MONASH University
Business and Economics

Centre for Health Economics

Research Paper 2012 (81)

Cross-national comparison of twelve quality of life instruments

MIC Paper 4

United States of America

Professor Jeff Richardson

Foundation Director, Centre for Health Economics
Monash University

Dr Munir A Khan

Research Fellow, Centre for Health Economics
Monash University

Angelo Iezzi

Research Fellow, Centre for Health Economics
Monash University

Aimee Maxwell

Research Fellow, Centre for Health Economics
Monash University

November 2012

Centre for Health Economics
ISSN 1833-1173
ISBN 1 921187 80 8

The Multi Instrument Comparison (MIC) survey is a project funded by a National Health and Medical Research Council (NHMRC) project grant (ID 1006334) ‘A cross national comparison of eight generic quality of life instruments’. (Since its inception, three additional instruments have been added).

Chief Investigators on the project are:

Professor Jeff Richardson (Monash University, Australia)
Professor Robert Cummins (Deakin University, Australia)
Professor Robert Kaplan (University of California Los Angeles, USA)
Professor Jan Abel Olsen (University of Tromso, Norway)
Professor Michael Schlander (University of Heidelberg, Germany)
Professor Joanna Coast (University of Birmingham, England)

Upon completion of the project all data will be made publicly available.

Professor Jeff Richardson
Centre for Health Economics
Faculty of Business and Economics
Monash University   Vic   3800
Australia

Ph: +61 3 9905 0754 Fax: +61 3 9905 8344
Jeffrey.Richardson@monash.edu

ABSTRACT

The Multi Instrument Comparison (MIC) project is the largest comparative study of health and wellbeing instruments undertaken worldwide. To date 6773 individuals have completed twelve instruments relating to their health or wellbeing. Data were collected from a representative healthy cohort and from patients in eight clinical areas in each of five countries.

This and other country-specific research papers report data related to the project study questions. They do not seek to interpret data or comment on the study questions. This will be the subject of later publications.

Countries, diseases and questionnaires included in the MIC are summarised in Boxes 1 to 4 below. The background study questions questionnaires and utility weights used are outlined in detail in MIC Paper 1, Background, Questions, Instruments (Richardson, Iezzi et al. 2012). Choice of weights is also discussed in Section 8.

Box 1 Country and disease area summary as at May 2012

Respondent numbers after editing

Total sample	
Australia	1436
UK	1358
USA	1467
Canada	1335
Norway	1177
Total	6773

Health state	
Arthritis	770
Asthma	709
Cancer	657
COPD	66
Depression	757
Diabetes	784
Chronic heart disease	791
Stroke	23
Hearing problems	716
Total disease	5273
Healthy	1500

Box 2 Main Questionnaire

Type	Title	Questions
Subjective Wellbeing (SWB)	Personal Wellbeing Index (PWI)	9
	Integrated Household Survey (IHS)	5
	Satisfaction with Life Survey (SWLS)	4
	subtotal	**18**
Multi Attribute Utility (MAU) Instruments	EQ-5D	5
	AQoL-8D and AQoL-4D	44
	HUI3	8
	15D	15
	QWB-^{SA}	77
	SF-6D (derived from SF-36)	
Non-Utility	SF-36	36
	Self TTO	1
	ICECAP-A	5
Demographics		18
	Total items in composite instrument	**227**

Box 3 Respondents with a chronic disease by disease and country

Diseases	Australia	UK	USA	Canada	Norway	Total
Asthma	141	150	150	138	130	709
Cancer	154	137	148	138	80	657
Depression	146	158	168	145	140	757
Diabetes	168	161	168	144	143	784
Hearing problems	161	128	163	149	115	716
Arthritis	163	159	179	139	130	770
Heart disease	149	167	170	154	151	791
COPD	66	x	x	x	x	66
Stroke	23	x	x	x	x	23
Disease sample	1171	1060	1146	1007	889	5273
'Healthy public'	265	298	321	328	288	1500
Total	1436	1358	1467	1335	1177	6773

Box 4 Sources of utility weights[1]

Instrument	Country and Respondents	Method of Calibration	Reference
EQ-5D-5L	US Public n=3691	TTO	Interim scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L value sets http://www.euroqol.org/about-eq-5d/valuation-of-eq-5d/eq-5d-5l-crosswalk-value-sets.html
SF6D	US Public n=611	SG	Brazier, J, Roberts J, Deverill M: The estimation of a preference-based measure of health From the SF-36. J Health Econ. 2002 mar;21(2)271-92
HUI3	Canada Public n= 256	SG	Furlong W, Feeny D, Torrance GW, et al. Multiplicative Multi-Attribute Utility Function for the Health Utilities Index Mark 3 (HUI3) System: A Technical Report, McMaster University Centre for Health Economics and Policy Analysis Working Paper No. 98-11, December 1998.
15D	Finland Public n=1255	VAS	Brazier, J., Ratcliffe, J., Salomon, JA. and Tsuchiya, A. (2007):'Measuring and Valuing Health Benefits for Economic Evaluation' Oxford University Press, page 195. http://www.15d-instrument.net/15d
QWB	USA Public n=435	VAS	Sieber W, Groessl E, David K, Ganiats T, Kaplan R. (2008): Quality of Well Being Self-Administered (QWB-SA) Scale, User's Manual, Health Services Research Centre, University of California, San Diego. https://hoap.ucsd.edu/qwb-info/QWB-Manual.pdf
AQoL-4D	Australia Public n=350	TTO	Hawthorne, G., Richardson, J., Day, N., Osborne, R., McNeil, H.(2000) Construction and Utility Scaling of the Assessment of Quality of Life (AQoL) Instrument. Monash University Centre for Health Economics Working paper 101. http://www.buseco.monash.edu.au/centres/che/pubs/wp101.pdf
AQoL-8D	Australia Public =347 Patient =323 n=670	TTO	Richardson J, Iezzi A: Psychometric validity and the AQoL-8D Multi Attribute Instrument. Research Paper 71 (2011). Centre for Health Economics, Monash University, Australia http://www.buseco.monash.edu.au/centres/che/pubs/researchpaper71.pdf

[1] Choice of weights is also discussed in Section 8.

Box 5 List of abbreviations

MA	Multi attribute
MAU	Multi attribute utility
MAUI	Multi attribute utility instrument
SWB	Subjective wellbeing (‘happiness’)
CUA	Cost Utility Analysis

TABLE OF CONTENTS

1 Introduction 1
2 Respondent characteristics 4
3 Summary statistics 6
4 Correlation 9
5 Linear relationships 16
6 Instrument content (sensitivity) 20
7 Pairwise comparison of instruments 34
8 Discussion and Conclusion 42
References 55

List of Boxes

Box 1 Country and disease area summary as at May 2012 iii
Box 2 Main Questionnaire iii
Box 3 Respondents with a chronic disease by disease and country iv
Box 4 Sources of utility weights iv
Box 5 List of abbreviations v

List of Tables

Table 1 Edit procedures – US patients and public 3
Table 2.1 'Healthy Public': Age and gender 4
Table 2.2 Healthy public: Highest education by gender 4
Table 2.3 Distribution of disease group by age and gender 5
Table 3.1 Summary statistics for the MAU instruments (Public n=321) 6
Table 3.2 Summary statistics for the MAU instruments (Total n=1467) 6
Table 3.3 Reliability of instruments 7
Table 4.1 Pearson correlation between MAU (Public n=321) 10
Table 4.2 Pearson correlation between MAU (Total n=1467) 10
Table 4.3a Pearson correlations between MAUI instruments (Public n=321) 10
Table 4.3b Pearson correlations between MAUI and non-MAU instruments (Total n=1467) 11
Table 4.4 Pearson correlations between non-MAU and non-MAU instrument (Total n=321) 11
Table 4.5 Pearson correlations between non-MAU instruments (Total n=1467) 12
Table 4.6 Intraclass correlation between MAU instrument (Total n=1467) 13
Table 5.1 Discrepancies in marginal change: slope, coefficient, b, in regression 19
Table 5.2 Difference in marginal change: public vs total (instrument A=a+b instrument B) 19
Table 6.1a Ceiling effects (MAU = 1) Ave value of other MAUI when an MAU=1.0 (Public n=321) 20
Table 6.1b Ceiling effects (MAU=1.0) Ave value of other MAUI when an MAU=1.0 (Total n=1467) 20
Table 6.1c Floor effects Ave value of other MAUI when an MAU < .40 (Total n=1467) 21
Table 6.2 Correlation of instruments with SF-36, AQoL-8D physical and psycho-social scales (Total n=1467) 21
Table 6.3a Ratio of scores in top and bottom 50% of total sample, ranked by MAUI (SF-36 dimensions) 25
Table 6.3b Ratio of scores in top and bottom 50% of total sample, ranked by MAUI (AQoL-8D dimensions, SWB and Self-TTO) 25
Table 6.4a Sensitivity to SF-36 dimensions: Beta coefficient and R^2 from the regression of MAU on single dimensions of the SF-36 (Total n=1467) 26
Table 6.4b Sensitivity to SF-36 dimensions: Beta coefficient from the regression of MAU on all dimensions of the SF-36 (Total n=1467) 27
Table 6.5a Sensitivity to AQoL-8D dimensions: Beta coefficient and R^2 from the regression of MAU on single dimensions of the AQoL-8D 28
Table 6.5b Sensitivity to AQoL-8D dimensions: Beta coefficient from the regression of MAU on all dimensions of the AQoL-8D 29
Table 6.6 Instrument content: regression of MAU on non-MAU instruments 30
Table 7.1 Dimension and instrument correlations with MAU residuals (Total 1467) and SF-36 dimensions and SWB instruments 35
Table 7.2 Dimension and instrument correlations with MAU residuals (Total 1467) and AQoL-8D dimensions **....36**
Table 7.3 Sensitivity: Summary of pairwise comparisons 42
Table 8.1 Summary of MAU order by criteria (USA) 45

List of Figures

Figure 1 Administration of the MIC online questionnaires 2
Figure 3.1 Mean of MAU instruments (Total n=1467) ....
Figure 3.2 Mean EQ-5D by disease group (Total n=1467) 8
Figure 4.1 Average Pearson correlation with other MAU Instruments (Total n=1467) 12
Figure 4.2 Average Intraclass correlation with other MAU Instruments (Total n=1467) 13
Figure 4.3 R^2: MAUI on MAU Instruments (Total n=1467 13
Figure 4.4 Pearson correlation of MAU instrument with PWI (Total n=1467) 14
Figure 4.5 Pearson correlation of MAU instrument with SWLS (Total n=1467)...........................14
Figure 4.6 Pearson correlation of MAU instrument with Self-TTO (Total n=1467) 14
Figure 4.7 Pearson correlation of MAU instrument with SF-36 (Public n=321) 15
Figure 4.8 Pearson correlation of MAU instrument with SF-36 (Total n=1467) 15
Figure 5.1 Geometric regression results (Public n=321)..17
Figure 5.2 Geometric regression results (Total n=1467)..18
Figure 6.1 Correlation with summary scores of SF-36 (PCS and MCS) and AQoL-8D (PSD and MSD) 22
Figure 6.2 Comparison of Summary Physical and Psycho-Social Dimensions (Average SF-36 and AQoL-8D summary scores) 23
Figure 6.3 Effect of SD change in dimension on standardised score (beta coefficient) 31
Figure 6.4 Instrument content: Disaggregated by AQoL-8D dimensions....................................32
Figure 6.5 Split half analysis: Ratios of values in top/bottom half of population ranked by instrument 33
Figure 7.1 Pairwise instrument sensitivity: Correlations of dimension scores with the residual of one MAU instrument regressed upon a second MAU instrument...........................37
Figure 8.1 Comparison of EQ-5D with US and UK weights 46
Figure A.1.1 Frequency distribution of MAU instruments (Total n=1467) 47
Figure A.1.2 Frequency distribution of MAU instruments (Public n=321)......................................48
Figure A.2.1 Frequency distribution of non-MAU instruments (Total n=1467)...............................49
Figure A.2.2 Frequency distribution of non-MAU instruments (Public n=321) 50
Figure A.3.1 Frequency distribution of residuals from pairwise regression of MAU instruments...52

Cross-national comparison of twelve quality of life instruments: MIC Paper 4 USA

1 Introduction

Objectives

The background and objectives of the MIC project are described in MIC Paper 1 (Richardson, Iezzi et al. 2012). In sum, the project is a response to the evidence that different MAU instruments produce different values for ‘utility’ and (despite the common label ‘utility’) measure different constructs. The principle objectives of the project are, firstly, to document the differences in the values produced by the instruments for different groups of patients in different countries; and, secondly, to determine what the different instruments measure – which dimensions of wellbeing explain variation in instrument scores.

To achieve these objectives we sought respondents with a diverse range of health states and, specifically, health states associated with major disease areas. This implies that the total sample is not representative of the population as the focus of the study is the relationships between instruments in different health states and not the wellbeing of the overall population. Despite this, comparisons may be made with population or other instrument norms. ‘Patients’ complete a disease-specific questionnaire for which there are norms and the non-patient sample may be weighted to correct for any mismatch between them and independently obtained norms if population values are needed.

The primary objectives relate to the content and validity of existing instruments, ie those which are currently used for cost utility analysis (CUA). While the investigation of the psychometric properties of the instruments are a further area of inquiry the main research, including results reported in this paper, use unadjusted MAU instruments irrespective of their reliability as indicated by the present data. The instruments are currently used irrespective of their properties.

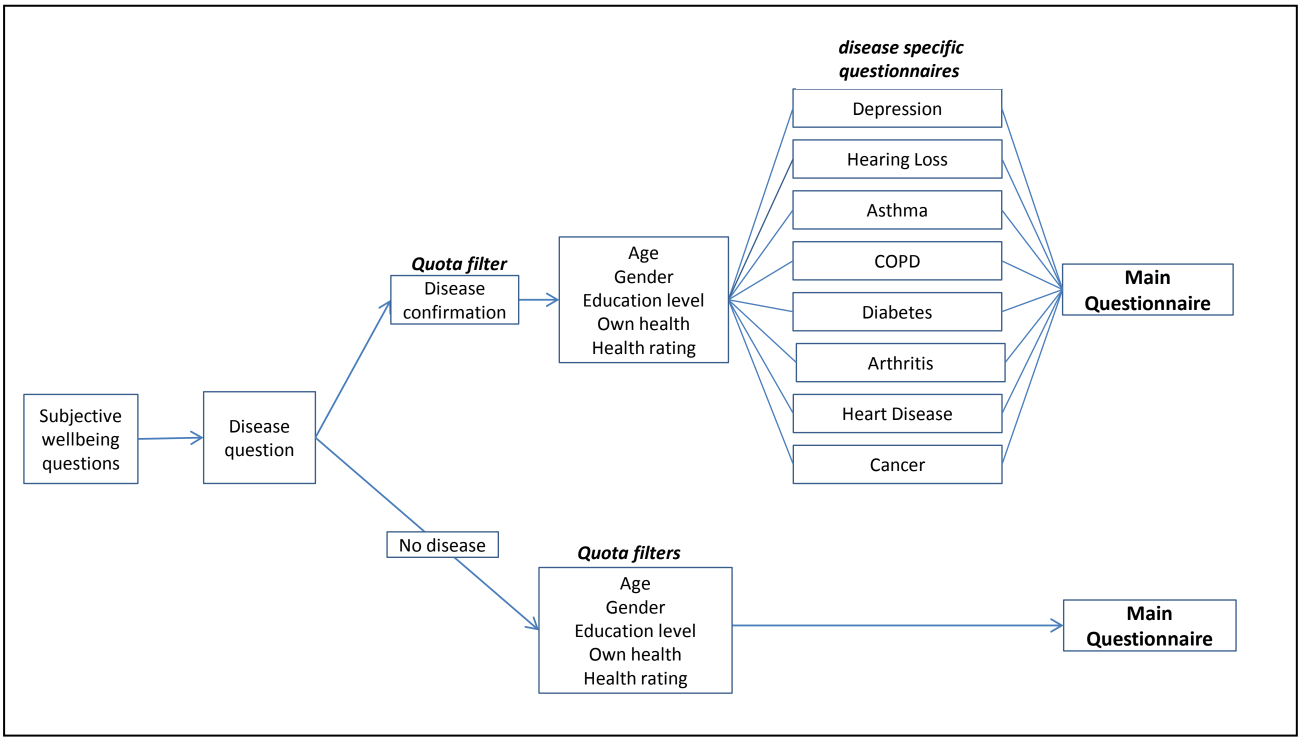
The administration of the MIC survey is illustrated in Figure 1. A survey company, CINT, invited individuals on their database to participate. A person accepting this invitation was first asked to complete the three subjective wellbeing questions: the Personal Wellbeing Index (PWI), the Integrated Household Survey (IHS) and the Satisfaction with Life Survey (SWLS). These questions were administered immediately as they seek to measure ‘affect’ – a person’s ‘undigested’ feelings. Asking the questions after ‘priming’ respondents with questions about their health (do you have one of the eight diseases of interest?) would potentially create biased responses.

After completion of these questions the respondent was asked the following question: ‘Have you got a current diagnosis of any of the following health problems? Please choose the most serious illness you have.’

Those nominating one of the survey diseases proceeded with the survey if and only if the quota – the target number of respondents – had not been reached. To confirm the patient’s status the first question was a repetition of the question above. Patients then completed the core questionnaire which was administered to all respondents within the quota. This was followed by the disease-specific questionnaire which applied to their particular disease.

Those who did not report a disease were questioned about their age, gender and education. Additionally they were asked to indicate their overall health on a visual analogue scale (VAS) where *‘Zero is the least desirable state of health you could imagine and 100 is the best possible health (physical, mental and social).*’ The individual was invited to proceed to the core questions only if their VAS score exceeded 70 and their age, gender and education quota had not been filled. The VAS criterion was included to ensure that the ‘healthy public’ excluded those whose self rating was very poor. The web-based procedure employed here attracts a disproportionate number of distressed respondents and the procedure was adopted to reduce this effect and increase the sample size of respondents in good health. The number 70 was selected judgementally to achieve this goal but to permit variation in ‘normal health’.

Figure 1 Administration of the MIC online questionnaires



Editing

Introductory comments from the panel company to their panellists were designed to deter unreliable respondents. Eight edit criteria were subsequently used to eliminate unreliable answers. These were:

Edit 1: Any response that was completed in less than 20 minutes was eliminated. The survey median completion time was 40 minutes (range 7.7-260.9 minutes). Times between 20-25 minutes were marked for subsequent inspection (Edit 7, 8). Records with duplicated IDs were eliminated.

Edit 2: The EQ-5D mobility question was duplicated in the survey. Anyone with a response that varied by more than +/- 1.00 was eliminated. Those differing by only +/- 1.00 were earmarked for subsequent inspection (Edit 7, 8).

Edit 3: The SF-36 question 1 and question concerning own health were identical. Those with responses greater than +/- 1.00 were eliminated. Those without identical answers but within +/- 1.00 were earmarked.

Edit 4: SF-36 question 1 and QWB question 9a were identical. The same procedure was followed as above.

Edit 5: Own health and QWB question 9a were identical. The same criterion was followed as above.

Edit 6: EQ-5D question 4 (pain) and AQoL-8D question 22 (pain) were very similar. Those with two response level differences were eliminated.

Edit 7: The number of inconsistencies from edits 2, 3, 4, 5 and 6 were summed. Those with two or more inconsistencies and a time less than 25 minutes were eliminated.

Edit 8: Those with three or more inconsistencies were eliminated.

The effect of these procedures on USA respondents with self-reported disease is shown in Table 1.

Table 1 Edit procedures – US patients and public

Stage	Deleted	Remaining	Stage	Deleted	Remaining
Patients		1282	Edit 5	2	1187
Edit 1	67	1279	Edit 6	5	1185
Edit 2	14	1215	Edit 7	28	1180
Edit 3	12	1201	Edit 8	6	1152
Edit 4	2	1189	Total	136 (11%)	1146 (89%)
Public		387			
Edits 1-5	66	321	Total	66 (17%)	321 (83%)

Utility weights

Utility weights for all instruments are not available for all countries. Box 4 reports the weights used in the initial analysis with the project. In principle the use of alternative weights for different countries may alter results. This is discussed further in Section 8 which presents a comparison of US and US weights for the EQ-5D data from the MIC project. It does not suggest that the explanatory power of the EQ-5D could alter with a choice between these weights.

2 Respondent characteristics

The healthy public

After conclusion of the edit procedures outlined above 1467 respondents were retained, 1146 patients' and 321 representing the 'healthy public'. Table 2.1 shows the distribution of respondents by age and gender. The highest level of education of the public respondents by gender is reported in Table 2.2.

Patient samples

1146 patient surveys were retained. The focus of the study is upon the comparison of instruments and the purpose of the patient samples was primarily to maximise the diversity of health states in the sample. Consequently, no age-gender quotas were used. Table 2.3 disaggregates respondents by age, gender and disease group. It indicates that the overall sample is highly skewed with respect to age reflecting the increasing probability of a chronic disease with age.

Table 2.1 'Healthy Public': Age and gender

Age group	US Public		Total
	Male (%)	Female (%)	
18-24	4.7	5.6	33
25-34	8.1	9.7	57
35-44	7.5	10.6	58
45-54	9.3	11.2	66
55-64	7.8	8.1	51
65+	7.8	9.7	56
Total	45.2	54.8	321

Table 2.2 Healthy public: Highest education by gender

Education	US Public		Total
	Male (%)	Female (%)	
High school	20.6	21.8	136
Diploma or certificate or trade	10.0	13.1	74
University	14.6	19.9	111
Total	45.2	54.8	321

Table 2.3 Distribution of disease group by age and gender

Diseases	Age group by gender												Total		
	18-24		25-34		35-44		45-54		55-64		65+				
	M	F	M	F	M	F	M	F	M	F	M	F	M	F	T
Asthma	3	13	5	24	11	22	7	22	4	24	4	11	34	116	150
Cancer	1	2	3	4	4	18	11	23	10	31	26	15	55	93	148
Depression	4	16	10	21	6	23	20	40	5	17	2	4	47	121	168
Diabetes	1	2	4	5	7	14	27	20	19	32	18	19	76	92	168
Hearing problems	1	2	2	9	7	10	14	20	25	19	34	20	83	80	163
Arthritis	1	2	3	7	6	11	14	33	8	38	11	45	43	136	179
Heart problems	2	5	1	3	3	8	15	21	29	29	31	23	81	89	170
Healthy people	15	18	26	31	24	34	30	36	25	26	25	31	145	176	321
Total	28	60	54	104	68	140	138	215	125	216	151	168	564	903	1467

3 Summary statistics

Mean values

Summary statistics for the twelve instruments are reported in Tables 3.1 and 3.2. MAU instruments purport to measure the same construct – utility. Consequently, direct comparison of their scores is appropriate. Other instruments may not be directly compared. The PWI, SWLS and IHS all measure facets of subjective wellbeing (SWB). However, they do not purport to measure the same construct and their correlation reflects this (see Table 4.4).

Differences between patient groups are not the principle focus of the present report. Nevertheless, the average utility using a single MAU – the EQ-5D – is shown in Figure 3.2.

Frequency distributions for each of the instruments are reported in Appendices 1 and 2.

Table 3.1 Summary statistics for the MAU instruments (Public n=321)

	EQ-5D[1]	HUI3	SF-6D	15D	QWB	AQoL-4D[2]	AQoL-8D[3]
Mean	.90	.89	.80	.94	.76	.81	.87
N	321	321	321	321	321	321	321
SE	.006	.008	.006	.004	.008	.011	.007
SD	.104	.151	.115	.070	.140	.193	.130
Minimum	.14	.03	.53	.46	.28	.08	.33
Maximum	1.00	1.00	1.00	1.00	1.00	1.00	1.00

(1) Kind et al. (1999)

(2) Hawthorne et al. (2012)

(3) Richardson et al. (2012)

Table 3.2 Summary statistics for the MAU instruments (Total n=1467)

	EQ5D	HUI3	SF-6D	15D	QWB	AQoL4D	AQoL8D
Mean	.73	.70	.70	.84	.63	.63	.72
N	1467	1467	1467	1467	1467	1467	1467
SE	.006	.007	.004	.003	.004	.007	.006
SD	.228	.274	.140	.133	.161	.269	.219
Minimum	-.32	-.33	.30	.31	.15	-.04	.03
Maximum	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Score	%						
1.00	17.1	7.5	2.1	7.4	2.9	6.1	2.7
0.95+	17.1	14.6	3.6	26.1	2.9	8.0	13.2
<0.4	8.9	14.5	1.4	0.3	7.4	22.0	10.1
<0.1	2.6	4.4	0.0	0.0	0.0	4.6	0.3
<0.0	1.6	2.0	0.0	0.0	0.0	1.0	0.0

Internal reliability

A test of scale reliability was carried out with public data using the Cronbach's alpha (Cronbach 1951). This determines the internal consistency or average correlation of items in a survey instrument. The reliability of a scale can vary depending on the sample that it is used with. Table 3.3 reports the alpha coefficient. If this is above 0.7, the scale can be considered reliable with the sample (Pallant 2010). The result shows that all of the scales pass this test except for the IHS.

Table 3.3 Reliability of instruments

Instrument	No of items	Cronbach's Alpha
AQoL-4D	12	.83
AQoL-8D	35	.96
HUI3	8	.70
EQ-5D	5	.81
15D	15	.88
QWB	251	.90
ICECAP	5	.86
SF-36	36	.70
IHS	4	.50
SWLS	5	.92
PWI	9	.91

* These values are below those generally accepted as indicating the reliability of a scale.

Figure 3.1 Mean of MAU instruments (Total n=1467)

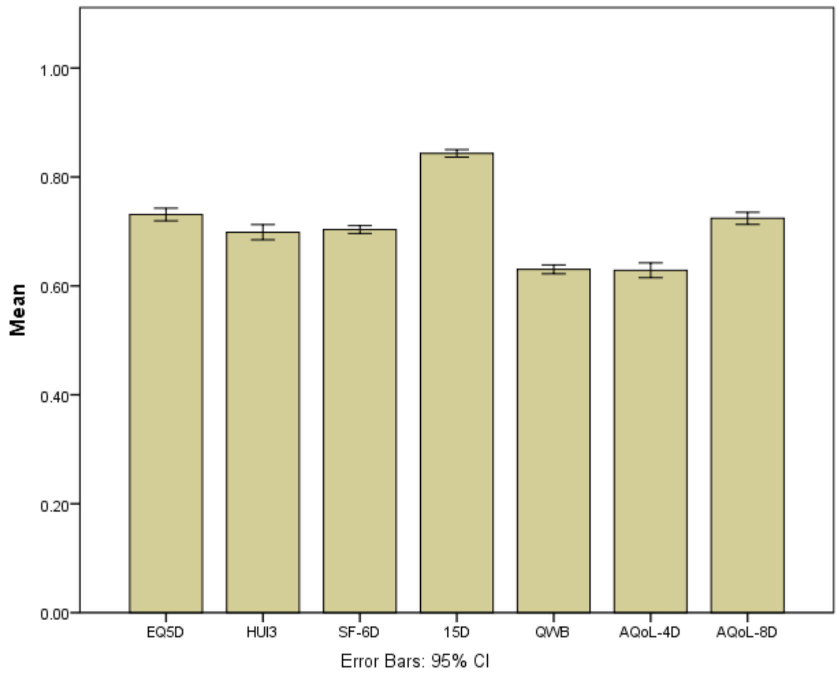
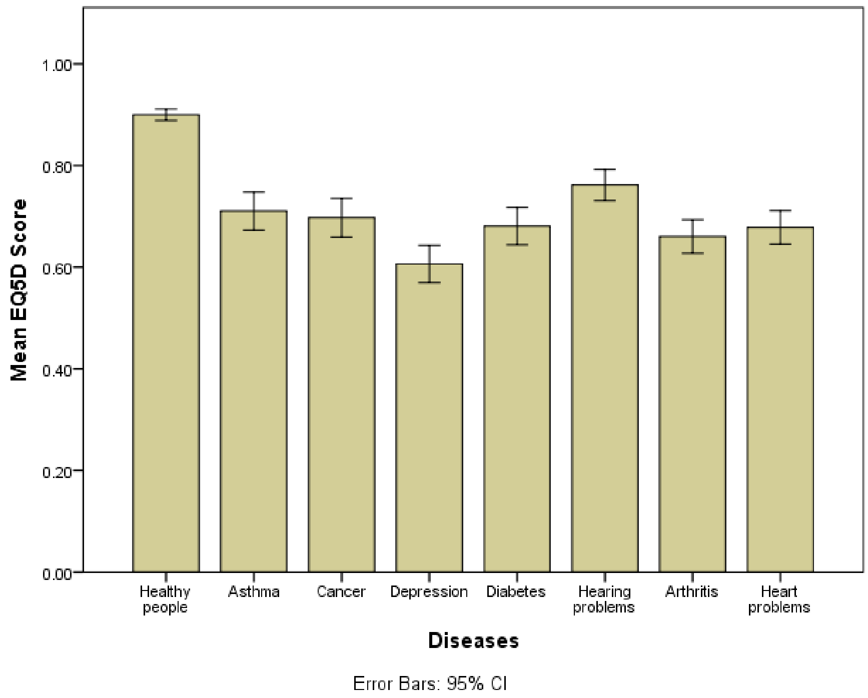


Figure 3.2 Mean EQ-5D by disease group (Total n=1467)



4 Correlation

Validation tests draw heavily upon correlation. In particular, convergent validity is established if an instrument correlates as predicted with other instruments or criteria scores which are believed to correlate with the construct. Higher correlation justifies greater confidence in overall validity. The MIC project collected several types of data to test convergent validity. These were:

1. *Other MAU scores*. As each MAU instrument is believed to reflect 'utility', the instruments can 'cross validate'. Confidence in one MAU instrument increases when it correlates with the other MAU instruments.
2. *Subjective Wellbeing (SWB) score*. Utility is commonly equated with SWB. This is not strictly correct as people's preferences do not always maximise happiness (Richardson, Maxwell et al. 2012). However the two constructs are undoubtedly related and high correlation with SWB is independently important if MAU instruments are to influence policy decisions. The three instruments used here – PWI, SWLS and IHS – are outlined in MIC Research Paper 1 (Richardson, Iezzi et al. 2012).
3. *Self TTO*. The concept and measurement of self TTO are also explained in MIC Research Paper 1. It is conceptually the same as a conventional TTO except that the health state evaluated is not 'external' as described to the respondent, but the respondent's own health state. The relationship between self and conventional TTO is the subject of ongoing research (Richardson, Iezzi et al. forthcoming).
4. *Disease-specific QoL instruments*. These are not utilised in the current report.

The Pearson correlation between MAU instruments, between MAU and non-MAU instruments and between non-MAU instruments are reported in Tables 4.1-4.6 and Fig 4.1. The Pearson correlation indicates the extent to which changes in one variable correspond with changes in another. It does not indicate that two variables are the same or even the same order of magnitude. The better measure of this is the intraclass correlation (ICC). This is reported in Table 4.7 and Figure 4.2. The difference is parenthesised by the relative score for the 15D. This has the highest average Pearson correlation but (reflecting significant differences in its predicted utility scores) it has the lowest ICC.

Overall the ICC reflects a poorer correspondence between instruments than the Pearson correlation. The imperfect correspondence is also illustrated by the use of R^2 coefficients in Figure 4.3 rather than Pearson correlation coefficients ($R^2 = \rho^2$). This is because a complete explanation of variation would imply $R^2 = 1$. The extent to which the R^2 falls short of 1.00 indicates the extent to which variance is explained by some unknown variable or variables.

Correlation with non-MAU instruments are shown in Table 4.7 and Figures 4.4–4.8. The low correlation between measures of utility and PWI and SWLS is in need of explanation. While preferences may differ from subjective wellbeing (SWB) their correlation might be expected to be higher than observed here.

Table 4.1 Pearson correlation between MAU (Public n=321)

Instrument	EQ-5D	HUI3	SF-6D	15D	QWB	AQoL-4D	AQoL-8D
EQ-5D	1	.648**	.550**	.689**	.504**	.627**	.532**
HUI3	.648**	1	.484**	.745**	.405**	.644**	.647**
SF-6D	.550**	.484**	1	.550**	.480**	.520**	.559**
15D	.689**	.745**	.550**	1	.491**	.658**	.654**
QWB	.504**	.405**	.480**	.491**	1	.468**	.424**
AQoL-4D	.627**	.644**	.520**	.658**	.468**	1	.713**
AQoL-8D	.532**	.647**	.559**	.654**	.424**	.713**	1
Ave	0.592	0.596	0.524	0.631	0.462	0.605	0.588

** Correlation is significant at the 0.01 level (2-tailed).

Table 4.2 Pearson correlation between MAU (Total n=1467)

	EQ-5D	HUI3	SF-6D	15D	QWB	AQoL-4D	AQoL-8D
EQ-5D	1	.798**	.740**	.815**	.646**	.755**	.763**
HUI3	.798**	1	.729**	.831**	.655**	.791**	.818**
SF-6D	.740**	.729**	1	.788**	.675**	.737**	.795**
15D	.815**	.831**	.788**	1	.708**	.800**	.840**
QWB	.646**	.655**	.675**	.708**	1	.658**	.649**
AQoL-4D	.755**	.791**	.737**	.800**	.658**	1	.854**
AQoL-8D	.763**	.818**	.795**	.840**	.649**	.854**	1
Ave	0.753	0.770	0.744	0.797	0.665	0.766	0.787

** Correlation is significant at the 0.01 level (2-tailed).

Table 4.3a Pearson correlations between MAUI instruments (Public n=321)

Instrument	PWI Sum	PWI	SWLS	IHS	ICECAP	Self-TTO	SF36
EQ5D	.138*	.176**	.209**	.087	.341**	.200**	.576**
HUI3	.177**	.265**	.246**	.194**	.373**	.306**	.577**
SF-6D	.286**	.313**	.250**	.203**	.408**	.387**	.868**
15D	.164**	.224**	.220**	.161**	.403**	.279**	.653**
QWB	.078	.162**	.156**	.099	.317**	.232**	.438**
AQoL-4D	.256**	.344**	.340**	.318**	.616**	.366**	.587**
AQoL-8D	.375**	.447**	.456**	.448**	.659**	.344**	.605**

*. Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

Table 4.3b Pearson correlations between MAUI and non-MAU instruments (Total n=1467)

	PWI	SWLS	HIS	Ave SWB	ICECAP	Self-TTO	SF36
EQ5D	.363**	.445**	.426**	.357**	.625**	.336**	.777**
HUI3	.416**	.496**	.488**	.438**	.678**	.337**	.774**
SF-6D	.418**	.504**	.464**	.384**	.636**	.380**	.930**
15D	.383**	.491**	.471**	.392**	.675**	.357**	.844**
QWB	.288**	.363**	.367**	.289**	.508**	.299**	.701**
AQoL-4D	.444**	.544**	.519**	.462**	.727**	.351**	.778**
AQoL-8D	.558**	.644**	.620**	.575**	.820**	.389**	.831**

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4.4 Pearson correlations between non-MAU and non-MAU instrument (Total n=321)

Non-MAUI	PWI Sum	PWI	SWLS	IHS	ICECAP	Self-TTO	SF36
PWI Sum	1	.748**	.681**	.642**	.411**	.194**	.230**
PWI	.748**	1	.687**	.640**	.483**	.187**	.302**
SWLS	.681**	.687**	1	.728**	.532**	.123*	.201**
IHS	.642**	.640**	.728**	1	.544**	.180**	.161**
ICECAP	.411**	.483**	.532**	.544**	1	.291**	.405**
Self-TTO	.194**	.187**	.123*	.180**	.291**	1	.368**
SF36	.230**	.302**	.201**	.161**	.405**	.368**	1

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4.5 Pearson correlations between non-MAU instruments (Total n=1467)

Non-MAUI	PWI Sum	PWI	SWLS	IHS	ICECAP	Self-TTO	SF36
PWI Sum	1	.791**	.751**	.708**	.578**	.248**	.420**
PWI	.791**	1	.788**	.735**	.645**	.276**	.530**
SWLS	.751**	.788**	1	.774**	.634**	.267**	.481**
IHS	.708**	.735**	.774**	1	.614**	.238**	.403**
ICECAP	.578**	.645**	.634**	.614**	1	.338**	.654**
Self-TTO	.248**	.276**	.267**	.238**	.338**	1	.377**
SF36	.420**	.530**	.481**	.403**	.654**	.377**	1

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 4.1 Average Pearson correlation with other MAU Instruments (Total n=1467)

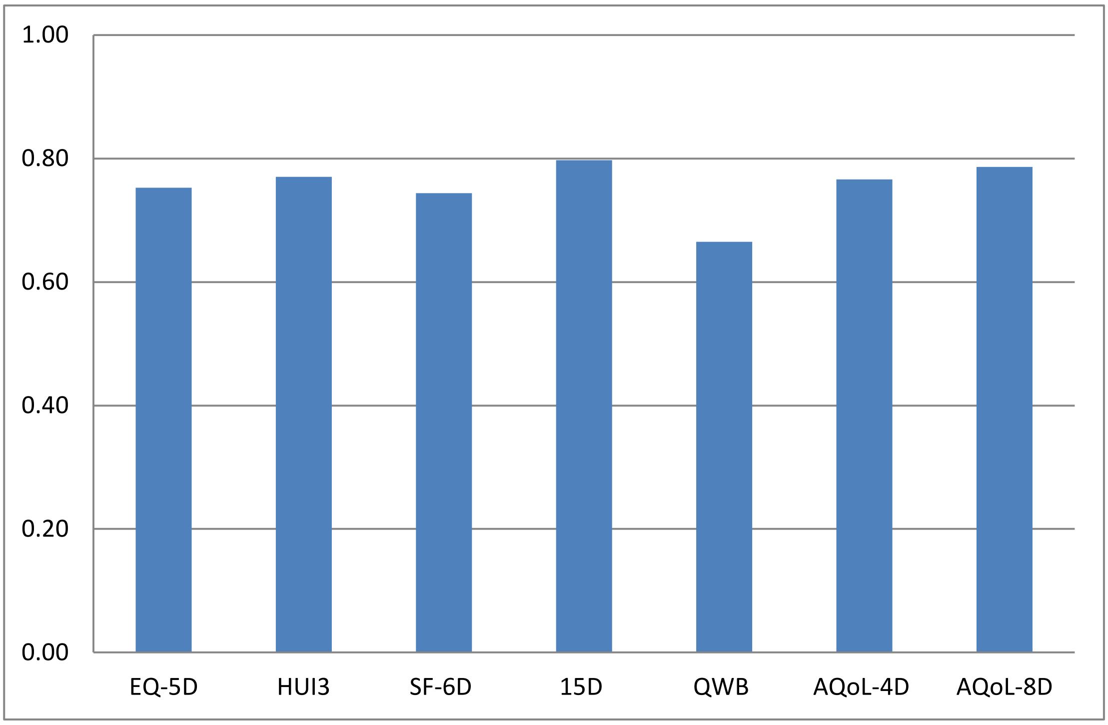


Table 4.6 Intraclass correlation between MAU instrument (Total n=1467)

	EQ5D	HUI3	SF-6D	15D	QWB	AQoL-4D	AQoL-8D
EQ-5D		0.78	0.65	0.60	0.54	0.69	0.76
HUI3	0.78		0.59	0.53	0.55	0.77	0.79
SF-6D	0.65	0.59		0.52	0.60	0.57	0.72
15D	0.60	0.53	0.52		0.34	0.42	0.61
QWB	0.54	0.55	0.60	0.34		0.58	0.55
AQoL-4D	0.69	0.77	0.57	0.42	0.58		0.78
AQoL-8D	0.76	0.79	0.72	0.61	0.55	0.78	
Ave	**0.67**	**0.67**	**0.61**	**0.50**	**0.53**	**0.63**	**0.70**

Figure 4.2 Average Intraclass correlation with other MAU Instruments (Total n=1467)

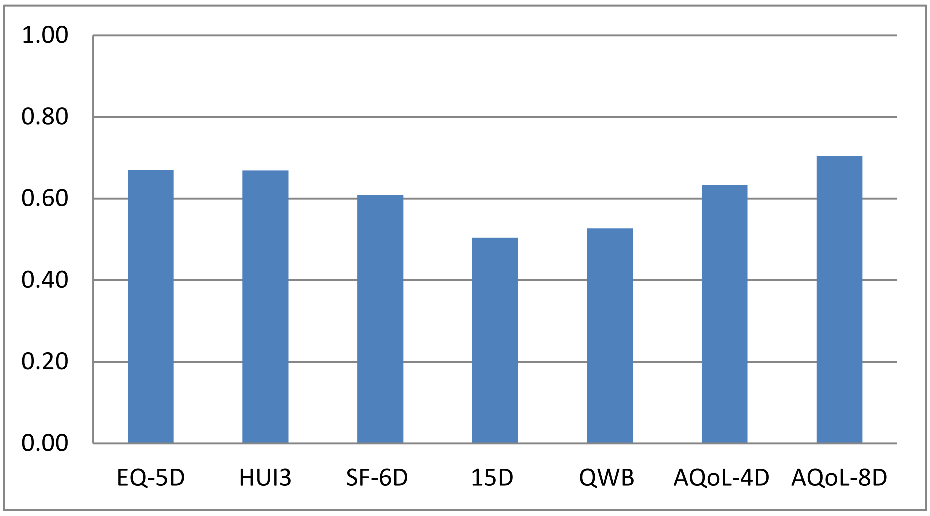


Figure 4.3 R^2: MAUI on MAU Instruments (Total n=1467

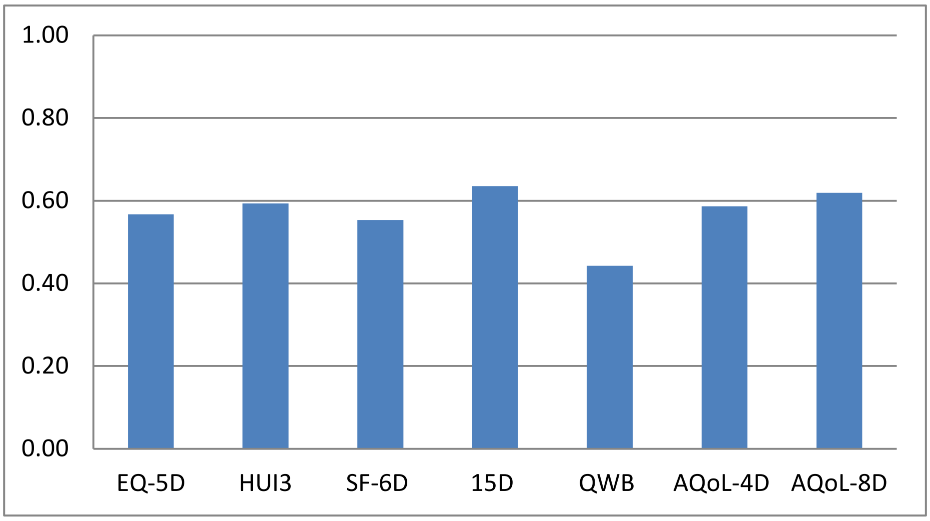


Figure 4.4 Pearson correlation of MAU instrument with PWI (Total n=1467)

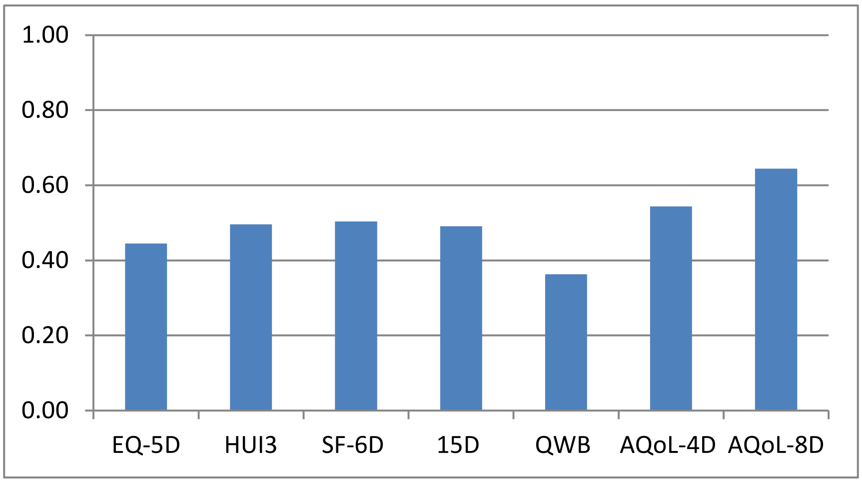


Figure 4.5 Pearson correlation of MAU instrument with SWLS (Total n=1467)

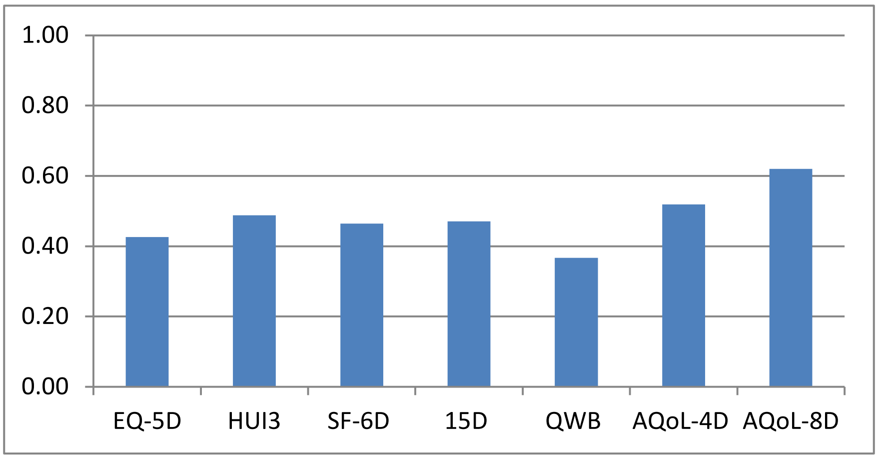


Figure 4.6 Pearson correlation of MAU instrument with Self-TTO (Total n=1467)

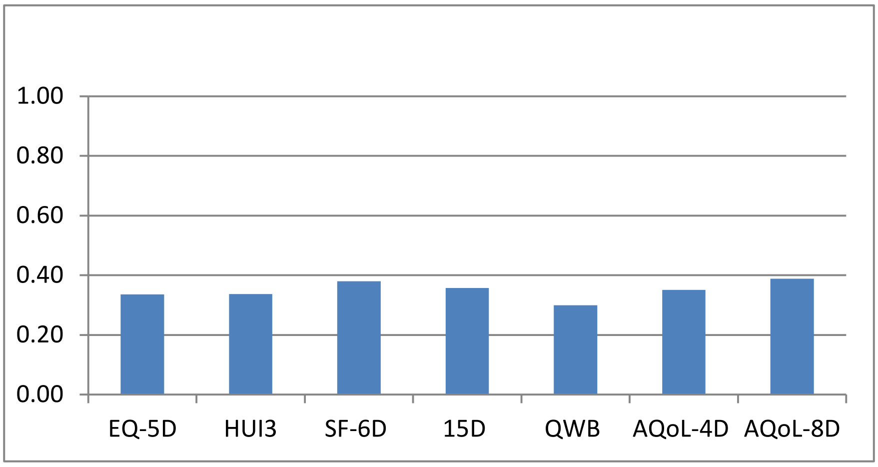


Figure 4.7 Pearson correlation of MAU instrument with SF-36 (Public n=321)

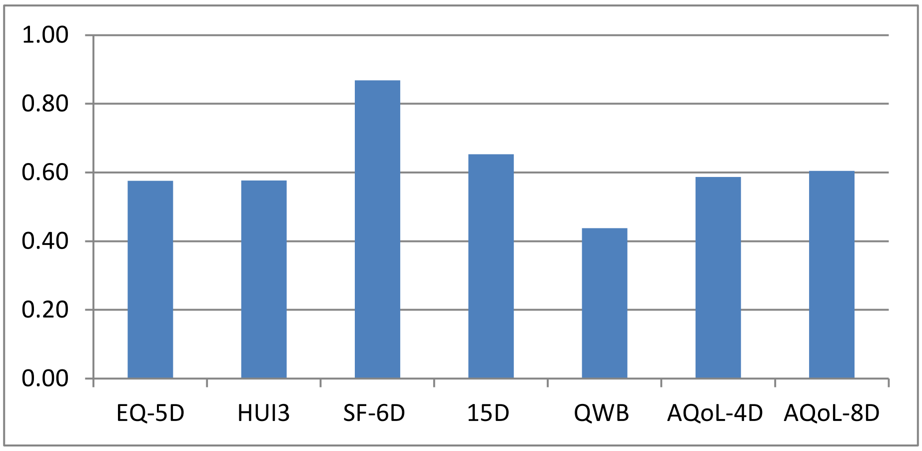
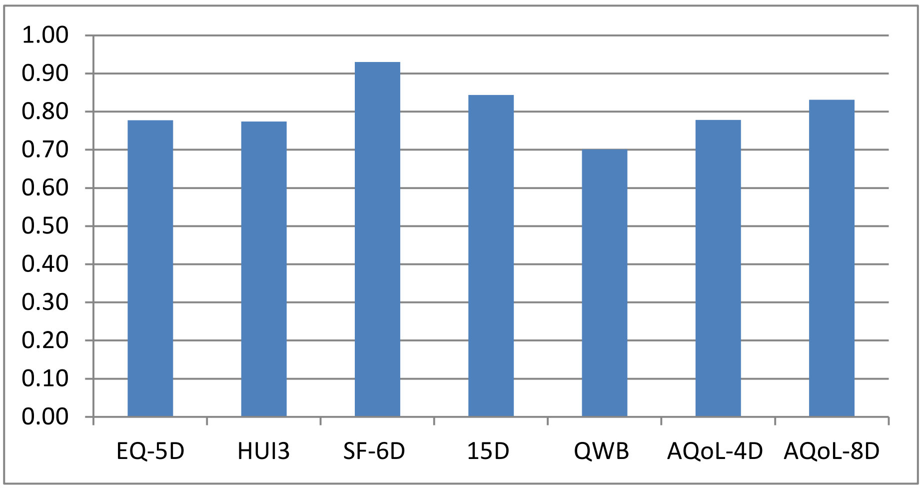


Figure 4.8 Pearson correlation of MAU instrument with SF-36 (Total n=1467)



5 Linear relationships

The MAU instruments were designed for use in cost utility analyses (CUA) in which, typically, utilities are measured before and after an intervention. This implies that it is the change in measured utilities, not their absolute values, which are important for validity. The comparative performance of the different instruments in this respect is not identified by either Pearson or intraclass correlations. It is however, easily measured with linear regression.

If instrument X is the criterion variable then the validity of the change predicted by instrument Y may be tested by the magnitude of the b coefficient in the linear relationship Y = a + bX. The absence of bias implies that b = 1.00. In the present case there is no criterion variable. However as with correlation, 'cross validation' may increase confidence: confidence rises if the b coefficients of an instrument are close to 1.00 in the linear relationships with the other MAU instruments. A technical problem which arises with this test is that, because both measured variables in the comparison are subject to error, the parameters will be sensitive to the choice of dependent and independent variable in OLS regressions. One solution to the problem is to use Geometric Mean Squares (GMS) regression. This is obtained by regressing Y on X then X on Y and deriving parameters from the geometric mean of the two regressions. Results are independent of the choice of dependent and independent variable. This technique was used in the present study.

Figure 5.1 reproduces the 21 pairwise GMS regressions, their scattergrams and the two GMS equations (Y on X; X on Y) using public data. Figure 5.2 gives the same results using the total sample.

Table 5.1 employs the results for the total sample to derive an average deviation away from b = 1 for each of the 6 regressions which include a particular MAUI. Depending upon the choice of left and right hand scale variable, 'b' may be greater than or less than 1.00. For consistency, the GMS regression was selected where b > 1. Thus from Figure 5.2 the linear relationship between the EQ-5D and HUI 3 for public respondents may be expressed either as
(1) EQ-5D = 0.15 + 0.833 HUI 3 or as (2) HUI 3 = -0.181 + 1.201 EQ-5D. Table 5.1 reports the b coefficient which is greater than 1.00 which, in this case, is 1.201. Table 5.1 indicates the instruments on the left and right of the selected equation using abbreviations (eg H = 1.201 EQ). From the bottom row in Table 5.1 the deviation for the MAUI vary from 34.5 percent (AQoL-8D) to 61.7 percent (15D). If these linear relationships were generally true (and not just for the present sample) the results would imply that the choice of AQoL-8D rather than one of the other six instruments would result in a 34.5 percent discrepancy in measured change. The choice of the 15D rather than one of the other six instruments would result in a 61.7 percent discrepancy.

Table 5.2 presents a different comparison using b coefficients. The bottom left of the table reports the b coefficients when instrument B is the left hand variable in the regression and instrument A is the right hand variable. The first figure is derived from the public regression and the second figure from the total sample. (Thus, in the public regression EQ-5D = 0.29 + 0.689 HUI 3 (Figure 5.1), the reported b coefficient is 0.689 rounded to 0.69. The b coefficient for the total sample, Figure 5.2 is 0.83. The difference between these coefficients (0.14) is shown in the top right hand side of Table 5.2 and the average difference involving each instrument is shown in the right hand column of the table. This is an indicator of the stability of the linear relationships involving an instrument when the severity of the health state changes. Thus for example, between the two samples the average of the 6 coefficients in equations with the EQ-5D as the dependent variable change by 39 percent.

Figure 5.1 Geometric regression results (Public n=321)

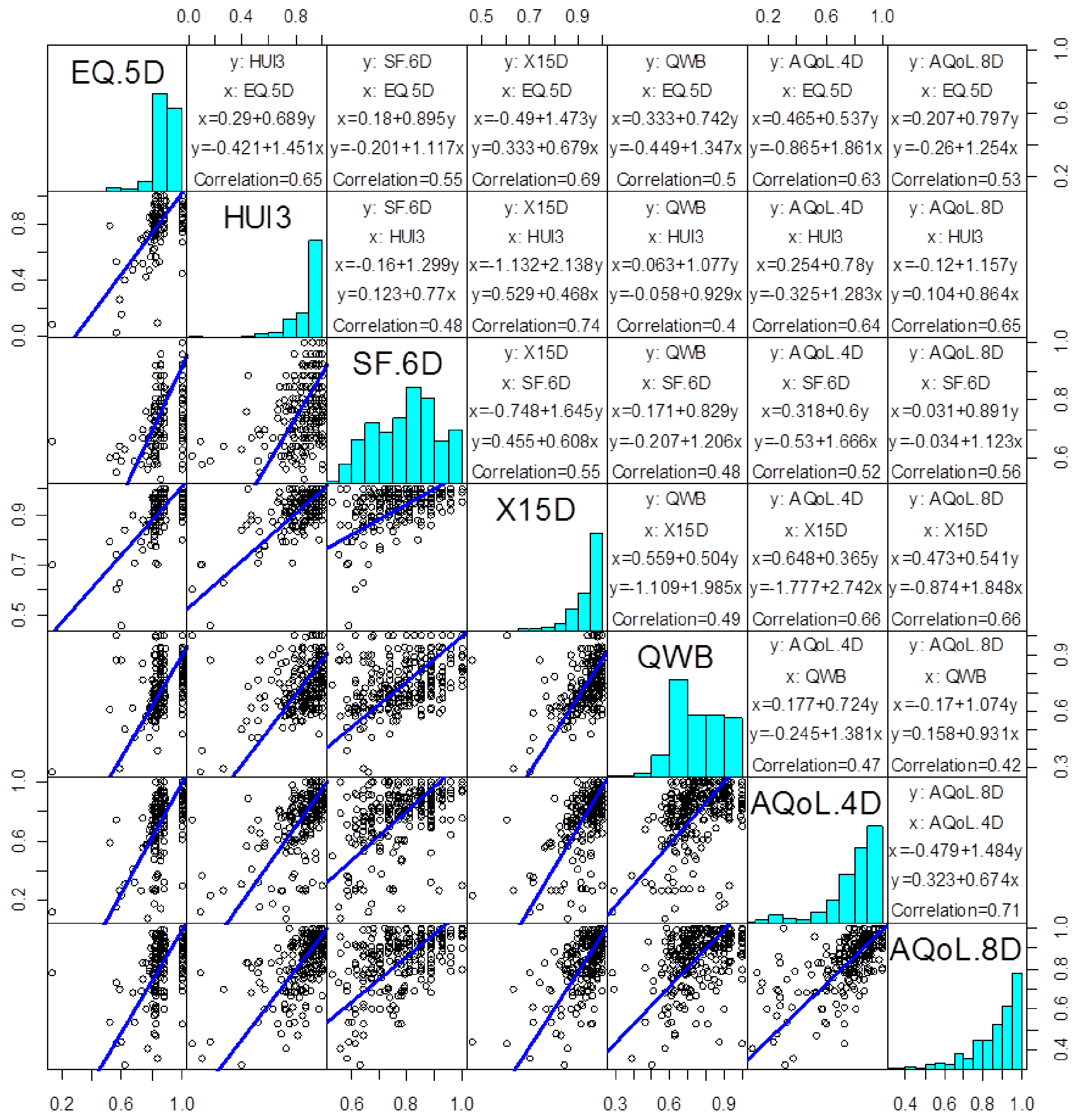


Figure 5.2 Geometric regression results (Total n=1467)

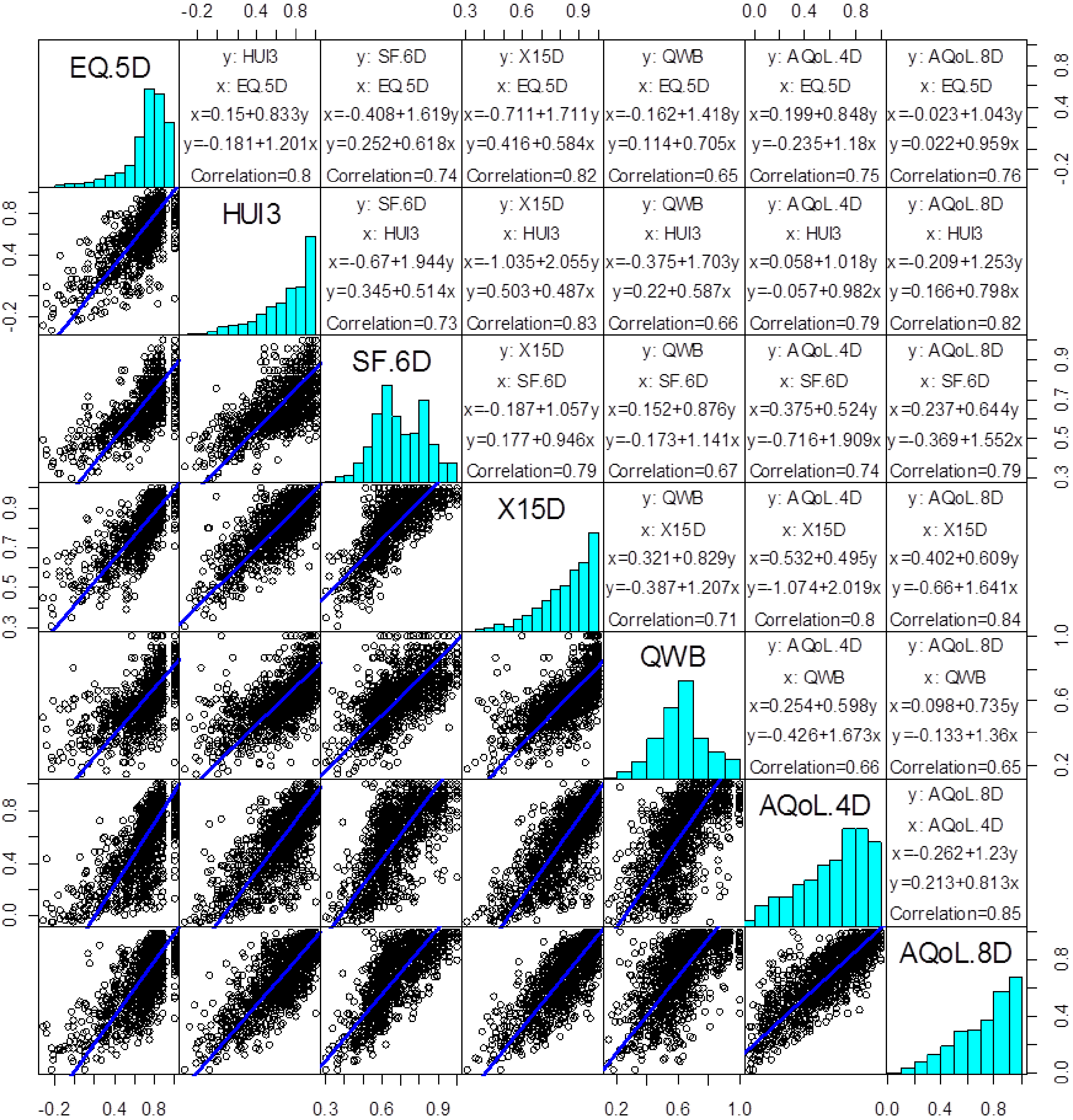


Table 5.1 Discrepancies in marginal change: slope, coefficient, b, in regression

(Instrument A=a+b instrument B)*

Instrument	EQ-5D	HUI3	SF-6D	15D	QWB	AQoL-4D	AQoL-8D
EQ-5D (EQ)	1.00						
HUI3 (H)	H=1.20(EQ)	1.00					
SF-6D (SF)	EQ=1.62(SF)	H=1.94(SF)	1.00				
15D (D)	EQ=1.71(D)	H=2.06(D)	SF=1.06(D)	1.00			
QWB (Q)	EQ=1.42(Q)	H=1.70(Q)	Q=1.14(SF)	Q=1.21(D)	1.00		
AQoL-4D (A4)	A4=1.18(EQ)	H=1.02(A4)	A4=1.91(SF)	A4=2.02(D)	A4=1.67(Q)	1.00	
AQoL-8D (A8)	EQ=1.04(A8)	H=1.25(A8)	A8=1.55(SF)	A8=1.64(D)	A8=1.36(Q)	A4=1.23(A8)	1.00
Ave % Diff	36.2	52.8	53.7	61.7	41.7	50.5	34.5

(NB Constant terms in the equations have been dropped)

*Equations arranged to obtain b>1 as a consistent index of deviation (Geometric Mean Regressions permit this)

Table 5.2 Difference in marginal change: public vs total (instrument A=a+b instrument B)

Instrument B	EQ5D		HUI3		SF6D		15D		QWB		AQoL-4D		AQoL-8D	
	Pub	Tot	Pub	Tot	Pub	Tot	Pub	Tot	Pub	Tot	Pub	Tot	Pub	Tot
EQ-5D	1.00		(.14)		(.72)		(.24)		(.68)		(.31)		(.24)	
HUI3	.69	.83	1.00		(.64)		(.08)		(.62)		(.24)		(.09)	
SF-6D	.90	1.62	1.30	1.94	1.00		(.59)		(.05)		(.08)		(.05)	
15D	1.47	1.71	2.14	2.06	1.65	1.06	1.00		(.33)		(.13)		(.07)	
QWB	.74	1.42	1.08	1.70	.83	.88	.50	.83	1.00		(.12)		(.33)	
AQoL-4D	.54	.85	.78	1.02	.60	.52	.37	.50	.72	.60	1.00		(.25)	
AQoL-8D	.80	1.04	1.16	1.25	.89	.64	.54	.61	1.07	.74	1.48	1.23	1.00	
Ave	0.39		0.30		0.36		0.24		0.36		0.19		0.17	

6 Instrument content (sensitivity)

Each MAU defines a ‘construct’. Results in this section seek to identify how clearly related dimensions of health/wellbeing are to the MAU constructs. Conversely the results seek to determine how sensitive the MAU constructs are to the dimensions. The dimensions used in the study are obtained from the SF-36 and AQoL-8D which have been independently shown to have construct validity (Richardson, Elsworth et al. 2011). Additionally, the widely used and validated SWB instruments, the PWI and SWLS are employed as is the yet unvalidated Self TTO. Similar results may be obtained for the IHS.

Ceiling effects: From Table 6.1a ceiling effects differ greatly. In the public sample the maximum score (the ‘ceiling’) was obtained by 42.1 percent and 7.2 percent on the EQ-5D and SF-6D respectively. Amongst the 255 respondents with an EQ-5D score of 1.00 the average scores on other instruments varied from 0.85 and 0.78 for SF-6D and QWB respectively to 0.97 for HUI 3.

Floor effects: Table 6.1b reveals similar differences in floor effects. For example, when EQ-5D < 0.4 its average score is 0.20. HUI 3, SF-6D and AQoL-8D have average scores of 0.21, 0.51 and 0.32 respectively. When HUI 3 < 0.4 average values for EQ-5D, HUI 3, SF-6D and AQoL-8D are 0.42, 0.19, 0.54 and 0.42 respectively.

Table 6.1a Ceiling effects (MAU = 1) Ave value of other MAUI when an MAU=1.0 (Public n=321)

MAU=1	Average value								
	EQ5D	**HUI3**	**SF-6D**	**15D**	**QWB**	**AQoL-4D**	**AQoL-8D**	**n**	**(%**
EQ5D	--	.95	.87	.98	.84	.92	.93	135	42.1
HUI3	.97	--	.86	.98	.84	.94	.94	69	21.5
SF-6D	1.00	.98	--	.99	.91	.95	.98	23	7.2
15D	.98	.97	.88	--	.88	.94	.96	75	23.4
QWB	.96	.94	.89	.98	--	.93	.93	29	9.0
AQoL-4D	.99	.98	.90	.99	.86	--	.97	58	18.1
AQoL-8D	1.00	.98	.93	1.00	.90	.95	--	26	8.1

Table 6.1b Ceiling effects (MAU=1.0) Ave value of other MAUI when an MAU=1.0 (Total n=1467)

MAU=1	Average value								
	EQ5D	**HUI3**	**SF-6D**	**15D**	**QWB**	**AQoL-4D**	**AQoL-8D**	**n**	**(%)**
EQ5D	--	.93	.85	.97	.80	.89	.92	251	17.1
HUI3	.97	--	.84	.98	.82	.92	.94	110	7.5
SF-6D	1.00	.98	--	.99	.88	.94	.98	31	2.1
15D	.98	.96	.87	--	.87	.94	.95	108	7.4
QWB	.95	.94	.87	.98	--	.92	.92	42	2.9
AQoL-4D	.98	.98	.88	.99	.86	--	.97	89	6.1
AQoL-8D	1.00	.98	.91	.99	.87	.96	--	39	2.7

Table 6.1c Floor effects Ave value of other MAUI when an MAU < .40 (Total n=1467)

MAU<0.4	**EQ5D**	**HUI3**	**SF-6D**	**15D**	**QWB**	**AQoL-4D**	**AQoL-8D**	**n**	**(%)**
EQ5D	<u>.18</u>	.19	.50	.61	.44	.20	.35	130	8.9
HUI3	.38	<u>.17</u>	.52	.63	.46	.24	.39	213	14.5
SF-6D	.09	.04	<u>.36</u>	.51	.38	.12	.23	20	1.4
15D	.00	-.15	.39	<u>.36</u>	.40	.00	.13	4	0.3
QWB	.40	.28	.52	.64	<u>.32</u>	.25	.42	108	7.4
AQoL-4D	.47	.37	.56	.68	.49	<u>.22</u>	.44	323	22.0
AQoL-8D	.34	.19	.50	.61	.44	.18	.28	148	10.1

Correlation with summary measures: Table 6.2 and Figure 6.1 report the correlation between MAU scores and the physical and psycho-social summary scores derived from the SF-36 and AQoL-8D. With one exception the correlations with the AQoL-8D (non-utility) super-dimensions are greater than with the SF-36 summary scores. With two exceptions correlation between MAU instruments and the physical summary score is greater than with the psycho-social summary scale. The Table suggests three groups of instruments. First, EQ-5D, HUI and 15D are relatively very sensitive to physical health (particularly EQ-5D). AQoL-8D is relatively very sensitive to psycho-social health. SF-6D, QWB and AQoL-4D are between these polar cases.

Table 6.2 Correlation of instruments with SF-36, AQoL-8D physical and psycho-social scales (Total n=1467)

SF-36/ AQoL-8D dimension	**EQ-5D**	**HUI3**	**SF-6D**	**15D**	**QWB**	**AQoL-4D**	**AQoL-8D**
SF-36	.777**	.774**	.930**	.844**	.701**	.778**	.831**
PCS	.661**	.621**	.673**	.694**	.599**	.569**	.514**
MCS	.503**	.547**	.709**	.574**	.445**	.607**	.758**
AQoL-8D	.763**	.818**	.795**	.840**	.649**	.854**	1
PSD	.787**	.797**	.728**	.822**	.676**	.790**	.775**
MSD	.587**	.640**	.726**	.683**	.577**	.736**	.889**

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 6.1 Correlation with summary scores of SF-36 (PCS and MCS) and AQoL-8D (PSD and MSD)

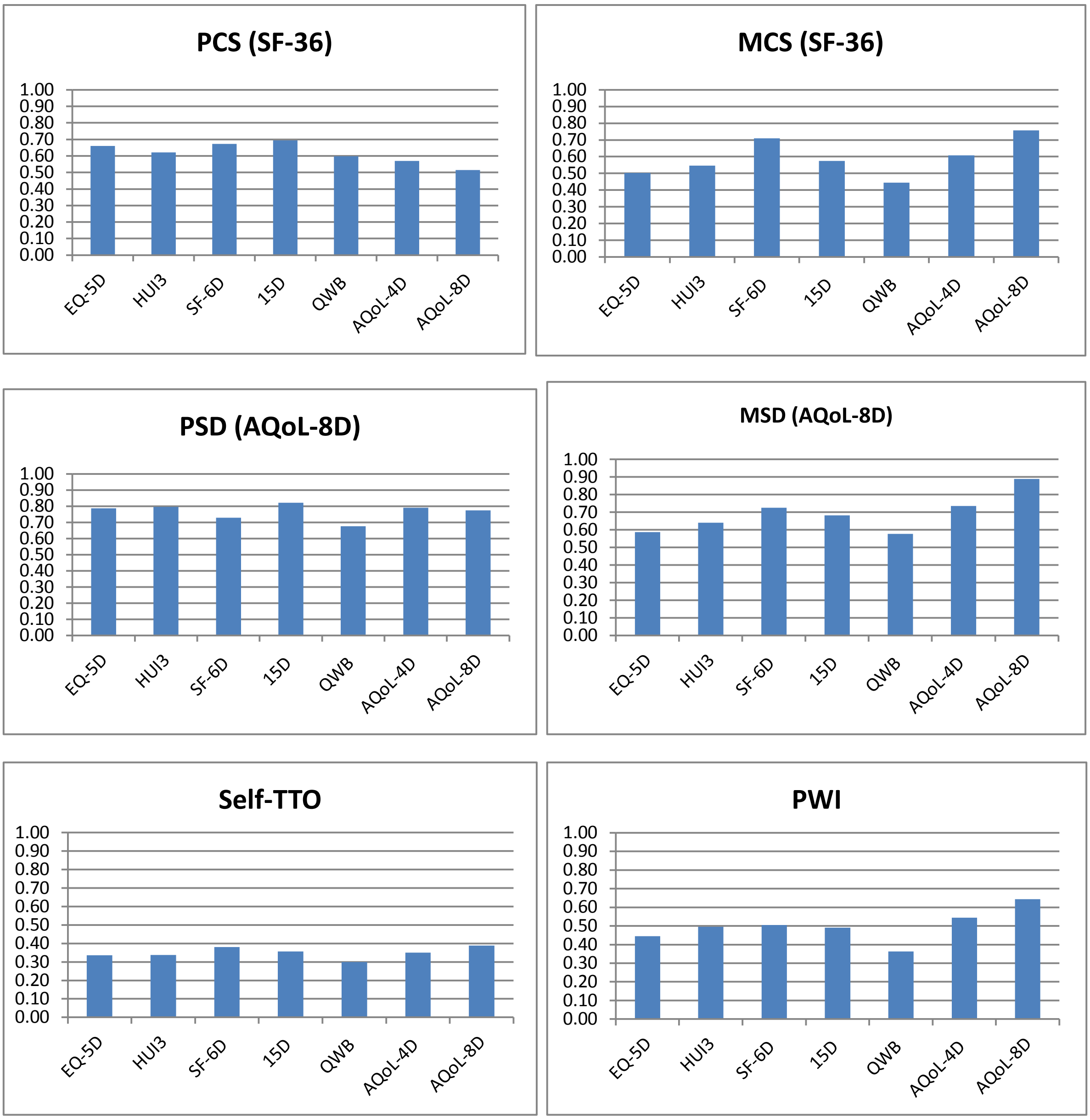
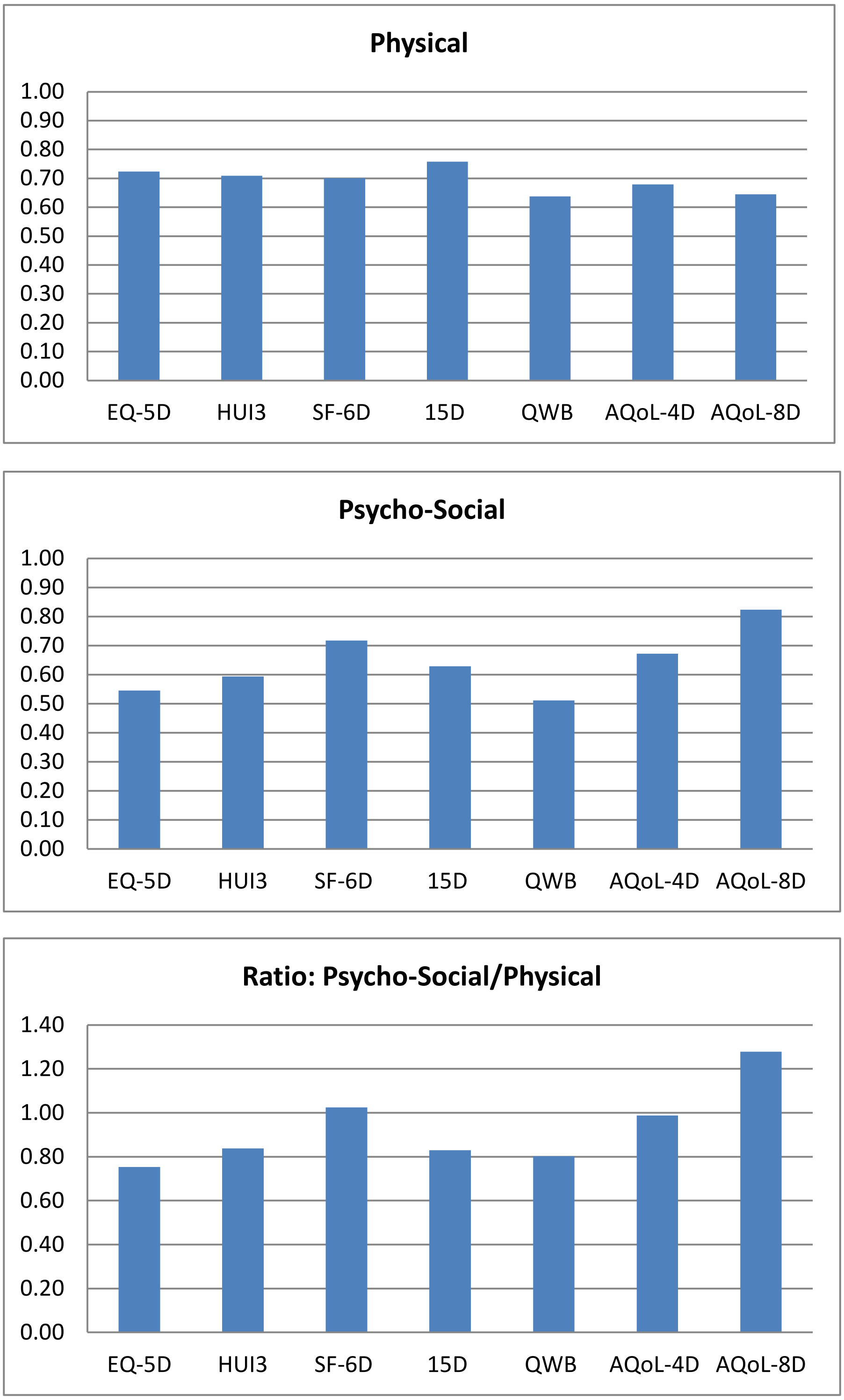


Figure 6.2 Comparison of Summary Physical and Psycho-Social Dimensions (Average SF-36 and AQoL-8D summary scores)



Split half analysis: Table 6.3 reports results from a comparison of two split halves of the full sample. Each MAU was used, in turn, to rank observations on the basis of which they were divided into a top and bottom half. Dimension and SWB scores were calculated for both halves. The table reports the ratio of these scores. Higher ratios indicate greater sensitivity of an instrument to a dimension or SWB.

Sensitivity to dimensions: Tables 6.4a, 6.4b; 6.5a, 6.5b and Figure 6.3a, 6.3b report beta coefficients from the regression of MAU scores on dimension scores. The coefficients show the change in the MAU score with a one standard deviation change in the dimension score. MAU scores are measured in standard deviations (of the MAU score) to allow comparison of sensitivity. This avoids the confusion of a large standard deviation with instrument sensitivity. Thus, for example, the 15D compresses scores. But this is offset in the calculation of beta coefficients by a correspondingly small standard deviation. A larger beta coefficient suggests greater sensitivity.

Tables 6.4a and 6.5a report results from regressions with a single explanatory variable. Because of its correlation with other explanatory variables (dimensions) interpretation of the beta score is ambiguous. Table 6.3b and 6.4b use multiple regressions to obtain the standardised beta. In principle this means that the beta coefficients represent the effect of the dimension after standardising for other dimensions in the regression. From the regressions employing the SF-36 dimensions (Table 6.4b) a one sd increase in each dimension would result in a 1.02 sd increase in the EQ-5D (ie Σ_i $Beta_i$) of which 68 percent would be attributable to physical function and pain (ie (0.28+41)/1.02). Mental health would contribute 24 percent and vitality 2 percent. The same increase in the dimension scores would increase AQoL-8D by 1.15 sd of which 38 percent would be attributable to mental health, 19 percent to vitality and only 21 percent to pain and physical function. This suggests that in the AQoL-8D the effects of pain and physical function may be largely mediated through psycho-social factors.

The percentage contribution to total change following a one sd increase in every dimension using data from Tables 6.4a,b and 6.5b is shown in the pie charts, Figure 6.4.

Table 6.3a Ratio of scores in top and bottom 50% of total sample, ranked by MAUI (SF-36 dimensions)

Ranking MAUI	SF-36 dimensions									
	GH	PF	RP	BP	VT	SF	RE	MH	PCS	MCS
EQ5D	1.71	1.69	3.24	1.75	1.72	1.46	1.96	1.33	1.47	1.23
HUI3	1.69	1.66	2.99	1.62	1.75	1.47	2.06	1.38	1.41	1.27
SF-6D	1.71	1.64	4.05	1.70	1.88	1.67	3.05	1.45	1.41	1.41
15D	1.80	1.66	3.47	1.64	1.86	1.51	2.14	1.37	1.45	1.29
QWB	1.62	1.58	2.71	1.58	1.66	1.41	1.95	1.29	1.39	1.23
AQoL-4D	1.64	1.56	2.76	1.58	1.79	1.50	2.15	1.41	1.35	1.32
AQoL-8D	1.71	1.50	2.80	1.57	2.02	1.55	2.41	1.53	1.31	1.42

Table 6.3b Ratio of scores in top and bottom 50% of total sample, ranked by MAUI (AQoL-8D dimensions, SWB and Self-TTO)

Ranking MAUI	AQoL-8D dimension										SWB		Self-TTO
	IL	Hap	MH	Cop	Rel	SW	Pain	Sen	PSD	MSD	PWI	SWLS	Self-TTO
EQ5D	1.29	1.25	1.28	1.28	1.23	1.22	1.37	1.10	1.52	1.86	1.23	1.25	1.22
HUI3	1.29	1.30	1.31	1.31	1.29	1.25	1.33	1.13	1.50	2.00	1.28	1.34	1.23
SF-6D	1.26	1.30	1.35	1.33	1.29	1.27	1.31	1.10	1.44	2.11	1.28	1.31	1.27
15D	1.29	1.29	1.34	1.35	1.29	1.26	1.32	1.12	1.50	2.11	1.29	1.32	1.26
QWB	1.25	1.23	1.26	1.27	1.22	1.21	1.28	1.10	1.42	1.80	1.20	1.25	1.20
AQoL-4D	1.27	1.32	1.35	1.32	1.36	1.28	1.31	1.15	1.50	2.21	1.31	1.34	1.24
AQoL-8D	1.25	1.43	1.48	1.41	1.43	1.36	1.29	1.13	1.45	2.78	1.38	1.45	1.29

Key

GH=general health; PF = physical functioning; RP = role limit physical; BP = bodily pain; VT = vitality; SF = social functioning; RE = role limit emotional; MH = mental health; PCS =physical component summary; MCS = mental component summary; IL = independent living; Hap = happiness; Cop = coping; Rel = relationships; SW = self worth; Pain=pain; Sen=senses; MSD = mental super dimension; PSD = physical super dimension;
PWI = Personal Wellbeing Index; SWLS = Satisfaction with Life Survey; TTO = Time- trade-off;

Table 6.4a Sensitivity to SF-36 dimensions: Beta coefficient and R^2 from the regression of MAU on single dimensions of the SF-36 (Total n=1467)

(MAU = a + b Dim)

SF-36 dimension	EQ5D	HUI3	SF-6D	15D	QWB	AQoL-4D	AQoL-8D
GH							
Beta	0.64	0.64	0.71	0.74	0.61	0.64	0.69
R^2	0.40	0.41	0.50	0.54	0.37	0.41	0.48
PF							
Beta	0.68	0.66	0.69	0.71	0.59	0.60	0.56
R^2	0.46	0.44	0.47	0.51	0.35	0.36	0.32
RP							
Beta	0.57	0.56	0.73	0.63	0.56	0.55	0.55
R^2	0.32	0.31	0.53	0.40	0.31	0.31	0.30
BP							
Beta	0.75	0.67	0.75	0.70	0.60	0.64	0.63
R^2	0.56	0.45	0.56	0.49	0.36	0.41	0.40
VT							
Beta	0.62	0.64	0.76	0.72	0.60	0.67	0.79
R^2	0.38	0.40	0.58	0.52	0.36	0.45	0.63
SF							
Beta	0.66	0.67	0.81	0.71	0.55	0.69	0.74
R^2	0.43	0.45	0.65	0.51	0.30	0.47	0.55
RE							
Beta	0.50	0.50	0.71	0.56	0.48	0.53	0.60
R^2	0.25	0.25	0.51	0.31	0.23	0.29	0.36
MH							
Beta	0.58	0.63	0.72	0.63	0.47	0.66	0.82
R^2	0.34	0.40	0.51	0.39	0.22	0.44	0.67
PCS							
Beta	**0.66**	**0.62**	**0.67**	**0.69**	**0.60**	**0.57**	**0.51**
R^2	**0.44**	**0.39**	**0.45**	**0.48**	**0.36**	**0.32**	**0.26**
MCS							
Beta	**0.50**	**0.55**	**0.71**	**0.57**	**0.45**	**0.61**	**0.76**
R^2	**0.25**	**0.30**	**0.50**	**0.33**	**0.20**	**0.37**	**0.57**

Table 6.4b Sensitivity to SF-36 dimensions: Beta coefficient from the regression of MAU on all dimensions of the SF-36 (Total n=1467)

$$(MAU = a + \sum_{u-1}^{8} b1Dimi)$$

(SF-36 dimension)	EQ5D	HUI3	SF-6D	15D	QWB	AQoL-4D	AQoL-8D
(GH) Beta	0.06	0.07	0.03	0.18	0.15	0.09	0.11
t	2.52	3.04	2.36	9.01	5.17	3.62	6.02
(PF) Beta	0.28	0.31	0.14	0.27	0.19	0.19	0.10
t	12.49	13.30	9.58	13.57	6.85	7.91	5.75
(RP) Beta	-0.07	-.03 (ns)	0.14	.01 (ns)	0.06 (ns)	.00 (ns)	-.03 (ns)
t	-2.87		9.88				
(BP) Beta	0.41	0.24	0.21	0.17	0.18	0.18	0.14
t	18.23	10.25	14.72	8.51	6.44	7.66	8.22
(VT) Beta	.002 (ns)	.01 (ns)	0.12	0.15	0.18	0.09	0.21
t			7.47	6.90	5.78	3.55	10.75
(SF) Beta	0.08	0.11	0.18	0.11	.00 (ns)	0.14	0.09
t	3.36	4.40	11.29	5.21		5.42	4.81
(RE) Beta	0.02 (ns)	-.01 (ns)	0.18	0.05	0.10	.02 (ns)	0.04
t			13.94	2.51	3.80		2.65
(MH) Beta	0.24	0.32	0.20	0.15	0.04	0.30	0.44
t	10.12	13.11	13.29	6.99	1.26	11.85	24.06
R^2	**0.70**	**0.67**	**0.88**	**0.76**	**0.52**	**0.66**	**0.82**
F	**418**	**377**	**1311**	**574**	**196**	**348**	**820**

Key

ns = not significant

1 Same as Table a

2 Direct comparison of the overall fit with the fit of SF-6D is invalid as it is derived from the SF-36

Table 6.5a Sensitivity to AQoL-8D dimensions: Beta coefficient and R^2 from the regression of MAU on single dimensions of the AQoL-8D

(MAU = a + b Dim)

AQoL-8D dimension	EQ5D	HUI3	SF-6D	15D	QWB	AQoL-4D	AQoL-8D
IL							
Beta	0.72	0.72	0.67	0.78	0.61	0.68	0.67
R^2	0.52	0.52	0.44	0.61	0.37	0.46	0.45
Hap							
Beta	0.58	0.65	0.65	0.64	0.50	0.70	0.88
R^2	0.34	0.42	0.43	0.41	0.25	0.48	0.77
MH							
Beta	0.57	0.60	0.68	0.65	0.53	0.68	0.84
R^2	0.32	0.36	0.46	0.42	0.28	0.46	0.71
Cop							
Beta	0.67	0.70	0.73	0.76	0.58	0.73	0.90
R^2	0.45	0.50	0.53	0.58	0.34	0.54	0.81
Rel							
Beta	0.45	0.55	0.58	0.55	0.45	0.68	0.76
R^2	0.21	0.30	0.33	0.30	0.20	0.46	0.58
SW							
Beta	0.61	0.67	0.66	0.67	0.51	0.70	0.89
R^2	0.37	0.44	0.43	0.45	0.26	0.49	0.79
Pain							
Beta	0.79	0.72	0.69	0.73	0.60	0.68	0.68
R^2	0.62	0.52	0.47	0.53	0.36	0.47	0.47
Sen							
Beta	0.36	0.49	0.36	0.49	0.39	0.54	0.54
R^2	0.13	0.23	0.13	0.24	0.15	0.29	0.29
PSD							
Beta	**0.79**	**0.80**	**0.73**	**0.82**	**0.68**	**0.79**	**0.78**
R^2	**0.62**	**0.64**	**0.53**	**0.68**	**0.46**	**0.62**	**0.60**
MSD							
Beta	**0.59**	**0.64**	**0.73**	**0.68**	**0.58**	**0.74**	**0.89**
R^2	**0.34**	**0.41**	**0.53**	**0.47**	**0.33**	**0.54**	**0.79**

Table 6.5b Sensitivity to AQoL-8D dimensions: Beta coefficient from the regression of MAU on all dimensions of the AQoL-8D

$$(MAU = a + \sum_{u-1}^{8} b1Dimi)$$

(AQoL-8D dimension)	EQ5D	HUI3	SF-6D	15D	QWB	AQoL-4D	AQoL-8D
(IL) Beta	0.25	0.26	0.23	0.36	0.27	0.19	0.10
t	12.47	13.10	10.65	20.93	9.59	10.00	23.48
(Pain) Beta	0.46	0.32	0.27	0.23	0.23	0.26	0.18
t	23.96	16.41	13.13	13.92	8.70	14.69	45.70
(Sen) Beta	.01 (ns)	0.14	-.01 (ns)	0.11	0.09	0.20	0.13
t		9.40		8.83	4.51	14.00	43.39
(Hap) Beta	0.08	0.14	.02 (ns)	-.01 (ns)	-.04 (ns)	0.08	0.19
t	2.89	4.93				3.06	32.10
(MH) Beta	0.10	.03 (ns)	0.25	0.17	0.19	0.08	0.17
t	4.06		9.85	7.99	5.81	3.35	33.96
(Cop) Beta	0.13	0.07	0.19	0.25	0.13	.05 (ns)	0.17
t	4.55	2.50	6.46	10.48	3.47		29.67
(SW) Beta	0.10	0.14	.04 (ns)	0.05	.00 (ns)	0.11	0.25
t	3.79	5.37		2.28		4.48	48.51
(Rel) Beta	-0.10	.01 (ns)	.03 (ns)	-.04 (ns)	.03 (ns)	0.22	0.07
t	-4.77					11.34	17.23
R^2	**0.74**	**0.74**	**0.70**	**0.81**	**0.51**	**0.78**	**0.99**
F	**525**	**521**	**434**	**766**	**188**	**633**	**16681**

(ns) = not significant

1 Beta coefficients are the change in the dependent variable, measured in standard deviations (of the dependent) when the independent variable changes by one standard deviation (after standardising for other variables in the regression). They allow direct comparison of the importance of independent variables.

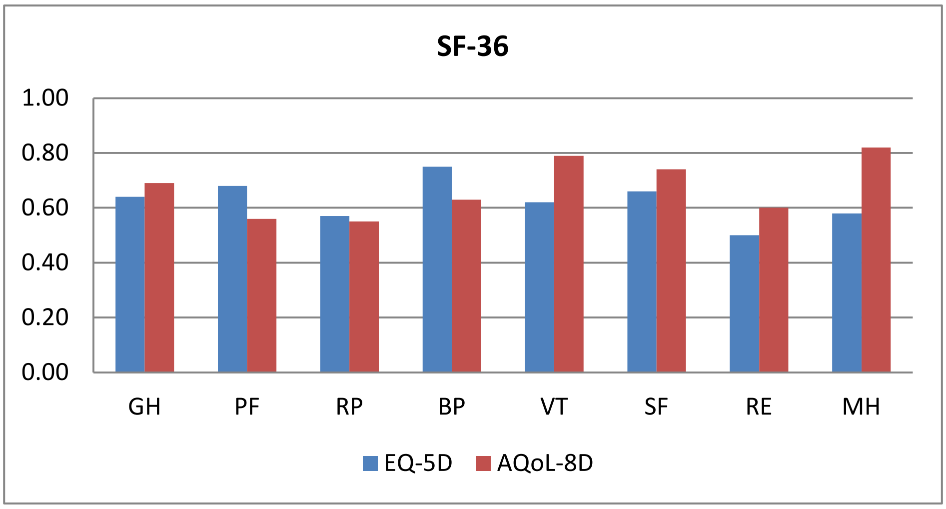
2 Direct comparison of the overall fit with the fit of AQoL-8D is invalid as it is an (exponential) function of the dimensions

Table 6.6 Instrument content: regression of MAU on non-MAU instruments

Dependent	EQ5D	HUI3	SF-6D	15D	QWB	AQoL-4D	AQoL-8D	
a	0.39	0.25	0.47	0.63	0.44	0.14	0.26	**PWI**
b	0.52	0.69	0.36	0.33	0.30	0.75	0.72	
Beta	0.45	0.50	0.50	0.49	0.36	0.54	0.64	
R^2	0.20	0.25	0.25	0.24	0.13	0.30	0.42	
F	362	478	498	465	222	617	1041	
MAU = a + b SWLS								
a	0.45	0.32	0.52	0.66	0.46	0.23	0.34	**SWLS**
b	0.45	0.62	0.21	0.29	0.28	0.65	0.63	
Beta	0.43	0.49	0.46	0.47	0.37	0.52	0.62	
R^2	0.18	0.24	0.22	0.22	0.13	0.27	0.38	
F	324	457	403	418	228	541	913	
MAU = a + b Self-TTO								
a	0.53	0.45	0.56	0.72	0.50	0.38	0.50	**Self-TTO**
b	0.25	0.31	0.18	0.16	0.16	0.31	0.28	
Beta	0.34	0.34	0.38	0.36	0.30	0.35	0.39	
R^2	0.11	0.11	0.14	0.13	0.09	0.12	0.15	
F	186	187	247	214	143	206	261	

Figure 6.3 Effect of SD change in dimension on standardised score (beta coefficient)

(a) Content of EQ-5D vs AQoL-8D (SF-36 Dimensions)



(b) Contrast of EQ-5D vs 15D (AQoL-8D Dimensions)

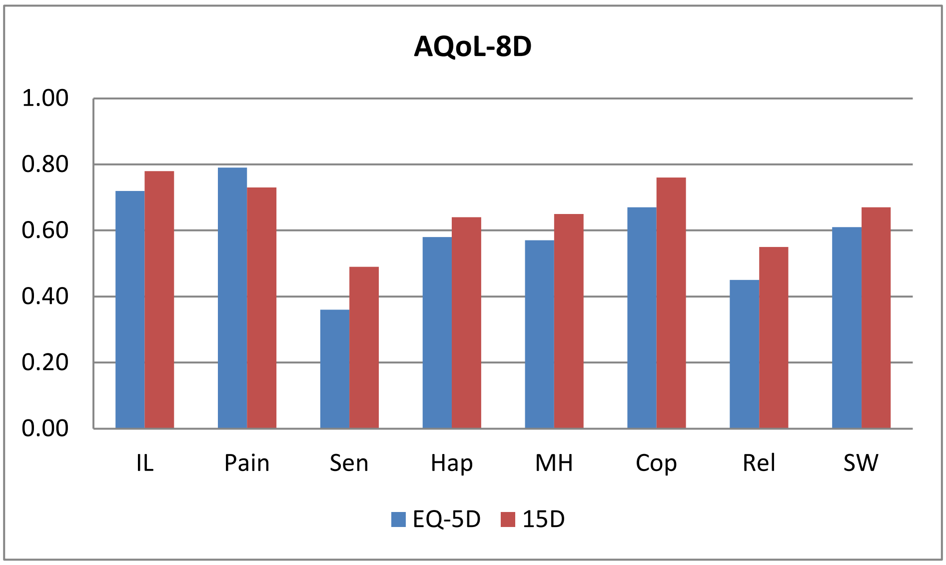


Figure 6.4 Instrument content: Disaggregated by AQoL-8D dimensions

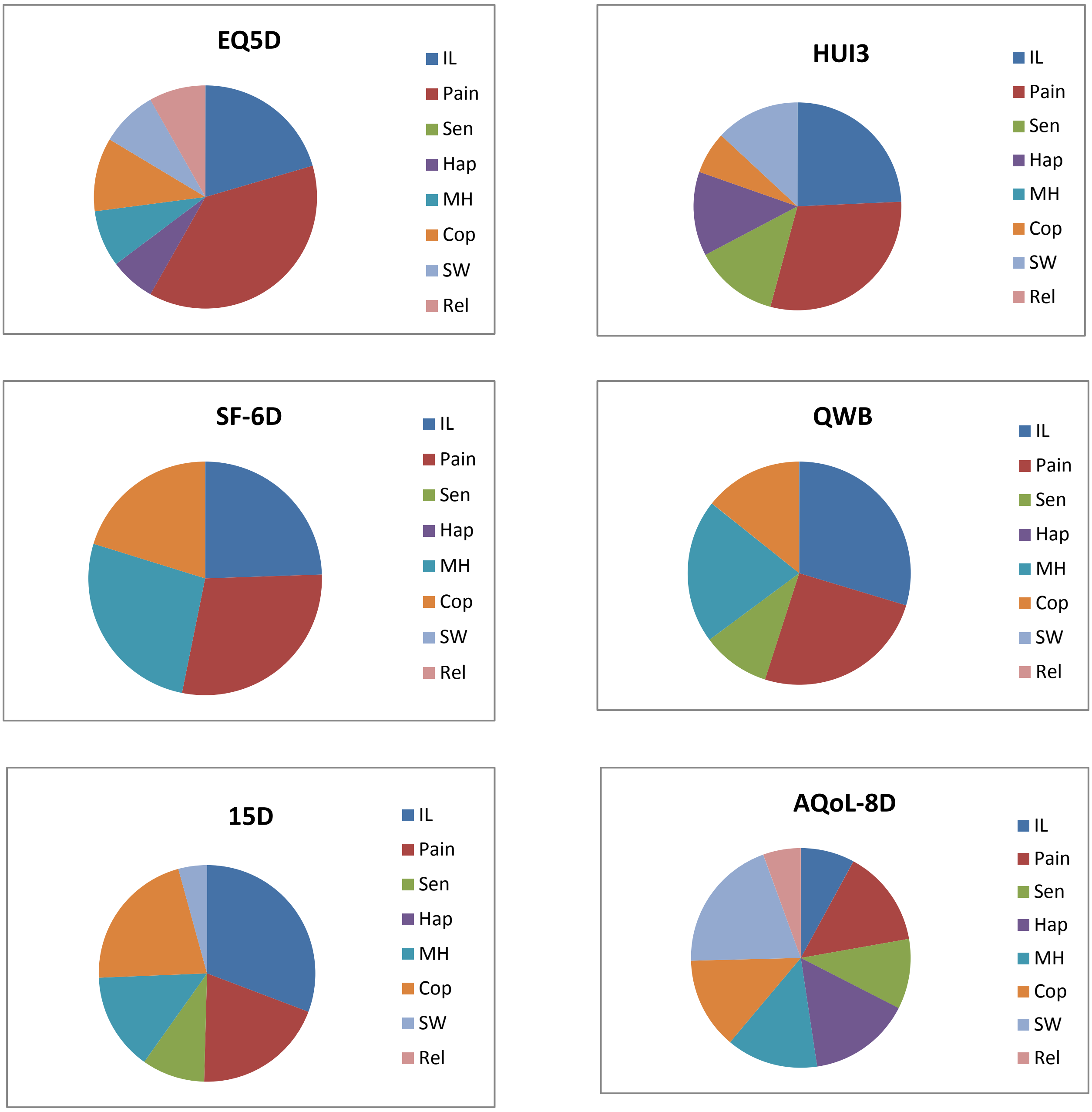
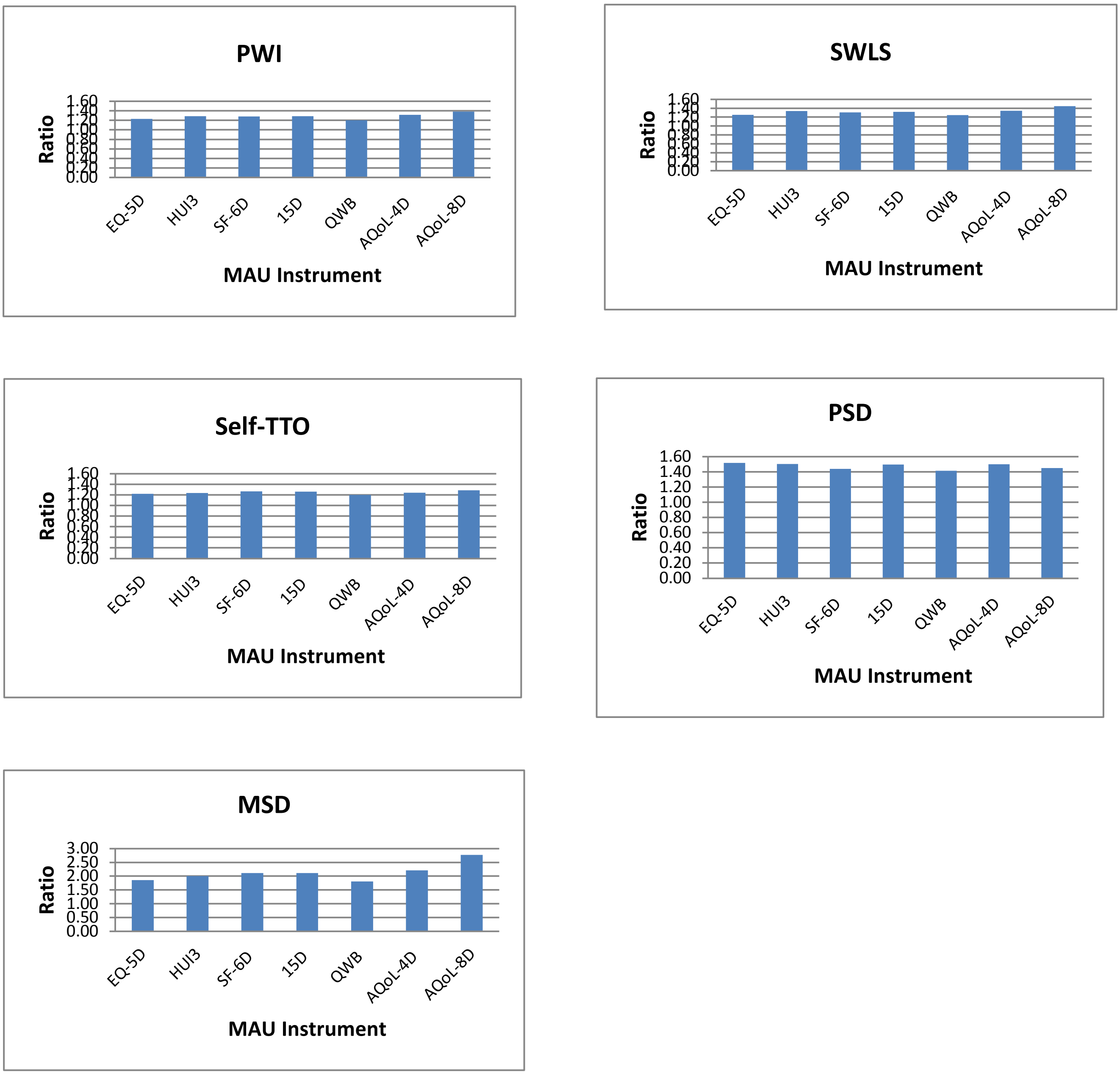


Figure 6.5 Split half analysis: Ratios of values in top/bottom half of population ranked by instrument



7 Pairwise comparison of instruments

The GMS regressions reported earlier were employed to help explain differences between the instruments' content. The residual from the regression of one instrument upon another was correlated with each of the major dimensions and non-MAU instruments A positive correlation between the residual of Y regressed upon X and a dimension, D or index, I, indicates a greater sensitivity of the instrument Y to dimension D or index I. A negative correlation implies the greater sensitivity of instrument X. Since regressions were calculated using geometric mean squares the results are independent of the choice of dependent and independent variable.

Results are given in Tables 7.1 and 7.2. The frequency distributions of the residuals are given in Appendix 3. To put the magnitude of the correlation coefficients in perspective, the average correlation between *unstandardised instruments* is 0.75; that is, a correlation between a *residual and a single dimension* of 0.25 is 0.25/0.75 or one third of this magnitude which is quantitatively large.

. Figure 7.1 presents the correlation results from Table 7.1 and 7.2. Table 7.3 summarises the results and therefore the implications of the data for the relative sensitivity of instruments.

Table 7.1 Dimension and instrument correlations with MAU residuals (Total 1467) and SF-36 dimensions and SWB instruments

	SF-36 Dimensions															
Residuals	**Gen**	**Phys**	**RoleP**	**Pain**	**SumP**	**Ind**	**Vital**	**Social**	**RoleE**	**MH**	**SumM**	**PWI**	**SWLS**	**IHS**	**Self-TTO**	**SF-36**
EQ5D-1.619*SF6D	-.098**	-0.014	-.216**	0	-0.014	.080**	-.197**	-.205**	-.296**	-.183**	-.284**	-.080**	-.053*	-0.037	-.060*	-.210**
EQ5D-.833*HUI3	-0.005	0.021	0.015	.114**	.064*	0.001	-0.03	-0.024	-0.009	-.078**	-.069**	-.081**	-.098**	-.127**	-0.002	0.003
EQ5D-1.418*QWB	0.034	.104**	0.016	.172**	.074**	.135**	0.024	.125**	0.025	.131**	.069**	.097**	.070**	.081**	0.044	.089**
EQ5D- 1.711*15D	-.164**	-.056*	-.101**	.070**	-.052*	-.093**	-.172**	-.089**	-.098**	-.069**	-.116**	-.075**	-.074**	-.056*	-0.035	-.111**
EQ5D- .848*AQOL4D	-0.006	.107**	0.021	.154**	.132**	.064*	-.082**	-0.044	-.054*	-.115**	-.148**	-.142**	-.134**	-.149**	-0.022	-0.002
EQ5D- 1.043*AQoL8D	-.082**	.167**	0.035	.163**	.213**	.070**	-.258**	-.121**	-.152**	-.337**	-.371**	-.291**	-.283**	-.317**	-.078**	-.081**
HUI3- 1.944*SF6D	-.091**	-0.031	-.225**	-.099**	-.069**	.078**	-.167**	-.180**	-.283**	-.112**	-.218**	-0.009	0.033	.074**	-.058*	-.209**
HUI3-1.703*QWB	0.038	.089**	0.005	.087**	0.026	.136**	0.047	.145**	0.032	.192**	.122**	.160**	.145**	.179**	0.046	.088**
HUI3- 2.055*15D	-.166**	-.081**	-.123**	-.052*	-.125**	-.099**	-.147**	-.068**	-.094**	0.01	-0.046	0.009	0.029	.080**	-0.035	-.119**
HUI3- 1.018*AQOL4D	-0.002	.096**	0.009	.055*	.081**	.068**	-.059*	-0.025	-0.05	-0.05	-.093**	-.075**	-0.049	-0.037	-0.022	-0.005
HUI3- 1.253*AQoL8D	-.089**	.168**	0.024	.065*	.175**	.078**	-.263**	-.113**	-.165**	-.303**	-.351**	-.247**	-.220**	-.228**	-.087**	-.096**
SF6D- .876*QWB	.123**	.121**	.210**	.180**	.090**	.069**	.202**	.314**	.291**	.300**	.326**	.173**	.120**	.117**	.100**	.281**
SF6D- 1.057*15D	-0.045	-0.037	.145**	.066*	-0.034	-.176**	.058*	.143**	.236**	.137**	.206**	0.019	-0.011	-0.012	0.034	.129**
SF6D- .524*AQOL4D	.091**	.117**	.235**	.149**	.141**	-0.018	.116**	.160**	.242**	.070**	.139**	-.058*	-.077**	-.108**	0.039	.207**
SF6D- .644*AQOL8D	0.022	.195**	.282**	.176**	.245**	-0.014	-.055*	.101**	.170**	-.156**	-.079**	-.222**	-.244**	-.299**	-0.015	.150**
15D-.829*QWB	.167**	.159**	.098**	.134**	.123**	.223**	.163**	.209**	.106**	.200**	.168**	.167**	.136**	.133**	.076**	.186**
15D- .495*AQOL4D	.152**	.174**	.122**	.105**	.198**	.162**	.076**	0.038	0.036	-.059*	-.051*	-.084**	-.076**	-.111**	0.01	.106**
15D- .609*AQOL8D	.077**	.264**	.153**	.124**	.316**	.187**	-.128**	-0.05	-.078**	-.334**	-.326**	-.272**	-.263**	-.324**	-.056*	0.022
QWB - .598*AQOL4D	-0.04	-0.015	0.001	-0.044	0.036	-.083**	-.094**	-.165**	-.071**	-.231**	-.196**	-.219**	-.184**	-.209**	-.063*	-.092**
QWB - .735*AQOL8D	-.101**	0.033	0.013	-0.038	.101**	-.078**	-.236**	-.224**	-.150**	-.408**	-.374**	-.336**	-.302**	-.341**	-.108**	-.155**
AQOL4D-1.123*AQOL8D	0.015	.169**	.108**	.113**	.189**	.119**	-.097**	0.023	-0.028	-.154**	-.164**	-.085**	-.090**	-.121**	-0.008	0.034

(1) Source: Figure 5.2 (constants omitted as they do not affect correlation)

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Table 7.2 Dimension and instrument correlations with MAU residuals (Total 1467) and AQoL-8D dimensions

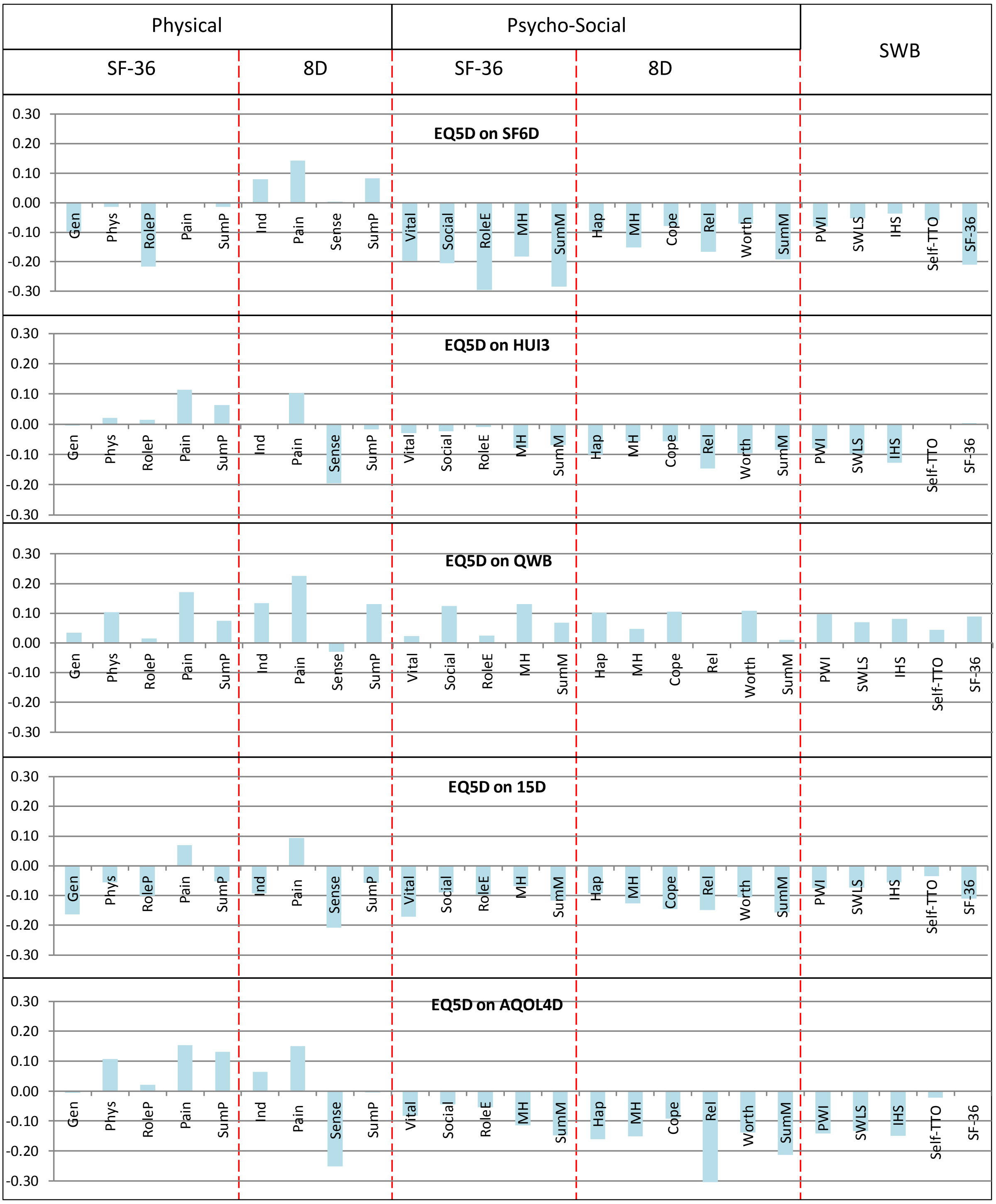
Residuals	AQoL-8D Dimensions									
	Ind	Pain	Sense	SumP	Hap	MH	Cope	Rel	Worth	SumM
EQ5D-1.619*SF6D	.080**	.143**	0.003	.083**	-.097**	-.152**	-.079**	-.166**	-.071**	-.192**
EQ5D-.833*HUI3	0.001	.102**	-.195**	-0.017	-.101**	-.056*	-.056*	-.146**	-.096**	-.085**
EQ5D-1.418*QWB	.135**	.226**	-0.029	.131**	.102**	0.047	.105**	0.002	.108**	0.01
EQ5D- 1.711*15D	-.093**	.094**	-.208**	-.057*	-.092**	-.127**	-.146**	-.149**	-.105**	-.157**
EQ5D- .848*AQOL4D	.064*	.150**	-.252**	-0.004	-.161**	-.152**	-.093**	-.319**	-.139**	-.214**
EQ5D- 1.043*AQoL8D	.070**	.151**	-.257**	0.016	-.426**	-.398**	-.335**	-.446**	-.410**	-.440**
HUI3- 1.944*SF6D	.078**	.053*	.170**	.095**	-0.008	-.101**	-0.029	-0.037	0.013	-.115**
HUI3-1.703*QWB	.136**	.151**	.119**	.145**	.180**	.090**	.148**	.114**	.183**	.075**
HUI3- 2.055*15D	-.099**	-0.013	-0.005	-0.043	0.014	-.072**	-.092**	0.004	-0.005	-.072**
HUI3- 1.018*AQOL4D	.068**	.063*	-.081**	0.012	-.075**	-.110**	-0.046	-.202**	-.056*	-.148**
HUI3- 1.253*AQoL8D	.078**	.065*	-.089**	0.035	-.380**	-.396**	-.324**	-.355**	-.367**	-.414**
SF6D- .876*QWB	.069**	.108**	-0.033	.063*	.193**	.185**	.180**	.151**	.176**	.182**
SF6D- 1.057*15D	-.176**	-.071**	-.197**	-.146**	0.021	0.05	-0.049	0.045	-0.02	.065*
SF6D- .524*AQOL4D	-0.018	0.002	-.246**	-.087**	-.059*	0.005	-0.012	-.143**	-.064*	-0.016
SF6D- .644*AQOL8D	-0.014	0.002	-.279**	-.075**	-.349**	-.256**	-.271**	-.291**	-.360**	-.257**
15D-.829*QWB	.223**	.174**	.133**	.190**	.185**	.152**	.231**	.120**	.202**	.136**
15D- .495*AQOL4D	.162**	.077**	-.078**	.052*	-.088**	-0.045	0.039	-.209**	-.052*	-.084**
15D- .609*AQOL8D	.187**	.084**	-.088**	.083**	-.418**	-.346**	-.249**	-.381**	-.384**	-.365**
QWB - .598*AQOL4D	-.083**	-.103**	-.183**	-.137**	-.240**	-.176**	-.185**	-.272**	-.228**	-.192**
QWB - .735*AQOL8D	-.078**	-.102**	-.181**	-.118**	-.452**	-.373**	-.380**	-.367**	-.444**	-.371**
AQOL4D-1.123*AQOL8D	.119**	.112**	.089**	.156**	-.200**	-.180**	-.167**	-0.033	-.205**	-.145**

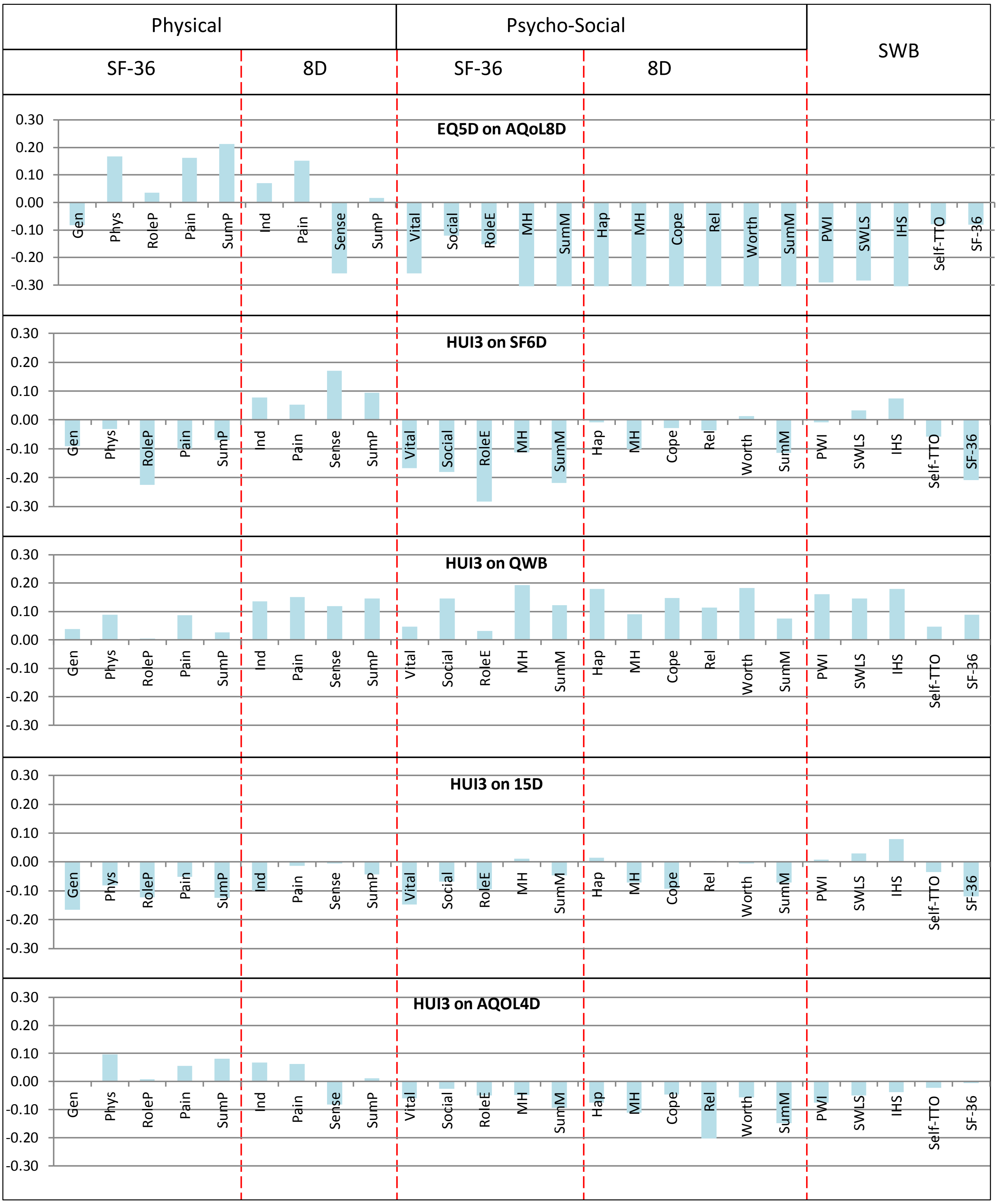
[(1)] Source: Figure 5.2 (constants omitted as they do not affect correlation)

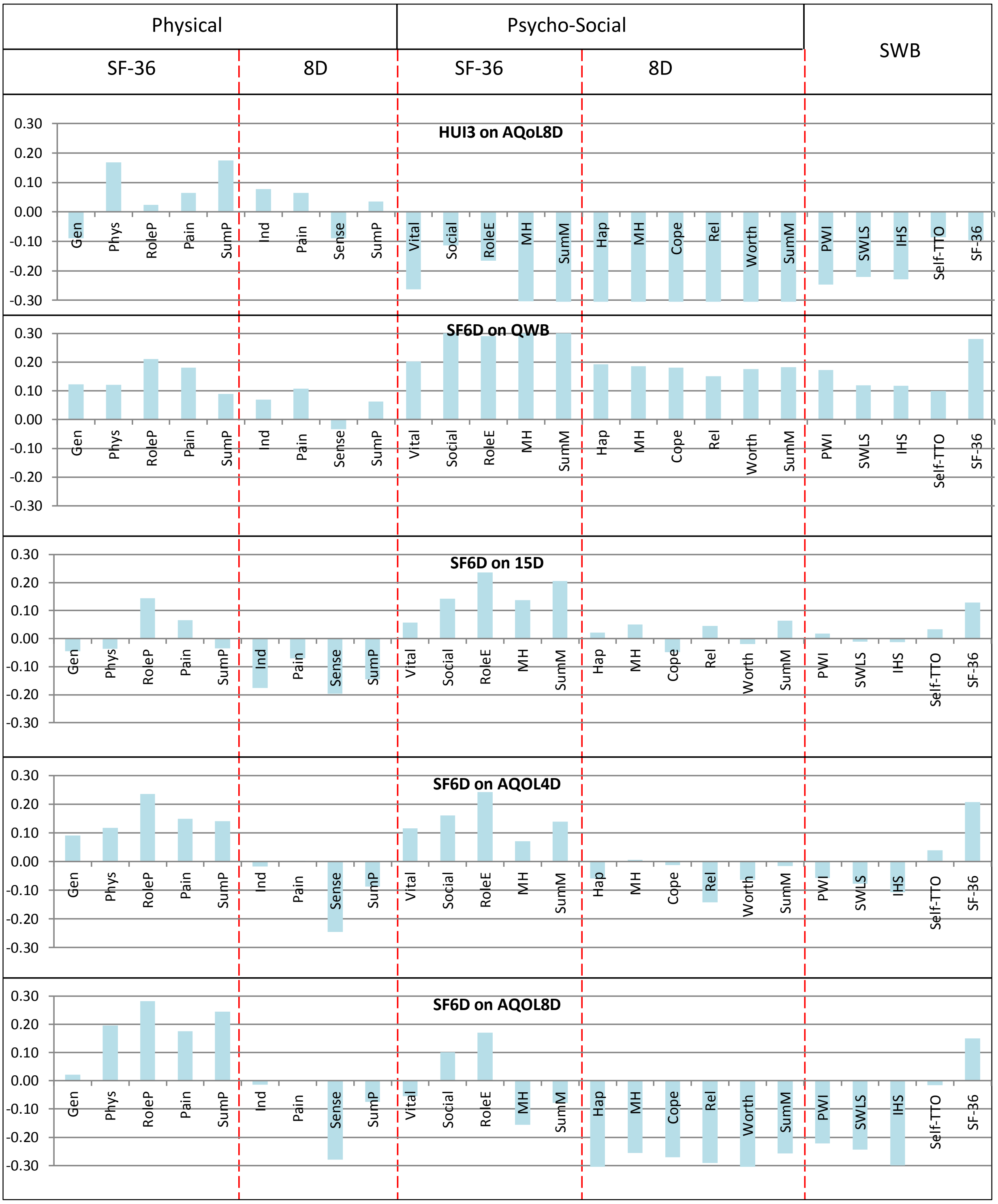
** Correlation is significant at the 0.01 level (2-tailed).

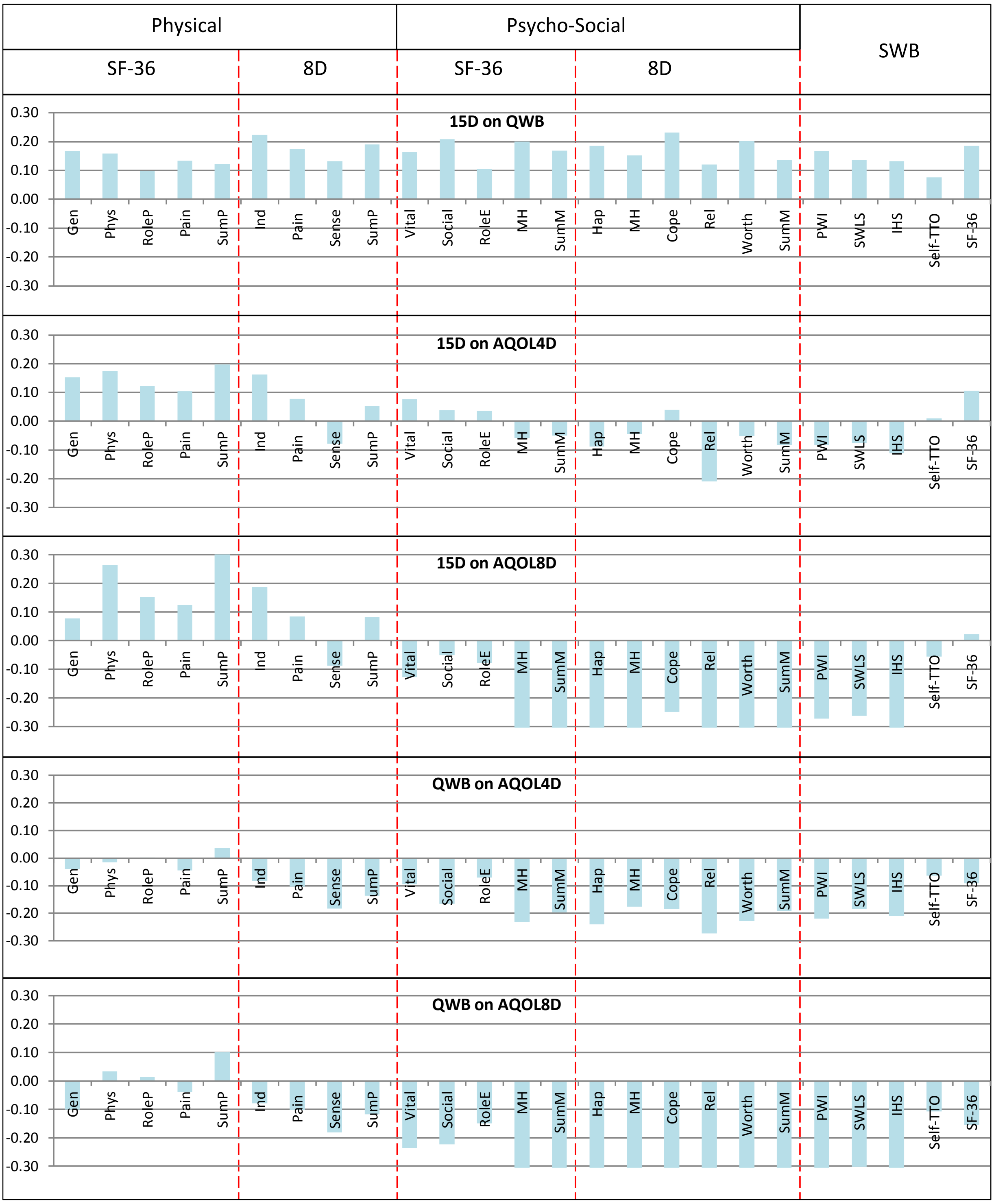
* Correlation is significant at the 0.05 level (2-tailed).

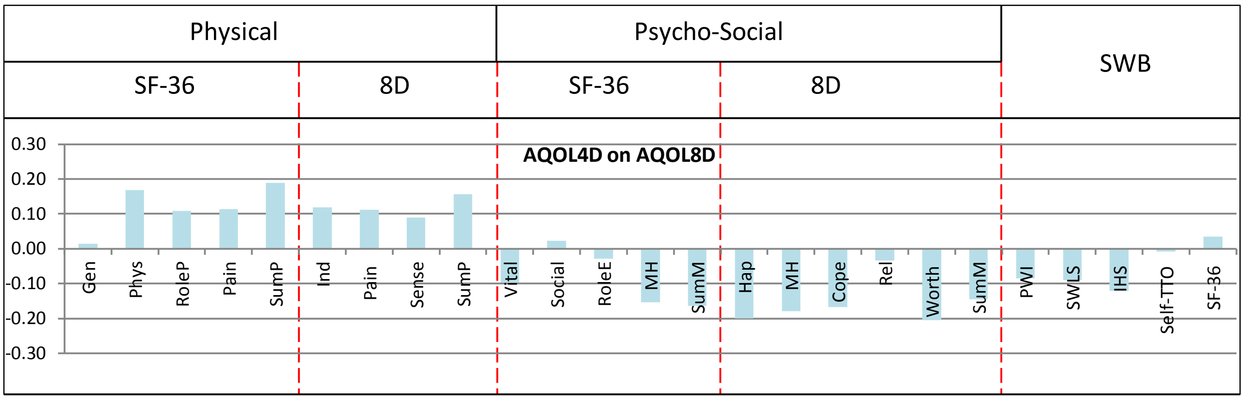
Figure 7.1 Pairwise instrument sensitivity: Correlations of dimension scores with the residual of one MAU instrument regressed upon a second MAU instrument




Physical
Psycho-Social
SWB
SF-36
8D
SF-36
8D
EQ5D on AQoL8D
HUI3 on SF6D
HUI3 on QWB
HUI3 on 15D
HUI3 on AQOL4D
0.30
0.20
0.10
0.00
-0.10
-0.20
-0.30
Gen
Phys
RoleP
Pain
SumP
Ind
Pain
Sense
SumP
Vital
Social
RoleE
MH
SumM
Hap
MH
Cope
Rel
Worth
SumM
PWI
SWLS
IHS
Self-TTO
SF-36


Physical
Psycho-Social
SWB
SF-36
8D
SF-36
8D
HUI3 on AQoL8D
SF6D on QWB
SF6D on 15D
SF6D on AQOL4D
SF6D on AQOL8D
0.30
0.20
0.10
0.00
-0.10
-0.20
-0.30
Gen
Phys
RoleP
Pain
SumP
Ind
Pain
Sense
SumP
Vital
Social
RoleE
MH
SumM
Hap
MH
Cope
Rel
Worth
SumM
PWI
SWLS
IHS
Self-TTO
SF-36


Physical
Psycho-Social
SWB
SF-36
8D
SF-36
8D
15D on QWB
15D on AQOL4D
15D on AQOL8D
QWB on AQOL4D
QWB on AQOL8D
0.30
0.20
0.10
0.00
-0.10
-0.20
-0.30
Gen
Phys
RoleP
Pain
SumP
Ind
Pain
Sense
SumP
Vital
Social
RoleE
MH
SumM
Hap
MH
Cope
Rel
Worth
SumM
PWI
SWLS
IHS
Self-TTO
SF-36



Key

Gen=general health; Phy = physical function; Role P = role limit physical; BP = bodily pain; Vit = vitality; Soc = social functioning; Role E = role limit emotional; MH = mental health; Cope = Coping; Rel = relationships; Worth = self worth; Pain=pain; Sen=senses; MSD = mental super dimension; PSD = physical super dimension; SF-36: 8 dimensions – 4 physical; 4 psycho-social. AQoL-8D: 8 dimensions - 3 physical; 5 psycho-social; S TTO = Self TTO; PWI = Personal Wellbeing Index; SWLS = Satisfaction with Life Survey; IHS = Integrated Household Survey

Table 7.3 Sensitivity: Summary of pairwise comparisons

Dimension where correlation with instrument exceeds ±0.1±0.2*

MAU with less sensitivity	**Instrument with greater sensitivity**						
	EQ-5D	**HUI3**	**SF-6D**	**15D**	**QWB**	**AQoL-4D**	**AQoL-8D**
EQ-5D		**Sense** Happiness Relations Worth SWLS IHS	General health **Role physical** **Vitality** **Social function** **Role emotional** Mental health Happiness Relations	General health Role physical Ind living **Sense** Vitality Social function Role emotional Mental health Happiness Coping Relations Worth		**Sense** Mental health Happiness Coping **Relations** Worth PWI SWLS IHS	General health **Sense** **Vitality** Social function Role emotional **Mental health** **Happiness** **Coping** **Relations** **Worth** **PWI** **SWLS** **IHS**
HUI3	Pain		Role physical Pain Vitality Social function **Role emotional** Mental health	General health Role physical Ind living Vitality Role emotional		Mental health **Relations**	**Vitality** Social function Role emotional **Mental health** **Happiness** **Coping** **Relations** **Worth** **PWI** **SWLS** **HIS** Self-TTO
SF-6D	Pain	Sense		Ind living **Sense**		**Sense** Relations IHS	**Sense** Mental health **Happiness** Coping **Relations** **Worth** PWI SWLS **IHS**
15D	Pain		Role physical Social function **Role emotional** Mental health			**Relations** IHS	Sense Vitality **Mental health** **Happiness** **Coping** **Relations** **Worth** **PWI** **SWLS** **HIS**

MAU with less sensitivity	Instrument with greater sensitivity						
	EQ-5D	**HUI3**	**SF-6D**	**15D**	**QWB**	**AQoL-4D**	**AQoL-8D**
QWB	Phys function **Pain** Social function Mental health Happiness Coping Worth PWI	Ind living Pain Sense Social function Mental health Happiness Coping Relations Worth PWI SWLS IHS	General health Phys function **Role physical** Pain **Vitality** **Social function** **Role emotional** **Mental health** Happiness Cope Relations Worth PWI SWLS IHS	General health Phys function Role physical Pain **Ind living** Sense Vitality **Social function** Role emotional **Mental health** Happiness **Cope** Relations **Worth** PWI SWLS IHS		Pain Sense Vitality Social function **Mental health** **Happiness** Coping **Relations** **Worth** **PWI** SWLS **IHS**	General health Sense **Vitality** **Social function** Role emotional **Mental health** **Happiness** **Cope** **Relations** **Worth** **PWI** **SWLS** **HIS** Self-TTO
AQoL-4D	Phys function Pain	Phys function	**Role physical** Pain Social function **Role emotional** Mental health	General health Phys function Role physical Pain Ind living			Vitality Mental health **Happiness** Cope **Worth** IHS
AQoL-8D	Phys function Pain	Phys function	**Phys function** **Role physical** Pain Social function Role emotional	**Phys function** Role physical Pain Ind living		Phys function Role physical Pain Ind living	

±0.1 = light text ±0.2 = **BOLD** text

8 Discussion and Conclusion

MAU instruments were scored for this paper using the algorithms summarised in Box 4. Prima facie the use of weights derived in one country in a second country may appear to invalidate the results. However this is not necessarily true and the issue of utility weights is complex. First there is very significant within country variation in preferences as found in the UK between social and demographic groups (Kind, Hardman et al. 1999). At best, national weights are themselves an average from heterogeneous groups.

The difference between national averages is presently of unknown importance. More significantly the evidence suggests the variance in scores is relatively insensitive to differences in weights. Using pilot data for this project Richardson and Khan (2012) found that 85 percent of the difference between instruments could be explained by unweighted instrument values, leaving little to be explained by differences in weights. As a further test of this, US and UK weights published by the EuroQOL group for the EQ-5D have been applied to the present data and the results reported in Figure 8.1. The R^2 of 0.9085 indicates that, overall, conclusions with respect to correlation and sensitivity could not change with the choice of weights. The significant difference in absolute score at the lower end of the scale suggests, prima facie, an error in the UK values. It appears very implausible that when UK citizens assign a score of 0.29, UK citizens would prefer to be dead.

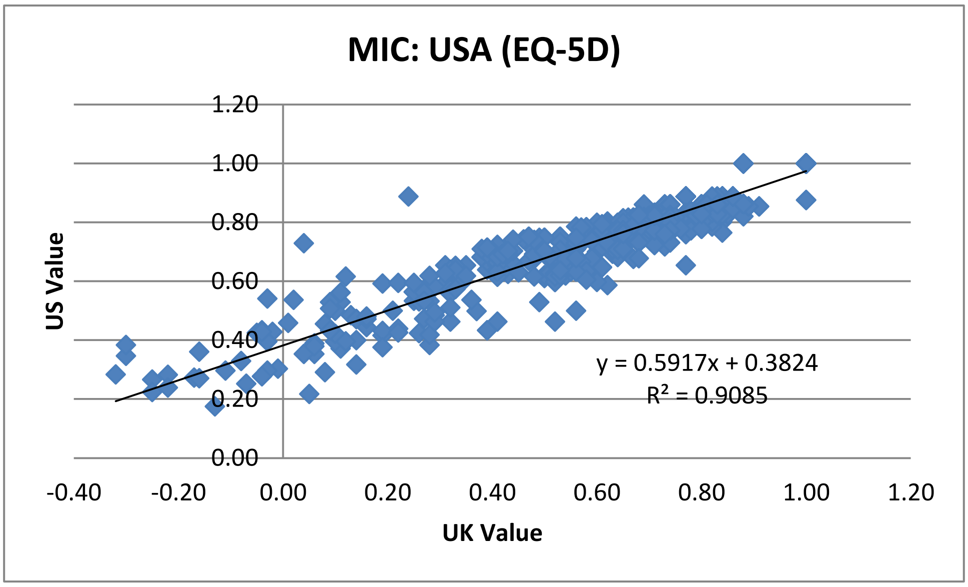
The two figures also indicate that the new five level EQ-5D does not overcome the problem of insensitivity in the region of good health (ceiling effects). The second highest possible UK and US utility scores are 0.906 and 0.888 respectively. This implies that moving 11 and 9 people respectively from the second highest health state to the highest would be equivalent to saving a life and returning a person to full health for the same period of time. Nevertheless some results might vary and the data available from this project could be reweighted with new scoring formula for difference countries.

The major conclusion to be drawn from this report is that, despite a similarity in the mean scores, the instruments are dissimilar with respect to virtually all other criteria used to compare them. Taking account of the fact that MAUI purport to measure the same quantity the correlation between instruments is low, the marginal relationships inconsistent and their relationship with health dimensions is variable. This suggests that, contrary to the impression generated by use of the generic term ‘utility’, the instruments are measuring different constructs. In effect each MAU instrument employs a different definition of ‘health’. The correlation which exists between instruments does not disconfirm this conclusion. Over a wide range of objects the height and weight of people correlate (the coefficient is about 0.81). But this does not demonstrate the existence of a common property (Chan 2003). A further important conclusion is that the evaluation of instruments is complex. Multiple criteria exist for their assessment many of which have not been discussed in this report.

Table 8.1 Summary of MAU order by criteria (USA)

Criteria	Instrument							Ratio highest/lowest
	EQ-5D	**HUI 3**	**SF-6D**	**15D**	**QWB**	**AQoL-4D**	**AQoL-8D**	
Distribution								
Mean value (Public)	0.90	0.89	0.80	0.94	0.76	0.81	0.87	1.24
Ceiling (% 1.00) Public	42.1	21.5	7.2	23.4	9.0	18.1	8.1	5.85
Floor (%<0.4) (All)	8.9	14.5	1.4	0.3	7.4	22.0	10.1	73.33
Correlation								
ICC (ave with other 7)	0.67	0.67	0.61	0.50	0.53	0.63	0.71	1.42
SWB (PWI)	0.45	0.50	0.50	0.49	0.36	0.54	0.64	1.78
SF-36	0.78	0.77	0.93	0.84	0.70	0.78	0.83	1.33
Self TTO	0.34	0.34	0.38	0.36	0.30	0.35	0.39	1.30
Discrepancies from b=1 in	36.2	52.8	53.7	61.7	41.7	50.5	34.5	1.79
Pairwise regression (ave %)								
Sensitivity								
b coefficient in mult reg on SF-36 dim (Table 6.4b)								
Pain	0.41	0.24	0.21	0.17	0.18	0.18	0.14	2.93
Gen Health	0.06	0.07	0.03	0.18	0.15	0.09	0.11	6.0
Physical function	0.28	0.31	0.14	0.27	0.19	0.19	0.10	3.1
Vitality	.002	.01	0.12	0.15	0.18	0.09	0.21	105.0-
Mental health	0.24	0.32	0.20	0.15	0.04	0.30	0.44	11.0
Rank order sensitivity using residuals								
Physical sum (SF-36)	3	4	2	1	5	6	7	
Physical sum (AQoL-8D)	4	2	6	1	7	3	5	
Mental sum (SF-36)	6	5	2	4	7	3	1	
Mental sum (AQoL-8D)	6	5	3	4	7	2	1	
Self TTO	6	5	2	3	7	4	1	
SWB (PWI)	6	3	4	5	7	2	1	

Figure 8.1 Comparison of EQ-5D with US and UK weights



Appendix 1 Frequency distribution of MAU instruments

Figure A.1.1 Frequency distribution of MAU instruments (Total n=1467)

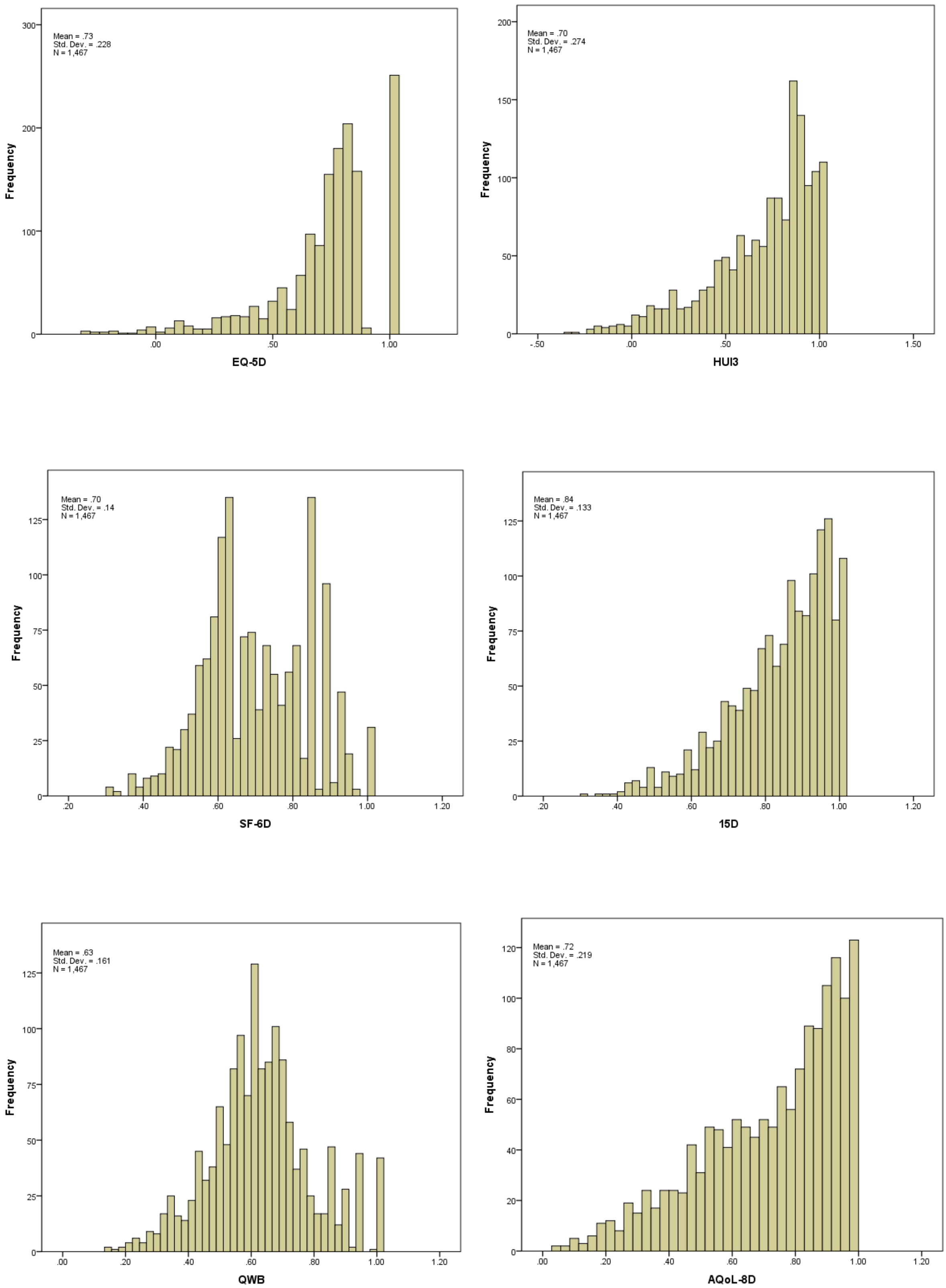
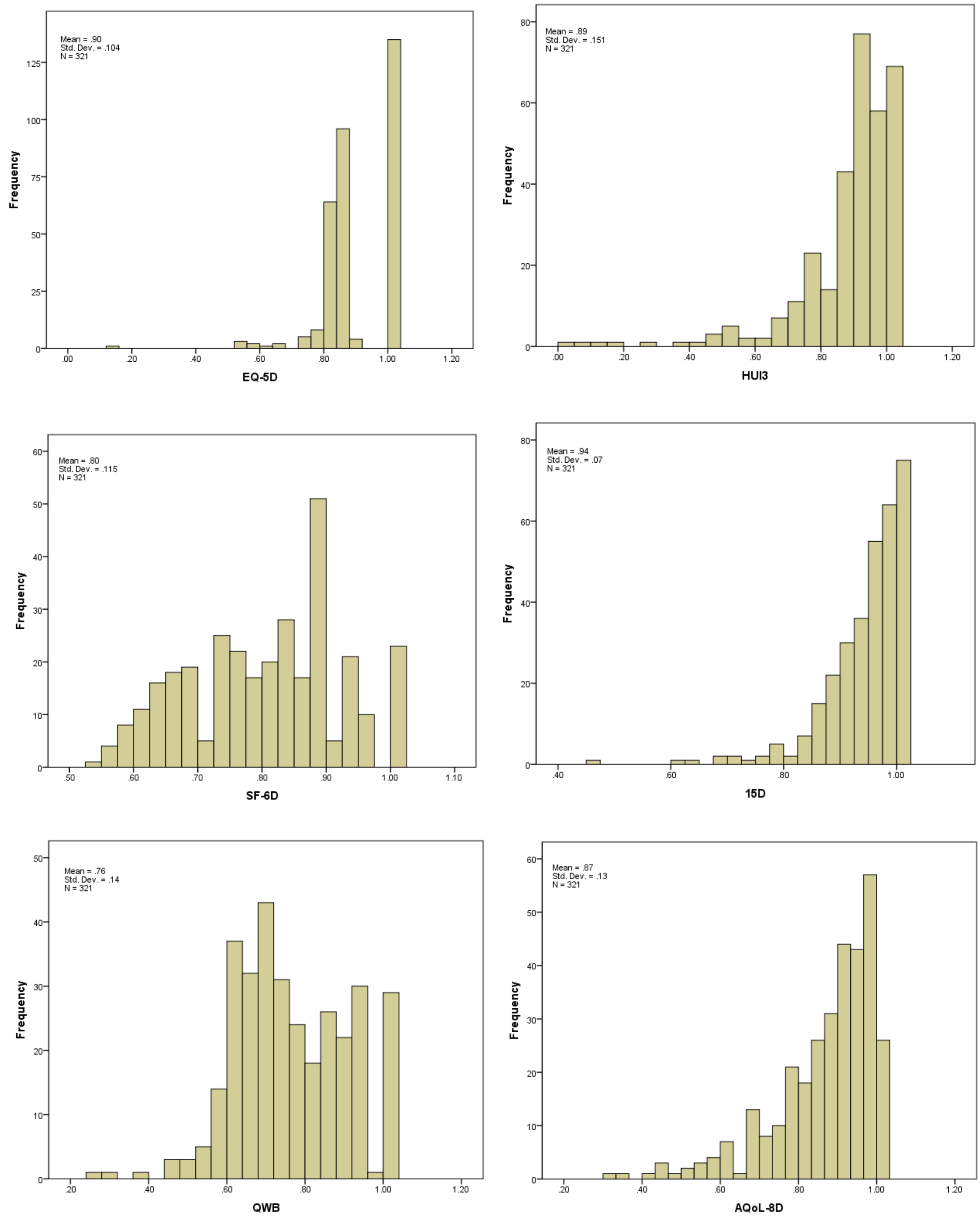


Figure A.1.2 Frequency distribution of MAU instruments (Public n=321)



Appendix 2 Frequency distribution of non-MAU instruments

Figure A.2.1 Frequency distribution of non-MAU instruments (Total n=1467)

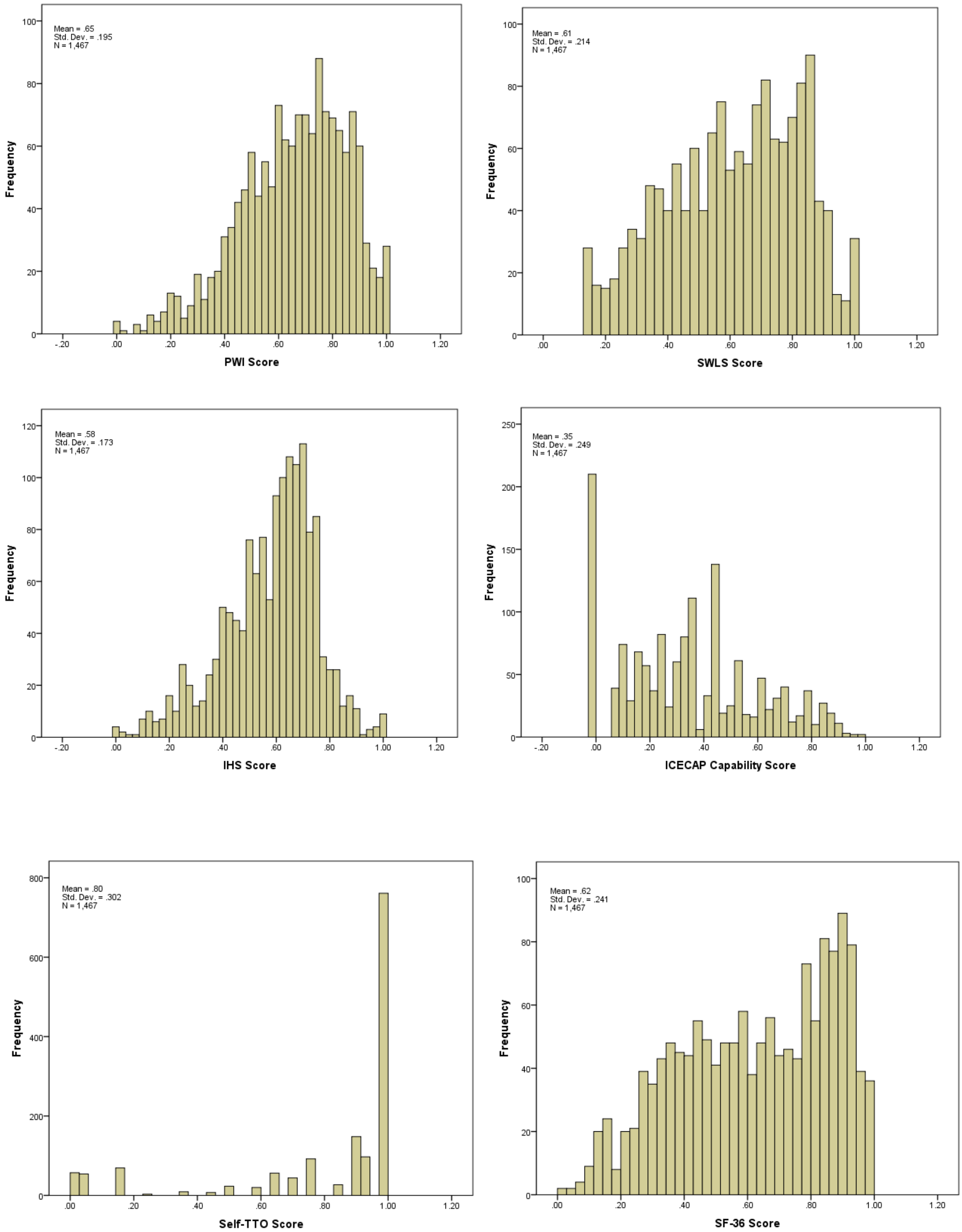
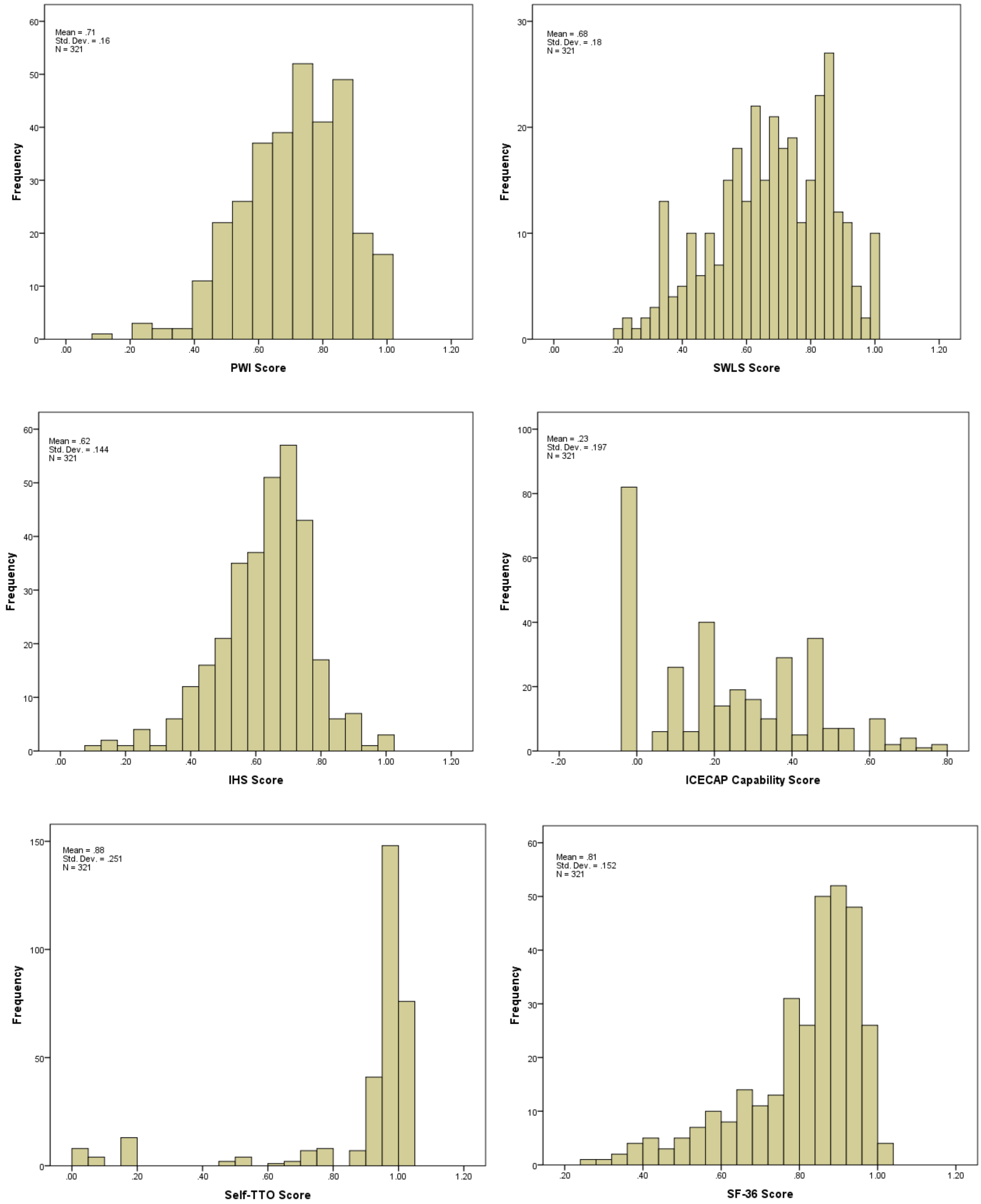
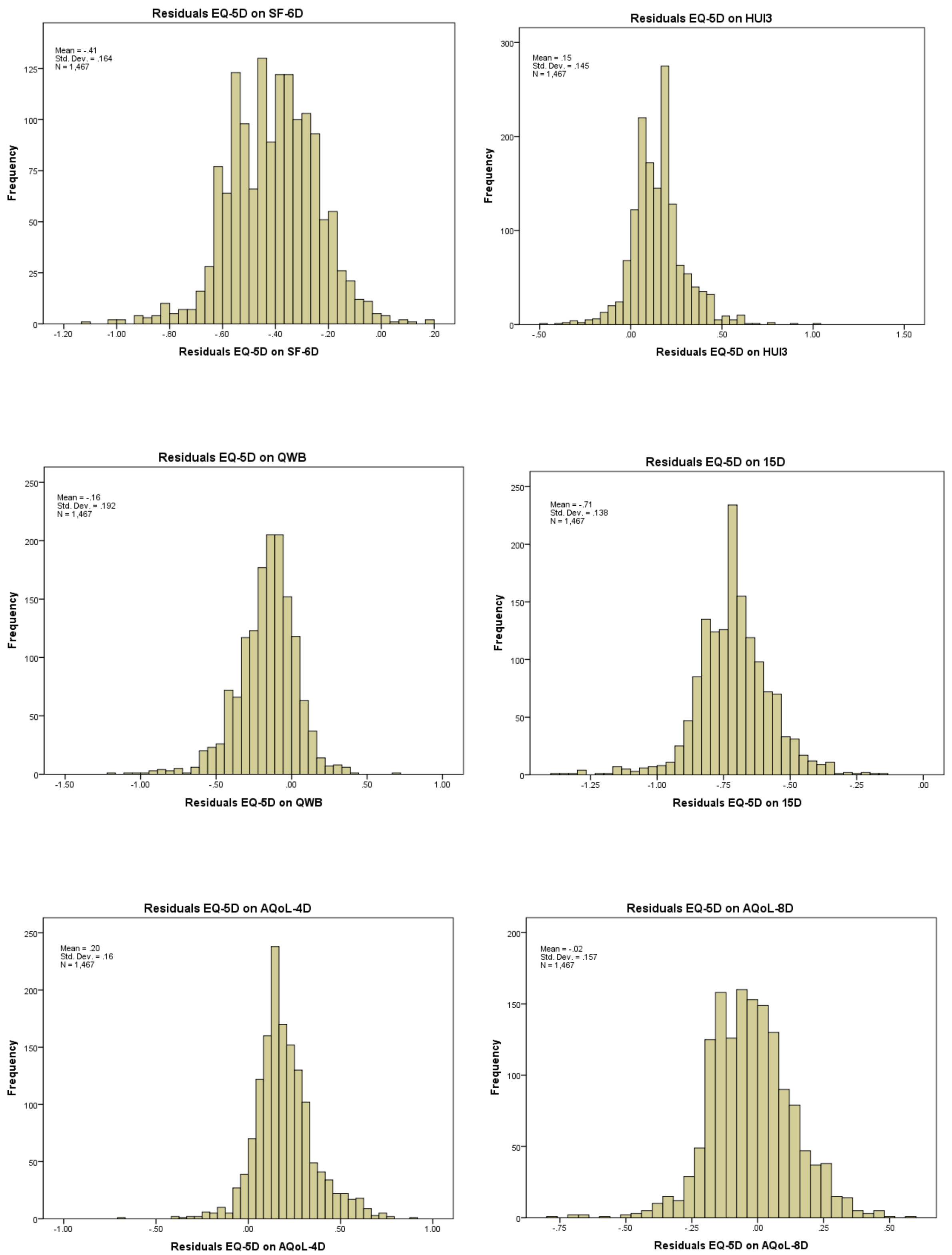


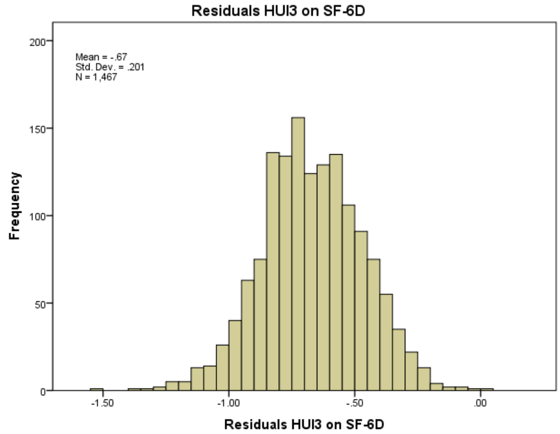
Figure A.2.2 Frequency distribution of non-MAU instruments (Public n=321)

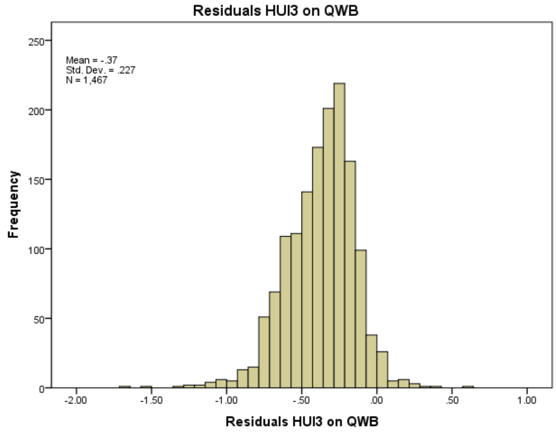


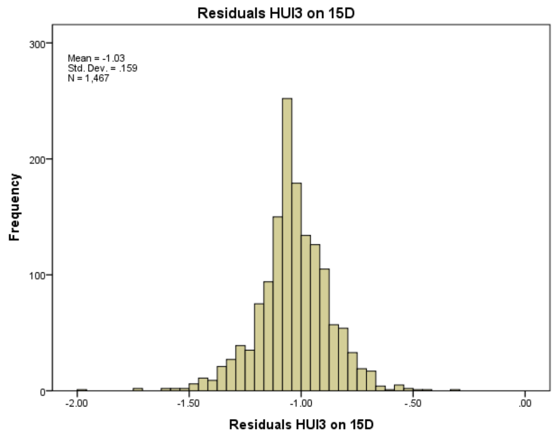
Appendix 3 Frequency distribution of residuals from pairwise regression of MAUI

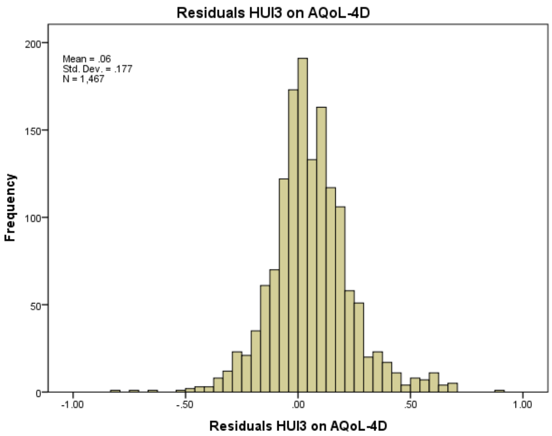
Figure A.3.1 Frequency distribution of residuals from pairwise regression of MAU instruments (Total n=1467)

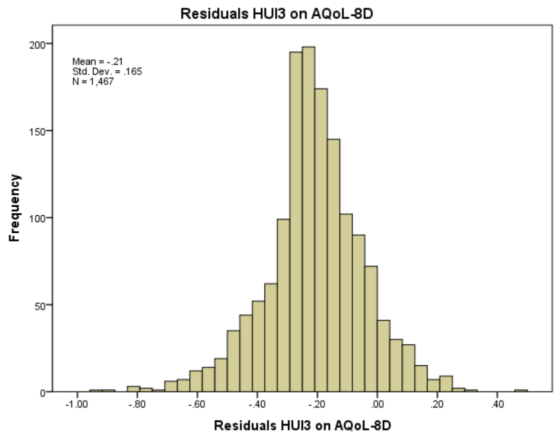


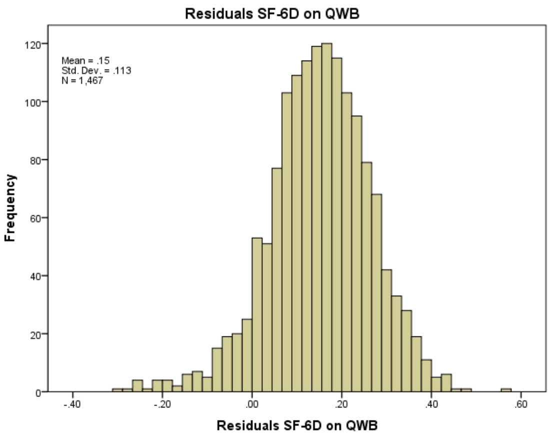

Residuals HUI3 on SF-6D
Mean = -.67
Std. Dev. = .201
N = 1,467
Frequency
Residuals HUI3 on SF-6D

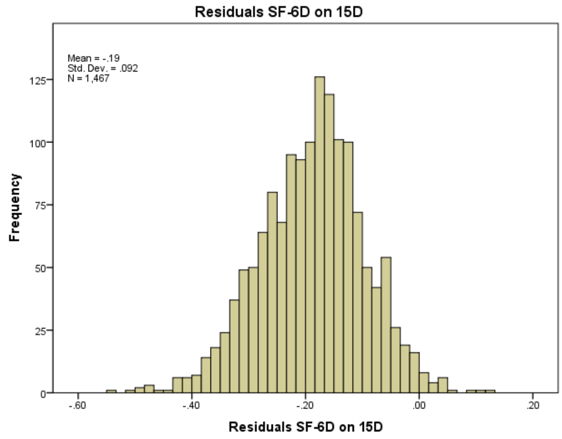

Residuals HUI3 on QWB
Mean = -.37
Std. Dev. = .227
N = 1,467
Frequency
Residuals HUI3 on QWB

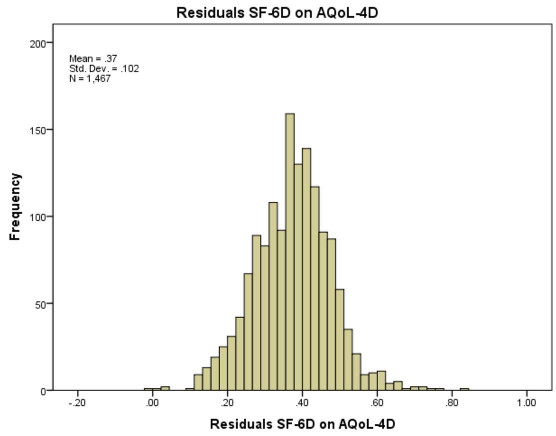

Residuals HUI3 on 15D
Mean = -1.03
Std. Dev. = .159
N = 1,467
Frequency
Residuals HUI3 on 15D

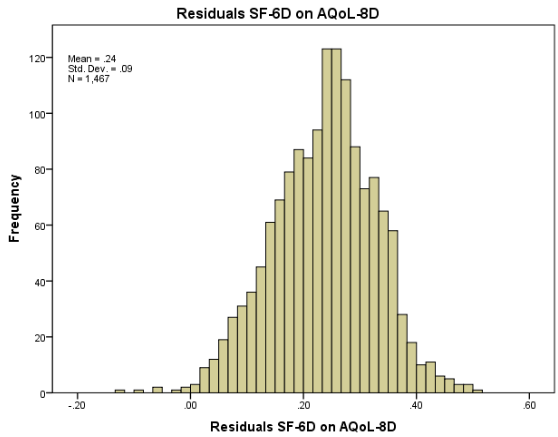

Residuals HUI3 on AQoL-4D
Mean = .06
Std. Dev. = .177
N = 1,467
Frequency
Residuals HUI3 on AQoL-4D

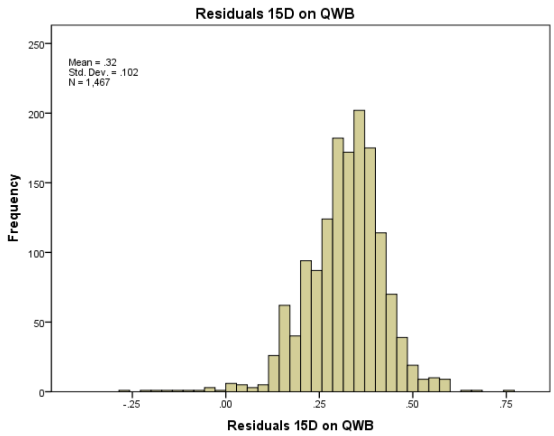

Residuals HUI3 on AQoL-8D
Mean = -.21
Std. Dev. = .165
N = 1,467
Frequency
Residuals HUI3 on AQoL-8D

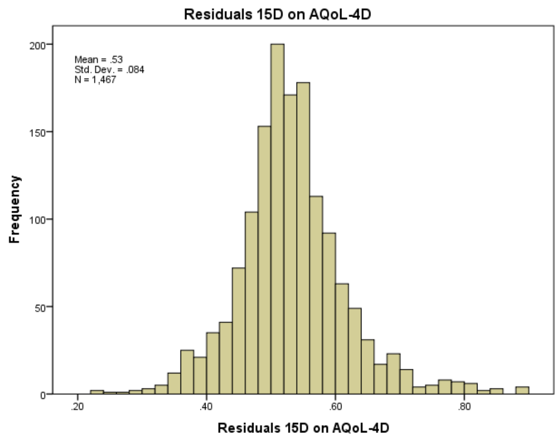

Residuals SF-6D on QWB
Mean = .15
Std. Dev. = .113
N = 1,467
Frequency
Residuals SF-6D on QWB

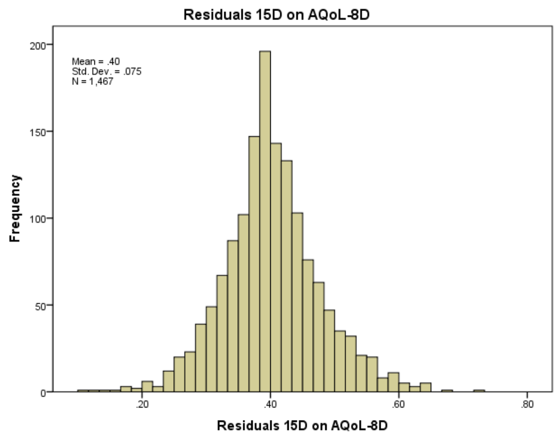

Residuals SF-6D on 15D
Mean = -.19
Std. Dev. = .092
N = 1,467
Frequency
Residuals SF-6D on 15D

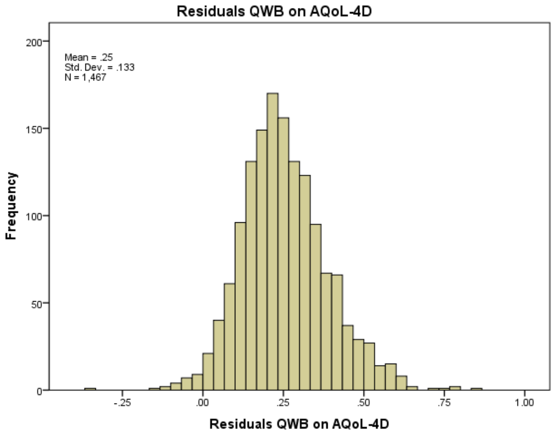

Residuals SF-6D on AQoL-4D
Mean = .37
Std. Dev. = .102
N = 1,467
Frequency
Residuals SF-6D on AQoL-4D

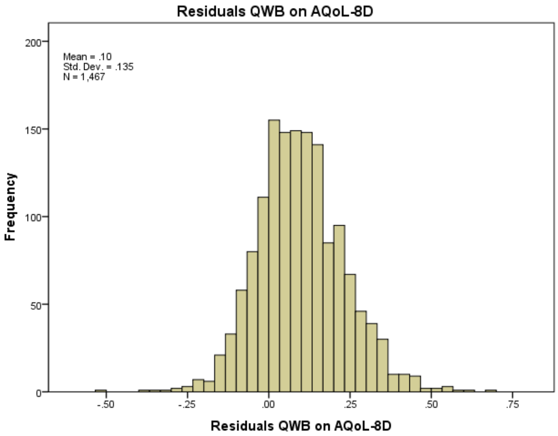

Residuals SF-6D on AQoL-8D
Mean = .24
Std. Dev. = .09
N = 1,467
Frequency
Residuals SF-6D on AQoL-8D

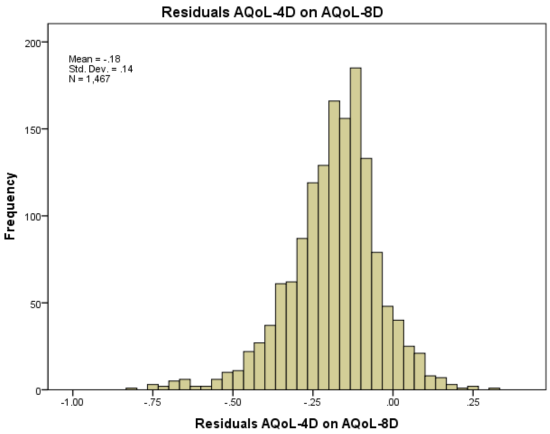

Residuals 15D on QWB
Mean = .32
Std. Dev. = .102
N = 1,467
Frequency
Residuals 15D on QWB


Residuals 15D on AQoL-4D
Mean = .53
Std. Dev. = .084
N = 1,467
Frequency
Residuals 15D on AQoL-4D


Residuals 15D on AQoL-8D
Mean = .40
Std. Dev. = .075
N = 1,467
Frequency
Residuals 15D on AQoL-8D


Residuals QWB on AQoL-4D
Mean = .25
Std. Dev. = .133
N = 1,467
Frequency
Residuals QWB on AQoL-4D


Residuals QWB on AQoL-8D
Mean = .10
Std. Dev. = .135
N = 1,467
Frequency
Residuals QWB on AQoL-8D


Residuals AQoL-4D on AQoL-8D
Mean = -.18
Std. Dev. = .14
N = 1,467
Frequency
Residuals AQoL-4D on AQoL-8D

References

Chan, Y. H. (2003). "Biostatistics 104: Correlational Analysis." Singapore Medical Journal **44**(12): 614-619.

Cronbach, L. (1951). "Coefficient alpha and the internal structure of tests." Psychometrika **16**(3): 297-334.

Hawthorne, G., S. Korn, et al. (2012). "Population norms for the AQoL derived from the 2007 Australian National Survey of Mental Health and Wellbeing." Australian and New Zealand Journal of Public Health (submitted).

Kind, P., G. Hardman, et al. (1999). US Population Norms for EQ-5D, Discussion Paper 172. York, Centre for Health Economics, University of York.

Pallant, J. (2010). SPSS Survival Manual: A step by step guide to data analysis using SPSS. Crows Nest, Allen and Unwin.

Richardson, J., G. Elsworth, et al. (2011). Increasing the Sensitivity of the AQoL Inventory for Evaluation of Interventions Affecting Mental Health, Research Paper 61. Melbourne, Centre for Health Economics, Monash University.

Richardson, J., A. Iezzi, et al. (forthcoming). The Self TTO: Its properties, validity, reliability and relationship with MAU scores, Research Paper. Melbourne, Centre for Health Economics, Monash University.

Richardson, J., A. Iezzi, et al. (2012). Cross-national comparison of twelve quality of life instruments: MIC Paper 1 Background, questions, instruments, Research Paper 76. Melbourne, Centre for Health Economics, Monash University.

Richardson, J., M. A. Khan, et al. (2012). Population norms and Australian profile using the Assessment of Quality of Life (AQoL) 8D utility instrument, Research Paper 72. Melbourne, Centre for Health Economics, Monash University.

Richardson, J., A. Maxwell, et al. (2012). The Myth of Maximisation: Quantifying trade-offs between social objectives using the Conflict Scale, Research Paper 74. Melbourne, Centre for Health Economics, Monash University.