
Final log likelihood		-25116.64
$Rho^2(0)$		0.209
Adjusted $Rho^2(0)$		0.207

The model for the LLS, three level version of the domains, showed a number of differences that were identified between respondents.

A difference was observable in the value placed on the lowest level of the *Accommodation Cleanliness and Comfort* domain "My home is not at all as clean or comfortable as I want". In this case we see that the utility associated with this level is lower for those that are not married than those that are married. This implies that the unmarried respondents place a greater value on moving from this low level of *Accommodation Cleanliness and Comfort* to the higher levels than their married counterparts.

Those who were living in a city or large town or a suburb, placed a higher value on *Safety* than those living in less urban areas – this difference applied across all domain levels, reflecting a higher value placed on the domain rather than just specific levels. Additional differences in the value placed on the *Safety* domain are observed for the first and third levels, suggesting that those on higher incomes place greater value on achieving a situation where they can state "I feel as safe as I want". We also observe that full-time workers have a

lower utility for the lowest level of this domain "I never feel as safe as I want", which implies that they placed a greater value on moving to the higher levels of this domain.

Also, female respondents placed less value on all levels of *Food and Nutrition* than male respondents, and that male respondents placed less value on all levels of *Personal Cleanliness and Comfort* than female respondents. However, male respondents placed higher value on the *Occupation and Employment* domain than the female respondents in our sample.

The *Control over daily life* domain was valued higher by those over 50 years of age, those that were in employment, and those who had been in a situation where someone close had not been able to care for themselves; these three terms are additive, so for example, an individual over 50 years of age who was in employment would have two of the terms applied. In addition, there was a covariate on the highest level which showed that those on higher incomes (greater or equal to £25k per annum) placed more value on achieving a level where they could state that "I have as much control over my daily life as I want".

It is also interesting to observe that those in full time employment placed greatest utility on obtaining a situation where they can state that "My social situation and relationships are as good as I want".

The *Dignity* domain was valued more by those that were married and less by those that reported to be single. We also observe that white British respondents placed higher value on the *Dignity* domain than other ethnic groups.

Those respondents that were married also placed higher value on *Anxiety* domain than other groups. The highest level of this domain, in which the individual attains a state where they "feel free from worry and concerns on a day-to-day basis", is valued lowest by the non White British respondents that lived in urbanised areas.

The results for the OSCA models are presented in Table 29 below.

Table 29: Estimated parameters in the full-version OSCA model (4-level domains)

Domain Level	Parameter value	Robust t-ratio
Accommodation cleanliness and comfort		
1. My home is as clean and comfortable as I want	3.23	22.23
2. My home is adequately clean and comfortable	2.97	21.54
3. My home is less than adequately clean or comfortable	1.04	10.36
4. My home is not at all clean or comfortable	0.61	7.32
Safety		
1. I feel as safe as I want * white British	3.36	21.64
1. I feel as safe as I want * non-white British	3.88	17.40
2. Generally I feel adequately safe, but not as safe as I would like	1.49	14.02
3. I feel less than adequately safe	0.67	7.68
4. I don't feel at all safe	-0.26	-2.99
All levels. Individual lives in a London borough or another city / large town or suburbs of a city / large town	0.28	3.65
Food and nutrition		
1. I get all the food and drink I like when I want	3.37	21.86
2. I get food and drink adequate for my needs	3.15	21.81
3. I don't get all the food and drink I need, but I don't think there is a risk to my health	0.81	8.26
4. I don't get all the food and drink I need, and I think there is a risk to my health	0.06	0.70
All levels. Female	-0.37	-4.60
All levels. Someone close had not able to care for themselves, but now can	0.31	2.24
Personal cleanliness and comfort		
1. I feel clean and am able to present myself the way I like	3.72	22.17
2. I feel adequately clean and presentable	3.01	20.68
3. I feel less than adequately clean or presentable	0.32	3.81
4. I don't feel at all clean or presentable	0.06	0.75

Control over daily life		
1. I have as much control over my daily life as I want	3.57	21.18
2. I have adequate control over my daily life	3.23	21.43
3. I have some control over my daily life but not enough	1.82	0.10
4. I have no control over my daily life	0.00	n/a
All levels. Male	0.24	2.86
All levels. Married	0.43	5.50
All levels. Full time or part time worker	0.30	4.19
All levels. Annual household income < £40,000 per year	0.28	3.80
All levels. Widowed	0.58	3.61
Social participation and involvement		
1. I have as much social contact as I want with people I like	3.51	21.52
2. I have adequate social contact with people	2.94	20.65
3. I have some social contact with people, but not enough	1.69	14.27
4. I have little social contact with people and feel socially isolated	0.03	0.24
Dignity		
1. The way I'm helped and treated makes me think and feel better about myself * married	3.42	18.34
1. The way I'm helped and treated makes me think and feel better about myself * non-married	3.06	16.87
2. The way I'm helped and treated does not affect the way I think or feel about myself	2.18	16.66
3. The way I'm helped and treated sometimes undermines the way I think and feel about myself	0.73	7.95
4. The way I'm helped and treated completely undermines the way I think and feel about myself	0.46	5.10
Occupation and employment		
1. I'm able to spend my time as I want, doing things I value or enjoy	3.76	21.55
2. I'm able do enough of the things I value or enjoy with my time * white British	3.68	20.98
2. I'm able do enough of the things I value or enjoy with my time * non-white British	3.03	13.36
3. I do some of the things I value or enjoy with my time but not enough	1.82	15.92
4. I don't do anything I value or enjoy with my time	0.16	2.12
All levels. Male	0.49	6.71
Domain position (in "best" or "second best" choices)		
Top B (1 for domain that appeared at the top, 0 otherwise)	0.31	5.56
Second Top B (1 for domain that appeared second from the top, 0 otherwise)	0.15	2.92
Second Bottom B (1 for domain that appeared second from the bottom, 0 otherwise)	*-0.10*	*-1.92*
Bottom B (1 for domain that appeared at the bottom, 0 otherwise)	*0.00*	*0.01*
Domain position (in "worst" or "second worst" choices)		
Top W (1 for domain that appeared at the top, 0 otherwise)	*-0.03*	*-0.53*
Second Top W (1 for domain that appeared second from the top, 0 otherwise)	*-0.06*	*-1.14*
Second Bottom W (1 for domain that appeared second from the bottom, 0 otherwise)	*0.02*	*0.40*
Bottom W (1 for domain that appeared at the bottom, 0 otherwise)	*0.11*	*1.91*
Scale parameters		
μ_{worst}: scale parameter if data relates to a "worst" or "second-worst" choice	0.89	3.32
$\mu_{excessive}$: scale parameter if data relates to case with "excessive" randomisation	0.81	3.87
Model diagnostics		
Number of observations (Hessian estimated across all choices)		14496
Number of individuals (BHHH matrix estimated on sequence of choices from each individual)		453
D.O.F.		53
Final log likelihood		-20803.982
$Rho^2(0)$		0.227
Adjusted $Rho^2(0)$		0.225

The model for the OSCA, four level version of the domains, shares a number of findings with the model based on the LLS measure, but also reveals a number of different trends in the data.

The similarities between the models are that:

- Individual that live in a London borough or another city / large town or suburbs of a city / large town place more importance on *Safety* than those living in more rural areas;
- Female respondents place less importance on the *food and nutrition* domain than male respondents;
- The *control over daily life* domain is valued more highly by those that are currently in employment; and

- Male respondents place higher value on the *occupation and employment* domain than the female respondents in our sample.

In the case of the OSCA measure we see that the respondents that were not white British placed higher value on attaining the level of "I feel as safe as I want" than the white British respondents.

We find no significant covariates on either *accommodation cleanliness and comfort*, *personal cleanliness and comfort* or *social participation and involvement*.

In addition to those currently in employment we can estimate significant terms on the *Control over daily life* domain for male, married, lower income (household income <£40k/annum) and widowed respondents. These terms are all additive, and demonstrate that there is considerable heterogeneity in the value placed on this domain.

We see that the white British respondents place higher value on reaching a level of *occupation and employment* at which they are able to state "I'm able do enough of the things I value or enjoy with my time", but as a result also place a lower value on the subsequent transition to the state of "I'm able to spend my time as I want, doing things I value or enjoy" which is valued equally by all ethnic groups.

We also see that those that have had someone close to them that formerly not been able to care for themselves for a period of time places more importance on *food and nutrition* than other respondents.

4.3 Adjusting domain-level weights to reflect population weights

The models that we have developed show that there are a number of areas in which we can detect significant variation in preference weights. A comparison of the sample with the national population (reported in Section 2.1) shows that the sample is not necessarily representative of the population n some of these key dimensions. As a result, the average values reported in the models in Section 4.1 are likely to be subject to some bias.

More appropriate mean values for the preference weights can therefore be obtained by using the results from the models incorporating the detected taste variation report in Section 4.2, and applying the appropriate population proportions to each of the segments to produce revised mean values for each domain level that take in to account the differences that we know exist in how different groups value the domain levels. The weighted average values from these calculations can then be applied with more confidence to situations where the analyst wishes to explain the mean preferences and values of the population. It is these values that we recommend being used in the current ONS Invest to Save funded Toolkit.

The following table presents the LLS measure models prior to the addition of any segmentation in preferences, and the results that come from the models where segmentation has been incorporated and the appropriate population weights have been applied to calculate a revised weighted average. The first set of columns in the table shows the utility weights of each domain level, and the second set of columns shows the relative value placed on improvements within each domain, relative to the lowest level (which is

informative for the ASCOT where the interest is in the value associated in changes in needs met, i.e. moving between levels within each domain).

Table 30: Comparison of LLS models without segmentation, and with segmentation following weighting (3-level domains)

Domain Level	LLS Model Results		Change within domain	
	Model without segmentation	Weighted model with segmentation	Model without segmentation	Weighted model with segmentation
Accommodation cleanliness and comfort				
1. My home is as clean and comfortable as I want	4.38	4.11	2.62	**2.77**
2. My home is less clean and comfortable than I want	2.47	2.12	0.71	**0.78**
3. My home is not at all as clean or comfortable as I want	1.76	1.34	0.00	**0.00**
Safety				
1. I feel as safe as I want	4.71	4.67	3.57	**3.70**
2. Sometimes I do not feel as safe as I want	1.71	1.59	0.57	**0.62**
3. I never feel as safe as I want	1.14	0.98	0.00	**0.00**
Food and nutrition				
1. I eat the meals I like when I want	4.16	3.86	2.20	**2.59**
2. I don't always eat the right meals I want, but I don't think there is a risk to my health	2.59	2.05	0.63	**0.78**
3. I don't always eat the right meals I want, and I think there is a risk to my health	1.96	1.27	0.00	**0.00**
Personal cleanliness and comfort				
1. I feel clean and wear what I want	4.54	4.26	3.45	**3.93**
2. I sometimes feel less clean than I want or sometimes can't wear what I want	1.87	1.23	0.78	**0.90**
3. I feel much less clean than I want, with poor personal hygiene	1.09	0.33	0.00	**0.00**
Control over daily life				
1. I have as much control over my daily life as I want	5.18	5.10	5.18	**4.47**
2. Sometimes I don't feel I have as much control over my daily life as I want	1.50	1.97	1.50	**1.34**
3. I have no control over my daily life	0.00	0.63	0.00	**0.00**
Social participation and involvement				
1. My social situation and relationships are as good as I want	4.67	4.44	3.91	**4.19**
2. Sometimes I feel my social situation and relationships are not as good as I want	2.36	1.99	1.60	**1.73**
3. I feel socially isolated and often feel lonely	0.76	0.26	0.00	**0.00**
Dignity				
1. I would be treated by other people with the dignity and respect that I want	4.25	4.03	3.07	**2.78**
2. Sometimes I would not be treated by other people with the dignity and respect that I want	1.63	1.65	0.45	**0.40**
3. I would never be treated with the dignity and respect that I want	1.18	1.25	0.00	**0.00**
Occupation and employment				
1. I do the activities I want to do	4.50	4.26	2.81	**2.77**
2. I do some of the activities I want to do	3.95	3.71	2.26	**2.22**
3. I don't do any of the activities I want to do	1.69	1.49	0.00	**0.00**
Anxiety				
1. I feel free from worry and concerns on a day-to-day basis	4.69	4.46	3.45	**3.43**
2. I sometimes feel worried and concerned	1.88	1.67	0.64	**0.64**
3. I feel very worried and concerned on a daily basis	1.24	1.03	0.00	**0.00**

Figure 6 provides a visual representation of the data presented in the final column of Table 30. This illustrates the differences in preference weight placed on improvements within each domain, and demonstrates the additional value added to the ASCOT by eliciting the population weights rather than assuming a uniform and equal interval weight across all improvements in needs met, across all domains.

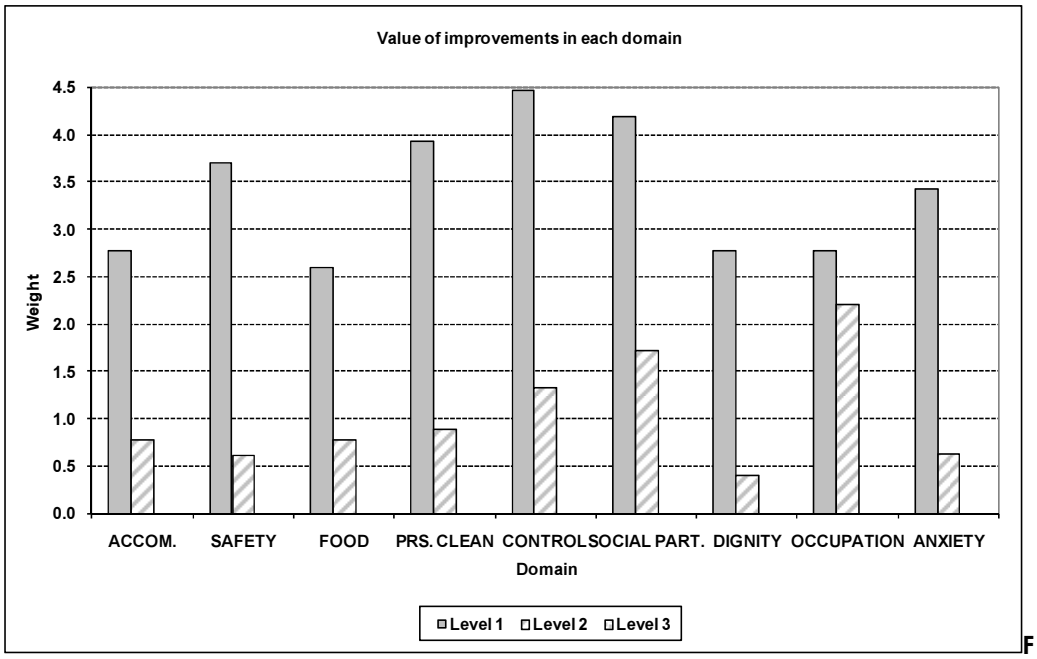

Figure 6: Value placed on improvements in needs met within each domain (LLS 3 level version)

Table 31 and Figure 7 present the same information for the OSCA 4-level measure. It should, however, be noted that the absolute values of the preference weights should not at this stage be compared between the LLS and OSCA measures as these are separate models, which are each estimated with their own scale.

Table 31: Comparison of OSCA models without segmentation, and with segmentation following weighting (4-level domains)

Domain Level	OSCA Model Results		Change within domain	
	Model without segmentation	Weighted model with segmentation	Model without segmentation	Weighted model with segmentation
Accommodation cleanliness and comfort				
1. My home is as clean and comfortable as I want	4.14	3.23	2.55	**2.62**
2. My home is adequately clean and comfortable	3.90	2.97	2.31	**2.36**
3. My home is less than adequately clean or comfortable	2.05	1.04	0.46	**0.43**
4. My home is not at all clean or comfortable	1.59	0.61	0.00	**0.00**
Safety				
1. I feel as safe as I want	4.49	3.65	3.93	**3.69**
2. Generally I feel adequately safe, but not as safe as I would like	2.46	1.72	1.90	**1.75**
3. I feel less than adequately safe	1.58	0.90	1.02	**0.93**
4. I don't feel at all safe	0.56	-0.03	0.00	**0.00**
Food and nutrition				
1. I get all the food and drink I like when I want	4.15	3.21	2.95	**3.31**
2. I get food and drink adequate for my needs	3.96	2.99	2.76	**3.09**
3. I don't get all the food and drink I need, but I don't think there is a risk to my health	1.87	0.66	0.67	**0.75**
4. I don't get all the food and drink I need, and I think there is a risk to my health	1.20	-0.10	0.00	**0.00**
Personal cleanliness and comfort				
1. I feel clean and am able to present myself the way I like	4.62	3.72	3.57	**3.66**
2. I feel adequately clean and presentable	3.93	3.01	2.88	**2.95**
3. I feel less than adequately clean or presentable	1.31	0.32	0.26	**0.26**
4. I don't feel at all clean or presentable	1.05	0.06	0.00	**0.00**
Control over daily life				
1. I have as much control over my daily life as I want	5.11	4.33	5.11	**3.57**
2. I have adequate control over my daily life	4.71	3.99	4.71	**3.23**
3. I have some control over my daily life but not enough	2.91	2.58	2.91	**1.82**
4. I have no control over my daily life	0.00	0.76	0.00	**0.00**
Social participation and involvement				
1. I have as much social contact as I want with people I like	4.42	3.51	3.10	**3.48**
2. I have adequate social contact with people	3.86	2.94	2.54	**2.91**
3. I have some social contact with people, but not enough	2.67	1.69	1.35	**1.66**
4. I have little social contact with people and feel socially isolated	1.32	0.03	0.00	**0.00**
Dignity				
1. The way I'm helped and treated makes me think and feel better about myself	4.16	3.25	2.72	**2.78**
2. The way I'm helped and treated does not affect the way I think or feel about myself	3.16	2.18	1.72	**1.72**
3. The way I'm helped and treated sometimes undermines the way I think and feel about myself	1.74	0.73	0.30	**0.27**
4. The way I'm helped and treated completely undermines the way I think and feel about myself	1.44	0.46	0.00	**0.00**
Occupation and employment				
1. I'm able to spend my time as I want, doing things I value or enjoy	4.88	4.00	3.91	**3.60**
2. I'm able do enough of the things I value or enjoy with my time	4.67	3.83	3.70	**3.43**
3. I do some of the things I value or enjoy with my time but not enough	2.80	2.06	1.83	**1.66**
4. I don't do anything I value or enjoy with my time	0.97	0.40	0.00	**0.00**

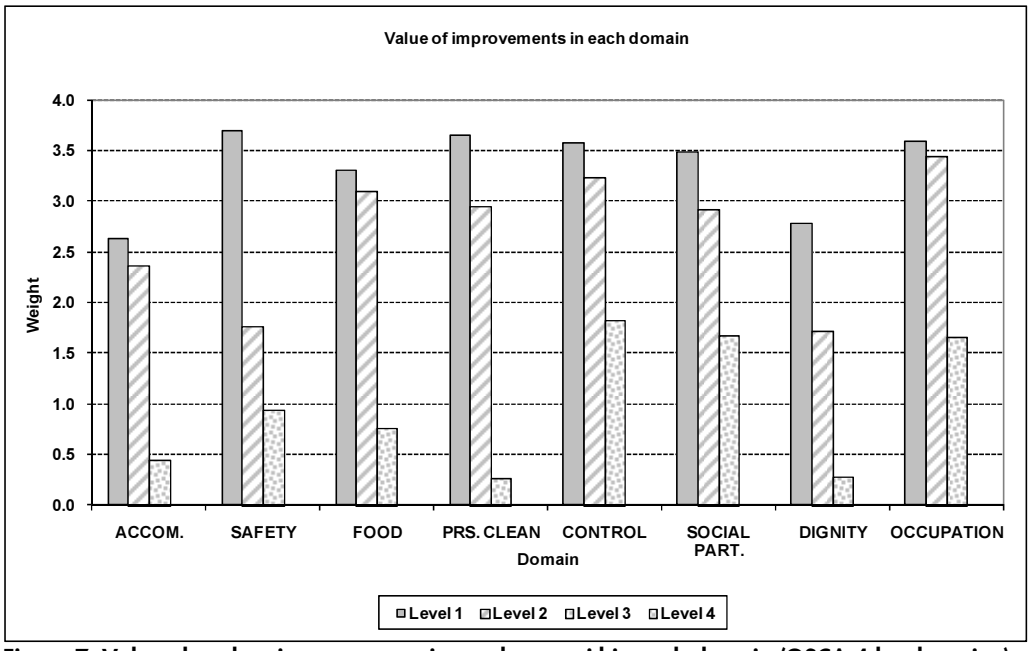

Figure 7: Value placed on improvements in needs met within each domain (OSCA 4 level version)

The following section now provides a comparison of the two measures, with a particular focus on how they may compare if it were assumed that the lowest level of needs on each domain in the LLS 3-level measure were approximately equal to the mean of the lowest two levels of needs of the same domains in the OSCA 4-level measure.

4.4 Comparison of domain weights between the 3-level version and the collapsed 4-level version

One of the research questions for this study is the extent to which the OSCA measure that uses 4-level domains may offer an improvement in the sensitivity of the measure at lower levels of need than the current LLS 3-level domains, and the extent to which the OSCA measure may be used to update the values used within the toolkit at a later date.

To investigate this question a further model has been estimated that collapses the OSCA measure to 3 levels, grouping the highest two levels to produce an average value for these lower levels of need. The resulting model is presented in Table 32.

Table 32: Domain weights when 4-level domains collapse into a 3-level version

Domain Level	Parameter value	Robust t-ratio
Accommodation cleanliness and comfort		
1. My home is as clean and comfortable as I want 2. My home is adequately clean and comfortable	4.01	22.55
3. My home is less than adequately clean or comfortable	2.09	14.91
4. My home is not at all clean or comfortable	1.63	13.63
Safety		
1. I feel as safe as I want 2. Generally I feel adequately safe, but not as safe as I would like	3.55	19.84
3. I feel less than adequately safe	1.60	13.36
4. I don't feel at all safe	0.55	5.01
Food and nutrition		
1. I get all the food and drink I like when I want 2. I get food and drink adequate for my needs	4.06	22.19
3. I don't get all the food and drink I need, but I don't think there is a risk to my health	1.91	14.56
4. I don't get all the food and drink I need, and I think there is a risk to my health	1.24	10.70
Personal cleanliness and comfort		
1. I feel clean and am able to present myself the way I like 2. I feel adequately clean and presentable	4.27	22.21
3. I feel less than adequately clean or presentable	1.32	10.89
4. I don't feel at all clean or presentable	1.08	9.24
Control over daily life		
1. I have as much control over my daily life as I want 2. I have adequate control over my daily life	4.88	22.97
3. I have some control over my daily life but not enough	2.92	16.89
4. I have no control over my daily life	0.00	n/a
Social participation and involvement		
1. I have as much social contact as I want with people I like 2. I have adequate social contact with people	4.15	22.02
3. I have some social contact with people, but not enough	2.70	17.02
4. I have little social contact with people and feel socially isolated	1.36	11.61
Dignity		
1. The way I'm helped and treated makes me think and feel better about myself 2. The way I'm helped and treated does not affect the way I think or feel about myself	3.68	19.79
3. The way I'm helped and treated sometimes undermines the way I think and feel about myself	1.80	13.36
4. The way I'm helped and treated completely undermines the way I think and feel about myself	1.48	11.66
Occupation and employment		
1. I'm able to spend my time as I want, doing things I value or enjoy 2. I'm able do enough of the things I value or enjoy with my time	4.76	22.29
3. I do some of the things I value or enjoy with my time but not enough	2.87	17.08
4. I don't do anything I value or enjoy with my time	1.00	9.59
Domain position (in "best" or "second best" choices)		
Top B (1 for domain that appeared at the top, 0 otherwise)	0.30	5.39
Second Top B (1 for domain that appeared second from the top, 0 otherwise)	0.16	3.21
Second Bottom B (1 for domain that appeared second from the bottom, 0 otherwise)	-0.11	-2.11
Bottom B (1 for domain that appeared at the bottom, 0 otherwise)	*0.01*	*0.22*
Domain position (in "worst" or "second worst" choices)		
Top W (1 for domain that appeared at the top, 0 otherwise)	*-0.03*	*-0.53*
Second Top W (1 for domain that appeared second from the top, 0 otherwise)	*-0.06*	*-0.97*
Second Bottom W (1 for domain that appeared second from the bottom, 0 otherwise)	*0.03*	*0.45*
Bottom W (1 for domain that appeared at the bottom, 0 otherwise)	*0.12*	*1.90*
Scale parameters		
μ_{worst}: scale parameter if data relates to a "worst" or "second-worst" choice	0.82	5.90
$\mu_{excessive}$: scale parameter if data relates to case with "excessive" randomisation	0.79	4.25
Model diagnostics		
Number of observations (Hessian estimated across all choices)		14496
Number of individuals (BHHH matrix estimated on sequence of choices from each individual)		453
D.O.F.		33
Final log likelihood		-21339.15
$Rho^2(0)$		0.207
Adjusted $Rho^2(0)$		0.206

We can now compare three models:

- The model using the OSCA measure with 4 levels[14];
- The model using the OSCA measure, but collapsing the top two levels of each domain; and
- The model using the LLS measure.

In order to compare the results from these models we need to take account of the fact that each model will have a different scale, resulting from having unexplained components within the utility function of different sizes. The impact of this is that the absolute values of the model coefficients cannot be compared directly, but we can take advantage of the fact that the difference is one of scale, so can normalise the results around one coefficient (in this case the lowest level of the *accommodation* domain, which has a level description which is very similar between the measures so may be expected to return similar values).

When interpreting these charts it is important to note that the levels are not directly comparable as there are subtleties in the wording of the levels of the two measures (see the appendix for the full level descriptions), and the differences in the wording of these are more pronounced for some domains than others.

Figure 8 shows all of the levels estimated in the three models, with the upper two levels of the OSCA measure being shown next to each other on the "level 1" bars. Figure 9 shows the "level 1" bars only to allow a better comparison of these differences.

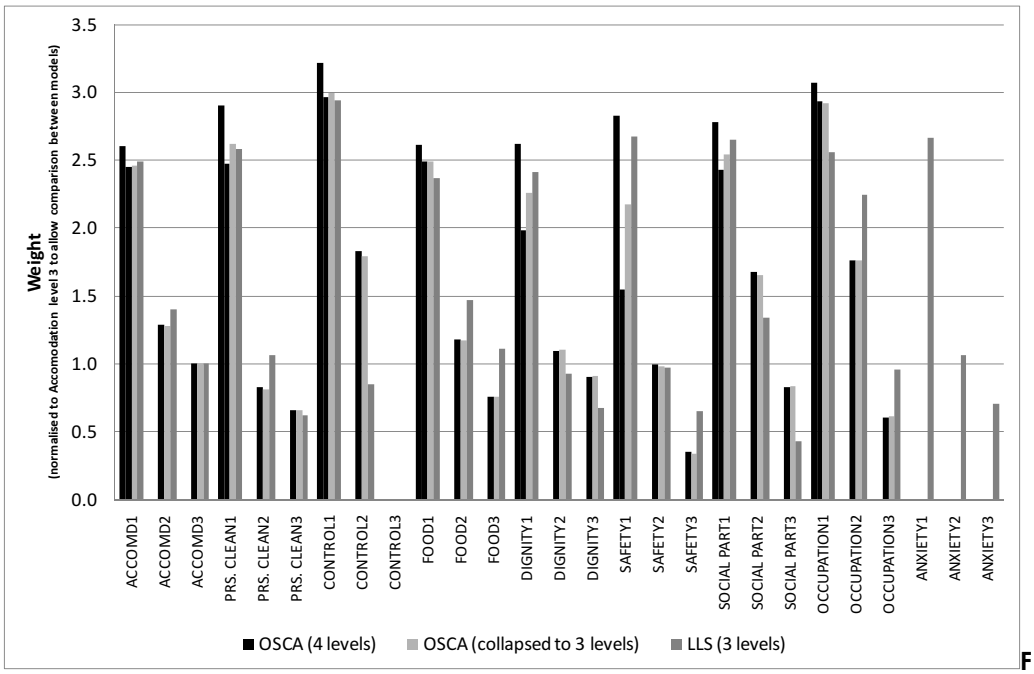

Figure 8: Comparison of measures with 4 and 3 levels

[14] For comparison purposes the top two levels of the domains in this model are shown as the two black bars on level 1. Levels 2 and 3 in the figure are then the lower two levels of the OSCA measure

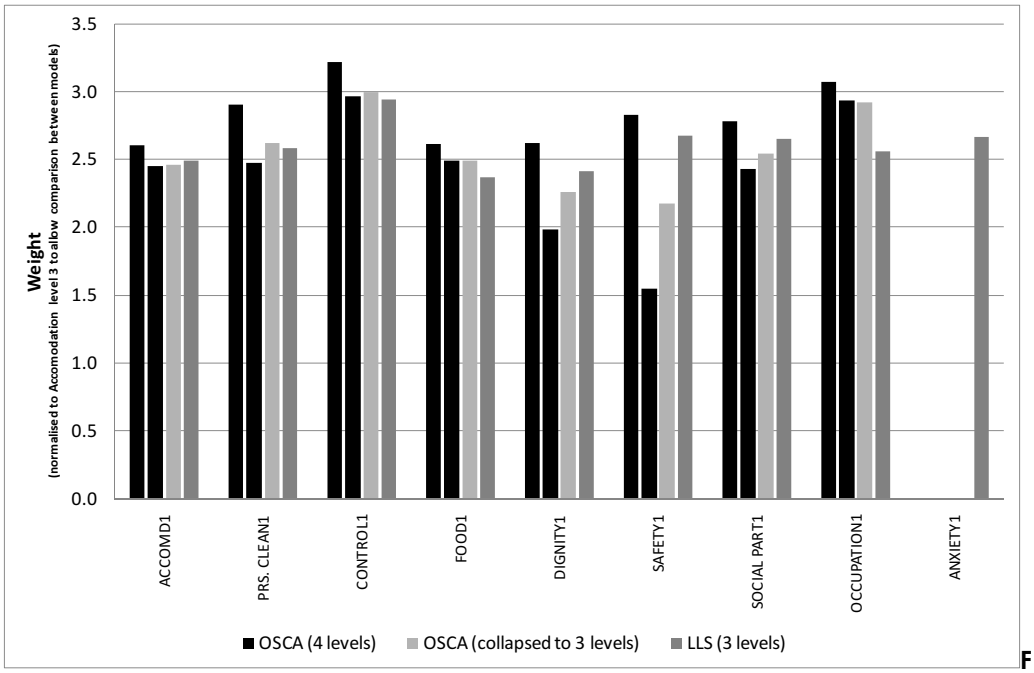

Figure 9: Comparison of upper levels from measures with 4 and 3 levels

As can be seen from Figure 9, there is relatively little difference in the utility that respondents place on the top two levels of the OSCA measure (shown as the two black bars). The notable exception to this, as noted earlier, is the *safety* domain where there is a very significant difference in the values attributed to these levels. Although the differences between the levels appear small, the model presented in Table 27 shows that these levels are strongly estimated and have values that are statistically different. Having said which, the sensitivity of any outcome related measure or toolkit developed around these levels would be relatively insensitive to differences between the two levels of need explained by these levels.

It can be seen that when the OSCA instrument is collapsed to a 3-level measure we get values that fall between the original two values for the separate levels (as would be expected) – this collapsed instrument is shown as light grey bars in Figure 9. It is interesting to note that on the whole there is a high level of consistency between this collapsed version of the OSCA levels and the lowest level of need on each domain from the LLS measure. This would suggest that it may be possible to use the OSCA measure to extend the LLS measure at a later date if it were decided that some elements of the OSCA instrument are useful in providing greater insight in to states of lower need.

It is also interesting to note that there is a reasonable level of consistency for the higher levels of need (levels 2 and 3 in Figure 8). There are notable differences are on levels 2 and 3 of *Food and Nutrition* and *Occupation and Employment*, where in each case the LLS measure values these levels higher than the OSCA measure. In the LLS measure, the *Food and Nutrition* levels the wording is based around "want", whereas in OSCA the wording is based around "need". It may be anticipated that "want" describes levels that are higher than "need", and the weights obtained would appear to be consistent with this interpretation. However, the interpretation of the wording *Occupation and Employment* levels is slightly less clear, with the LLS measure using "want" and the OSCA measure

using "value or enjoy". In this case it is harder to make a judgement as to an *a priori* ordering of these definitions. We also observe differences in the second level of *Control over daily life* and in the *Social participation* domain.

This discussion also illustrates the importance of remembering that there are subtleties in the wording of the levels of the two measures (see the appendix for the full level descriptions), and the differences in the wording of these are more pronounced for some domains than others. However, on the basis of the analysis undertaken there would appear to be sufficient overlap between the two measures, and consistency in the respondents' preferences in these cases, to allow some interchange of values between the two models (once the scale of the models have been taken in to account). It may therefore be possible to utilise some of the later OSCA findings within the ASCOT once the OSCA measure has been developed further.

CHAPTER 5 Summary

This preference study has provided new evidence on the preference weights of the population across a range of 'quality of life' domains that social care services may hope to influence through different care interventions.

Two complementary definitions of quality of life domains have been tested: the LLS measure which incorporates 9 domains at 3 levels (currently used within ASCOT), and the OSCA measure, currently under development, which incorporates 8 domains at 4 levels. Best-worst scaling exercises were developed for inclusion within a face-to-face survey in which respondents were asked to make trade-offs between the different domains when different levels of need were met.

Data has been collected from 1,000 members of the general population, with a sample that is broadly (although not entirely) representative of the observed population. Unfortunately a proportion of the sample were surveyed with an incorrect version of the questionnaire, but the implications of this have been identified, and the additional randomisation that will have occurred within the best-worst choice exercises has been taken in to account when estimating the choice models.

Within the survey, the respondents and the interviews were asked a number of diagnostic questions to ascertain whether each respondent had understood the choice task and had engaged with it. These revealed a high level of understanding and engagement, and allowed the specification of a number of data exclusions prior to the modelling.

A series of discrete choice models have been estimated from the best-worst scaling data, and results are presented for the following models:

- LLS 3-level measure with no socio-demographic segmentation on preferences;
- OSCA 4-level measure with no socio-demographic segmentation on preferences;
- LLS 3-level measure with socio-demographic segmentation on preferences; and
- OSCA 4-level measure with socio-demographic segmentation on preferences.

In each of these models the order in which the domains are presented is taken in to account, along with the specification of different model scales for the choice situations where the respondents were asked to rate the "worst" domain levels (compared to those where they rated the "best" domain levels), and different model scales for the cases where the respondents were presented with the survey which included excessive randomisation of the domains within the best-worst scaling exercise. The models have been specified to allow the computation of robust standard errors that take in to account the panel nature of the data across the choices provided by each respondent.

The models incorporating socio-economic differences in preferences have been used to calculate preference weights for the population (taking in to account the segmentation identified within the sample and apply the appropriate population weights to each segment). The final differences in preference weight placed on improvements *within* each domain of the LLS measures are shown below in Figure 10 (these reflect the value associated with moving between the lowest and other levels within each domain).

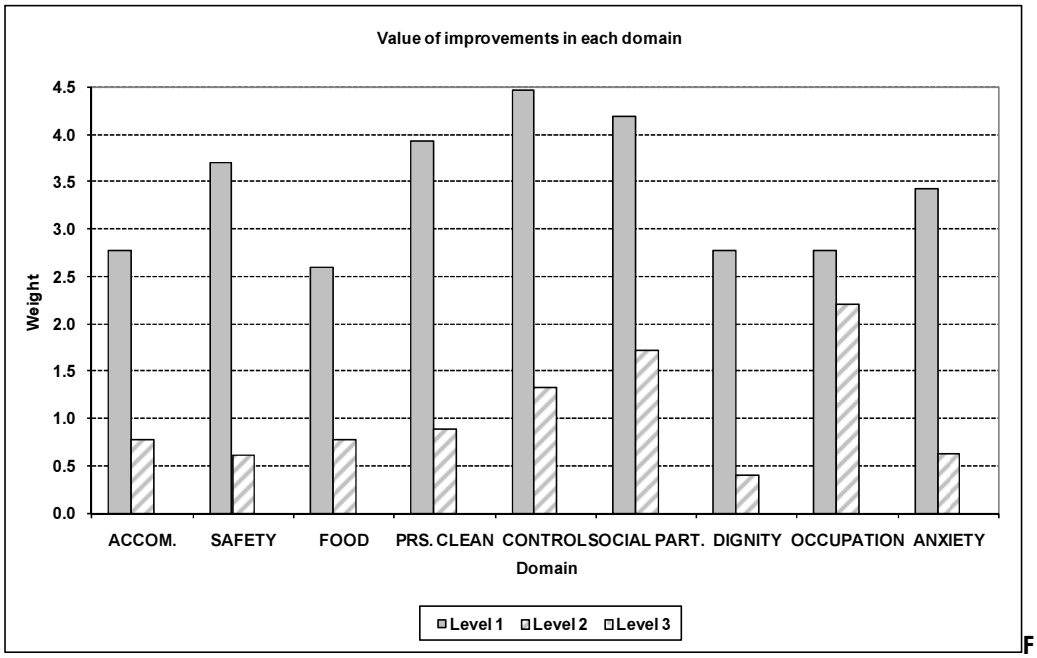

Figure 10: Value placed on improvements in needs met within each domain (LLS 3 level version)

An additional model has been estimated in which the OSCA 4-level measure has been collapsed to a 3-level instrument for comparison with the LLS measure. It is interesting to note that on the whole there is a high level of consistency between the values estimated in this collapsed version of the OSCA levels and the lowest level of need on each domain from the LLS measure. This would suggest that it may be possible to use the OSCA measure to extend the LLS measure at a later date if it were decided that some elements of the OSCA instrument are useful in providing greater insight in to states of lower need.

REFERENCES

Reference List

ALOGIT (2005) ALOGIT, HCG Software, http://www.alogit.com, London.

Ben-Akiva, M. and S. Lerman (1985) Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge, MA.

Ben-Akiva, M. and T. Morikawa (1990) "Estimation of switching models from revealed preferences and stated intentions", Transportation Research A 24, 485-496.

Bierlaire, M. (2003) BIOGEME: A free package for the estimation of discrete choice models. *Proceedings of the 3rd Swiss Transportation Research Conference.* Ascona, Switzerland.

Train, K.E. (2003), Discrete choice methods with simulation, Cambridge University Press, Cambridge, UK.

APPENDICES

Appendix A: Wording of domain levels

Table 33: Wording of domain levels in the LLS and OSCA measures

	LLS			OSCA			
	1	2	3	1	2	3	4
Accommodation Cleanliness and Comfort	My home is as clean and comfortable as I want	My home is less clean and comfortable than I want	My home is not at all as clean or comfortable as I want	My home is as clean and comfortable as I want	My home is adequately clean and comfortable	My home is less than adequately clean or comfortable	My home is not at all clean or comfortable
Safety	I feel as safe as I want	Sometimes I do not feel as safe as I want	I never feel as safe as I want	I feel as safe as I want	Generally I feel adequately safe, but not as safe as I would like	I feel less than adequately safe	I don't feel at all safe
Food and Nutrition	I eat the meals I like when I want	I don't always eat the right meals I want, but I don't think there is a risk to my health	I don't always eat the right meals I want, and I think there is a risk to my health	I get all the food and drink I like when I want	I get food and drink adequate for my needs	I don't get all the food and drink I need, but I don't think there is a risk to my health	I don't get all the food and drink I need, and I think there is a risk to my health
Personal Care	I feel clean and wear what I want	I sometimes feel less clean than I want or sometimes can't wear what I want	I feel much less clean than I want, with poor personal hygiene	I feel clean and am able to present myself the way I like	I feel adequately clean and presentable	I feel less than adequately clean or presentable	I don't feel at all clean or presentable
Control over Daily Life	I have as much control over my daily life as I want	Sometimes I don't feel I have as much control over my daily life as I want	I have no control over my daily life	I have as much control over my daily life as I want	I have adequate control over my daily life	I have some control over my daily life but not enough	I have no control over my daily life

	LLS			OSCA			
	1	2	3	1	2	3	4
Social Participation & Involvement	My social situation and relationships are as good as I want	Sometimes I feel my social situation and relationships are not as good as I want	I feel socially isolated and often feel lonely	I have as much social contact as I want with people I like	I have adequate social contact with people	I have some social contact with people, but not enough	I have little social contact with people and feel socially isolated
Dignity	I would be treated by other people with the dignity and respect that I want	Sometimes I would not be treated by other people with the dignity and respect that I want	I would never be treated with the dignity and respect that I want	The way I'm helped and treated makes me think and feel better about myself	The way I'm helped and treated does not affect the way I think or feel about myself	The way I'm helped and treated sometimes undermines the way I think and feel about myself	The way I'm helped and treated completely undermines the way I think and feel about myself
Occupation & Employment	I do the activities I want to do	I do some of the activities I want to do	I don't do any of the activities I want to do	I'm able to spend my time as I want, doing things I value or enjoy	I'm able do enough of the things I value or enjoy with my time	I do some of the things I value or enjoy with my time but not enough	I don't do anything I value or enjoy with my time
Anxiety	I feel free from worry and concerns on a day-to-day basis	I sometimes feel worried and concerned	I feel very worried and concerned on a daily basis	-	-	-	-

Table 34: Distribution of best, worst, second best, and second worst choices in both experiments (%)

		LLS experiment				OSCA experiment			
Domain	**Level**	**Best**	**Worst**	**Second best**	**Second worst**	**Best**	**Worst**	**Second best**	**Second worst**
Accommodation, Cleanliness and Comfort	1	6.4	1.1	8.5	0.7	4.9	0.7	5.6	0.6
	2	1.9	2.2	1.9	3.3	3.2	0.6	5.0	0.7
	3	1.0	3.5	0.9	4.9	0.7	2.7	0.9	4.4
	4					0.9	4.4	0.8	5.6
Safety	1	9.4	1.0	8.5	0.7	6.3	0.6	5.4	0.5
	2	1.0	4.4	1.0	5.4	1.4	3.3	2.0	4.2
	3	0.9	7.0	1.0	6.5	0.5	4.9	0.8	5.5
	4					0.4	9.6	0.2	6.3
Food and Nutrition	1	5.4	1.1	8.0	0.9	4.5	0.5	5.2	0.6
	2	1.2	1.7	1.8	2.2	3.6	0.7	5.3	0.7
	3	0.8	3.1	1.3	4.4	0.5	3.4	0.8	4.8
	4					0.4	6.1	0.5	5.7
Personal Cleanliness and Comfort	1	8.1	0.7	8.2	0.6	6.8	0.6	6.1	0.7
	2	0.3	2.7	0.7	4.7	3.8	0.8	5.2	0.8
	3	0.4	6.3	0.9	6.4	0.4	6.1	0.6	5.6
	4					0.3	6.4	0.6	6.5
Control over Daily Life	1	13.9	1.5	8.5	1.0	10.5	0.9	6.0	0.8
	2	1.1	5.7	1.5	6.3	8.5	0.9	6.1	0.9
	3	1.1	13.8	0.7	7.6	3.6	3.0	2.9	3.3
	4					0.8	12.5	0.5	5.8
Social Participation and Involvement	1	9.4	1.0	8.4	1.1	5.7	0.6	6.7	0.7
	2	1.7	2.9	2.1	3.9	3.8	0.9	4.7	0.9
	3	0.6	8.4	0.8	7.5	1.4	1.9	2.3	3.6
						0.6	5.4	0.9	6.4
Dignity	1	7.3	1.4	6.8	1.6	4.4	0.8	4.7	1.1
	2	1.1	5.5	1.2	5.0	1.4	1.0	2.4	1.7
	3	0.9	7.4	0.7	5.8	0.8	5.2	0.7	4.3
	4					0.8	5.4	0.8	5.4
Occupation and Employment	1	8.7	1.0	9.2	1.1	8.9	0.5	5.9	0.6
	2	5.4	0.7	6.2	0.8	7.7	0.7	6.2	1.0
	3	1.0	4.6	1.3	4.9	2.2	2.5	3.7	3.6
	4					0.5	6.4	0.6	6.9
Anxiety	1	9.5	1.4	8.1	1.3				
	2	0.8	3.9	1.3	4.7				
	3	0.8	6.0	0.8	6.9				
Total		**100**	**100**	**100**		**100**	**100**	**100**	**100**

Appendix B: Choice elicitation and model estimation procedure in Best-Worst Scaling data

Each best-worst task involved four sequential choices (i.e., best, worst, second best, second worst). Firstly, respondents were presented with 9 (in the LLS experiment) or 8 (in the OSCA experiment) domain levels and were asked to choose what they thought would be the best. Next, the chosen domain level was taken off the best-worst task and respondents were asked to choose the worst of the remaining alternatives. Following that, respondents were again asked to choose the best (what we call second best) of the remaining alternatives (7 in the LLS experiment, 6 in the OSCA experiment, respectively). Finally, respondents chose the worst alternative (what we call second worst) of the domain-levels remaining in the given best-worst task.

There are two ways of analysing this; either by looking at it as respondents aiming to maximise their utility over four related choices (and set up equations where the alternatives are all possible combinations of best-worst-second best-second worst sequences), or starting from the point of considering these as four separate choice processes and then take account of the fact they come from the same original BWS scenario.

The *Maxdiff* is the first approach described above. However, the estimation procedure of the *Maxdiff* significantly more complicated to analyse a data set that includes second best and second worst choices. Specifically, the problem lies around the very higher number of possible alternatives. In this study, the number of alternatives should include all possible combinations of best-worst pairs, namely $2^{(9*3)}$ in the LLS data and $2^{(8*4)}$ in the OSCA data). When we add the second best and second worst choices, the number of alternatives quickly becomes impossible to handle - even for a simple estimation. Also, the *Maxdiff* approach is used less by researchers and the majority of applications model best-worst scaling data as a series of independent choices.

Therefore the analysis started from the assumption that all choices made were independent from the previous choices even if those came from the same best-worst task (a correction to the standard errors was applied at a later stage of the model estimation). Using the LLS experiment as an example, the model coding involves the following:

- Choice of best: respondent chooses one out of 9 alternatives
- Choice of worst: respondent chooses one out of 8 alternatives
- Choice of second best: respondent chooses one out of 7 alternatives

- Choice of second worst: respondent chooses one out of 6 alternatives

In the first case the respondent is indicating the alternative with the highest utility, in the second the alternative with the lowest utility, then next highest utility, then next lowest utility.

In principle the alternatives in the model are every possible domain level, but then with availability criteria that say that in a given choice scenario a respondent only sees 9 (or 8 or 7 or 6) of the possible domain levels. It is possible to further simplify the coding (and gain significant efficiencies in estimating the model) since only one level appears for any given domain at the same time, so it is possible to set up a model with 9 possible alternatives to choose between, which have utility functions defined to take account of which level is presented for that domain. This is done using dummy coding to reflect the domain levels which were presented.

The terms in the utility are positive when the outcome is a best or second best choice (i.e. choice is maximising positive utility) and negative where the outcome is a worst or second worst choice (i.e., choice is maximising negative utility).

Additional availability criteria are applied moving through the best, worst, second best, second worst sequence to eliminate the previously chosen domains from the choice set for the subsequent choices.

Therefore, the dependent variable in the model is the domain level chosen (from those presented) and the utility of the alternative is a function of the level at which the domain was presented. The coefficient on just one of all domain levels needs to be constrained to zero to allow the model to be identified, and all other coefficients are therefore relative to this domain level.

This multinomial logit model then has the constraint that the observations are independent of each other relaxed by specifying the model to allow the BHHH matrix to be estimated on the sequence of choices from each individual. This provides more robust estimates of the standard errors, taking in to account the panel nature of the data.