ISPOR 7th Asia-Pacific Conference, Suntec Convention & Exhibition Center, Singapore
Patient-Reported Outcomes & Patient Preference Research
Sunday, 4 September 2016, 5:00 PM - 6:00 PM @ Room 326

W4: COMBINING TWO TYPES OF VALUATION DATA TO ESTIMATE HEALTH STATE UTILITIES: THE HYBRID REGRESSION MODEL

Discussion Leaders:
Nan Luo
National University of Singapore, Singapore
Kim Rand-Hendriksen
University of Oslo, Norway
Mark Oppe
Juan M. Ramos-Goñi
EuroQol Research Foundation, Rotterdam, The Netherlands

www.euroqol.org

Conflict of interest & disclaimer

- The discussion leaders are members of the EuroQol Group, a not-for-profit international research organization
- The views of the discussion leaders expressed in the workshop do not necessarily reflect the views of the EuroQol Group

www.euroqol.org

EQ-5D

Health-State Valuation & Data Collection

Nan Luo
(nan_luo@nuhs.edu.sg)
National University of Singapore, Singapore

EQ-5D

Health-state utility

- Value of a health state to a stakeholder such as the general public
- Measured on a cardinal scale anchored by 0 (death) and 1 (full health)
- Quality-of-life weights for calculating quality-adjusted life years (QALYs) in cost-utility analysis
- Determination of utility values
  - Direct method: eliciting the utility value of a health state from a respondent using a valuation technique such as standard gamble
  - Indirect method: preference-based instrument

EQ-5D

The EQ-5D-5L instrument

- The questionnaire
  - A new version of the EQ-5D instrument
  - Consist of a descriptive system and a visual analog scale for measuring health related quality of life
  - Defines 3,125 unique health states with 5 dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) and 5 levels of problems for each dimension (no/slight/moderate/severe/extreme)
- The value set
  - Consist of utility values for all 3,125 EQ-5D-5L health states
  - Values anchored by 0 (dead) and 1 (full health)
  - Reflecting the average values of the health states to the general public
  - Country-specific value set

www.euroqol.org

EQ-5D

Deriving utility values using EQ-5D-5L

“12345”

Coding:
No = 1
Slight = 2
Moderate = 3
Severe = 4
Extreme = 5

Value Set

State	Value
11111	1.00
….	
12344	0.13
12345	0.10
12351	0.05
12352	0.07
12353	0.03
…	
55555	-0.60

www.euroqol.org

Estimating an EQ-5D utility value set

- A 2-step procedure is used:
  - First, the values for a subset of health states are derived from stakeholders using the direct valuation approach
  - Second, a function is estimated by modeling the directly measured values to predict values for all health states

EQ-5D-5L Valuation Study

- To establish EQ-5D-5L value sets using a standardized data collection protocol
- The data collection protocol has been used in many countries including 8 Asian countries
- Main features:
  - Face-to-face personal interview using a computer program called EuroQol Valuation Technology (EQ-VT)
  - Interviewer training and monitoring
  - Use of two valuation methods: time trade-off (TTO) & discrete choice experiment (DCE)

The EQ-VT interview

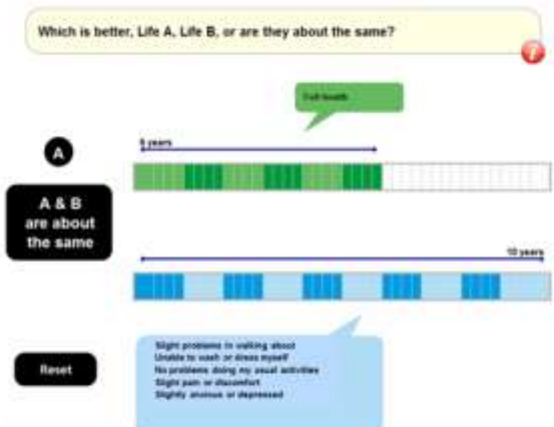
1. Introduction
    a. Self-reported health on the EQ-5D-5L descriptive system
    b. Self-reported health on the EQ-VAS
    c. Background questions
2. Composite Time Trade-Off (TTO)
    a. Instructions and example
    b. Three practices
    c. Valuation of 10 EQ-5D-5L states
    d. Feedback module
    e. Debriefing questions
3. Discrete Choice Experiment
    a. Instructions
    b. Valuation of 7 pairs of EQ-5D-5L states
    c. Debriefing questions
4. Demographic/additional questions

Composite TTO

- Conventional or lead-time TTO procedure is used contingent on the value of a health state to the respondent
- If the health state is considered as better than dead, the procedure used in the EQ-5D-3L valuation studies is used
- If the health state is considered as worse than dead, a new, lead-time procedure is used

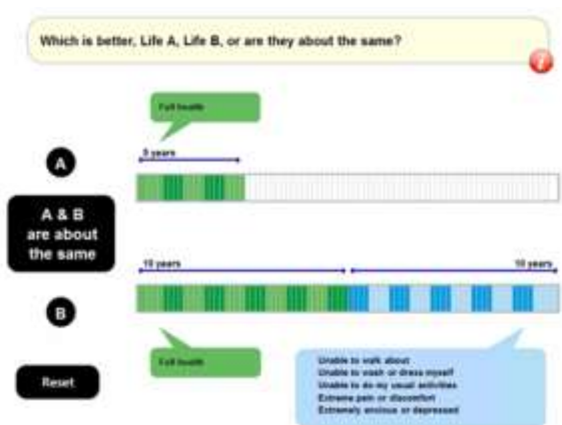
EQ-5D

Conventional TTO

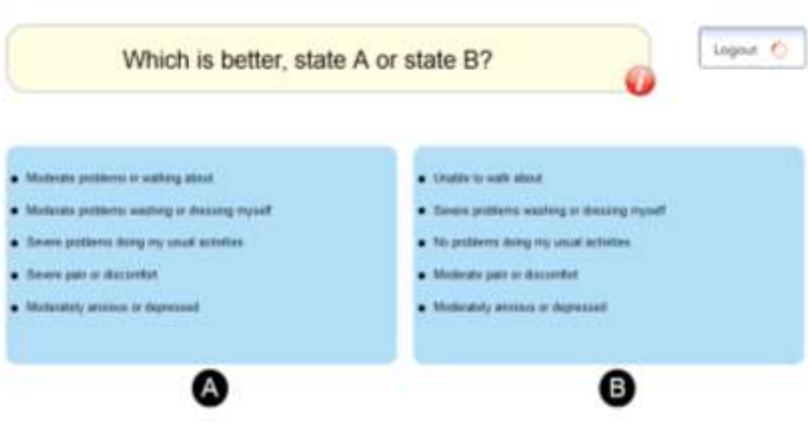


EQ-5D

Lead-Time TTO



DCE task



Health states for direct valuation

TTO

- 86 states were divided into 10 blocks, with each block containing 10 states
  - 1 mildest state
  - 55555 (all-worst)
  - 8 other states
- Each respondent was randomized to value one block

DCE

- 196 pairs were randomly divided into 28 blocks, with each block containing 7 pairs of states
- Each respondent was randomized to value one block

Overview of modelling analysis

Data	Model function form	Model parameter estimator	Model prediction evaluation	Model validation
TTO	Linear models Non-linear models	OLS; GLS; etc.	MAE	Out-of-sample health states Out-of-sample respondents
DCE		Conditional logit; Conditional probit	MAE	
TTO + DCE		“Hybrid” estimator	???	

OLS – ordinary least square; GLS – generalized least square; MAE – mean absolute error

The 20-parameter main-effects model

$$y = \alpha +$$

$$\beta_{MO2}x_{MO2} + \beta_{SC2}x_{SC2} + \beta_{UA2}x_{UA2} + \beta_{PD2}x_{PD2} + \beta_{AD2}x_{AD2} +$$

$$\beta_{MO3}x_{MO3} + \beta_{SC3}x_{SC3} + \beta_{UA3}x_{UA3} + \beta_{PD3}x_{PD3} + \beta_{AD3}x_{AD3} +$$

$$\beta_{MO4}x_{MO4} + \beta_{SC4}x_{SC4} + \beta_{UA4}x_{UA4} + \beta_{PD4}x_{PD4} + \beta_{AD4}x_{AD4} +$$

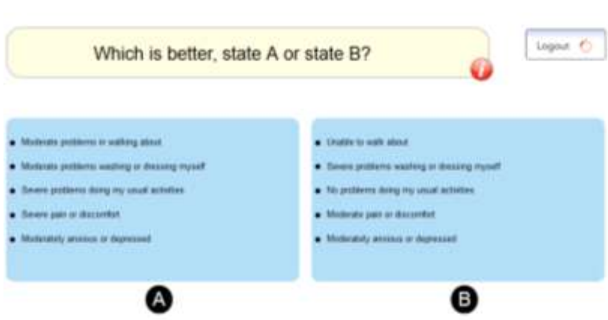
$$\beta_{MO5}x_{MO5} + \beta_{SC5}x_{SC5} + \beta_{UA5}x_{UA5} + \beta_{PD5}x_{PD5} + \beta_{AD5}x_{AD5}$$

DCE models and non-linear (constrained) models

Kim Rand-Hendriksen, PhD
University of Oslo

Singapore, September 2016

DCE task - reminder



www.euroqol.org

DCE models (rationale)

- We assume that respondents will select the alternative that maximizes their utility, with some degree of random error.
- If u(A) is the utility of alternative A, and u(B) is the utility of state B, the probability of selecting A is
- Since both u(A) and u(B) are subject to random variance, u(X) = v(X) + ε_X

$$P(Y_i = A) = P(U_{iA} > U_{iB}) = P(V_{iA} + \varepsilon_{iA} > V_{iB} + \varepsilon_{iB}) = P(V_{iA} - V_{iB} > \varepsilon_{iB} - \varepsilon_{iA})$$

$$P(Y_i = A) = \frac{e^{\lambda V_{iA}}}{e^{\lambda V_{iA}} + e^{\lambda V_{iB}}}$$

$$V_i = \sum_{j=1}^{J} X_{i,j} \cdot \beta_j$$

- Conditional logistic model
- Specific conditional logit/probit model for pair comparisons
- GLS random intercept logit/probit model

DCE Conditional logistic model

STATE_ID	ORDER_ID	MO	SC	UA	PD	AD	ORIGINAL_RESPONSE	Value	METHOD	TASK_TIME_ME	Profile	Severity
18	4	1	4	5	3	3	B	0	DCE_A	19.6	14533	16
18	4	2	1	5	4	2	A	1	DCE_B	19.6	21542	14
42	6	2	3	1	3	4	A	1	DCE_A	18.2	23134	13
42	6	1	4	3	1	4	B	0	DCE_B	18.2	14314	13
75	5	5	3	4	3	1	B	0	DCE_A	9.4	53431	16
75	5	5	2	2	5	5	A	1	DCE_B	9.4	52255	19
76	2	5	1	5	2	2	B	0	DCE_A	21.4	51522	15
76	2	4	5	2	4	4	A	1	DCE_B	21.4	45244	19
127	7	1	4	2	2	4	A	1	DCE_A	33.2	14224	13
127	7	3	2	3	2	2	B	0	DCE_B	33.2	32322	12
180	3	4	4	1	4	5	A	1	DCE_A	7.2	44145	18
180	3	4	5	4	3	2	B	0	DCE_B	7.2	45432	18
196	1	1	1	2	2	1	A	1	DCE_A	52.1	11221	7
196	1	2	2	1	2	2	B	0	DCE_B	52.1	22122	9

Stata code:

```
generate dce_group = concat(respondent_id state id)
clogit choice dummy_variables, group(dce_group)
```

R code:

```
df$dce_group <- interaction(df$respondent_id, df$state_id)
Clogit(Value ~ [dummy_variables] + strata(dce_group))
```

www.euroqol.org

EQ-5D

```
. clogit value mo2-ad5 if method!="TTO" & state_id <197, group( dce_group)

Iteration 0:   log likelihood = -4009.8239
Iteration 1:   log likelihood = -3797.2072
Iteration 2:   log likelihood = -3794.7782
Iteration 3:   log likelihood = -3794.7752
Iteration 4:   log likelihood = -3794.7752

Conditional (fixed-effects) logistic regression   Number of obs   =     14474
                                                  LR chi2(20)     =   2415.34
                                                  Prob > chi2     =    0.0000
Log likelihood = -3794.7752                       Pseudo R2       =    0.2414
```

value	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
mo2	.4127326	.0578585	7.13	0.000	.298332	.5261332
mo3	.1202229	.0646377	1.86	0.063	-.0064847	.2469305
mo4	.5920213	.0623833	9.49	0.000	.4697522	.7142904
mo5	.5503489	.0666046	8.26	0.000	.4198063	.6808916
sc2	.150481	.0613851	2.45	0.014	.0301689	.2707936
sc3	.1820482	.0653989	2.80	0.005	.0544587	.3096396
sc4	.3663121	.068909	5.31	0.000	.2312216	.5014026
sc5	.3452481	.0648272	5.33	0.000	.2181893	.4723072
ua2	.2342737	.0589688	3.97	0.000	.118697	.3498504
ua3	.0605232	.063089	0.96	0.337	-.0631384	.1841832
ua4	.4118989	.0634811	6.49	0.000	.287537	.5362628
ua5	.3216284	.0668057	4.83	0.000	.1910837	.4521732
pd2	.2816590	.0608909	4.64	0.000	.1627072	.4006111
pd3	.1241066	.0646124	1.92	0.055	-.0025314	.2507446
pd4	.9133304	.0657241	13.90	0.000	.7848135	1.042147
pd5	.3751377	.0689618	5.44	0.000	.2399781	.5103002
ad2	.2247589	.0626363	3.59	0.000	.1019628	.3475237
ad3	.2671108	.0656561	4.07	0.000	.1384272	.3957945
ad4	.5540466	.0703022	7.90	0.000	.4166489	.6914443
ad5	.4625092	.0670387	6.91	0.000	.3315148	.5943017

EQ-5D

Specific conditional logit/probit model for DCE pair comparisons

STATE_ID	ORDER_ID	MO A	SC A	UA A	PD A	AD A	MO B	SC B	UA B	PD B	AD B	ORIGINAL_RESPONSE	Value	METHOD	TASK_TIME_sec	Profile A	Severity A	Profile B	Severity B
18	4	1	4	5	3	3	2	1	5	4	2	B	0	DCE_A	19.6	14533	16	21542	14
42	6	2	3	1	3	4	1	4	3	1	4	A	1	DCE_A	18.2	23134	13	14314	13
75	5	5	3	4	3	1	5	2	2	5	5	B	0	DCE_A	9.4	53431	16	52255	19
76	2	5	1	5	2	2	4	5	2	4	4	B	0	DCE_A	21.4	51522	15	45244	19
127	7	1	4	2	2	4	3	2	3	2	2	A	1	DCE_A	33.2	14224	13	32322	12
180	3	4	4	1	4	5	4	5	4	3	2	A	1	DCE_A	7.2	44145	18	45432	18
196	1	1	1	2	2	1	2	2	1	2	2	A	1	DCE_A	52.1	11221	7	22122	9

- **Generate "A-B"**
- **logit** choice dummy_variables, nocons
- **probit** choice dummy_variables, nocons
- **hetprobit** choice dummy_variables, nocons het(dummy_variables)

- Random intercept/random slope logistic models can be fitted in STATA:
  - **xtlogit** choice dummy_variables
  - **xtprobit** choice dummy_variables
  - **melogit** choice dummy_variables
  - **meprobit** choice dummy_variables

Stata output

```
. logit value _mo2-_ad5 if method=="DCE_A" & state_id <137, nocons

Iteration 0:   log likelihood = -5002.4432
Iteration 1:   log likelihood = -3800.0306
Iteration 2:   log likelihood = -3794.7784
Iteration 3:   log likelihood = -3794.7752
Iteration 4:   log likelihood = -3794.7752

Logistic regression                             Number of obs   =       7217
                                                Wald chi2(20)   =    1525.43
Log likelihood = -3794.7752                     Prob > chi2     =     0.0000
```

value	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_mo2	.4127324	.0578589	7.13	0.000	.299332	.5261333
_mo3	.1202029	.0646377	1.86	0.063	-.0064847	.2468905
_mo4	.5920013	.0623833	9.49	0.000	.4697521	.7142904
_mo5	.5503469	.0666344	8.26	0.000	.4198063	.6808916
_sc2	.150481	.0613551	2.45	0.014	.0302185	.2707336
_sc3	.1825482	.0650889	2.80	0.005	.0549567	.3096396
_sc4	.3663121	.068935	5.31	0.000	.2312216	.5014026
_sc5	.3452483	.0648272	5.33	0.000	.2181893	.4723072
_ua2	.2342737	.0589688	3.97	0.000	.118697	.3498504
_ua3	.0605232	.063093	0.96	0.337	-.0631349	.1841832
_ua4	.4118989	.0634511	6.49	0.000	.287537	.5362608
_ua5	.3216284	.0666057	4.83	0.000	.1910837	.4521732
_pd2	.2816592	.0606909	4.64	0.000	.1627073	.4006111
_pd3	.1241066	.0646134	1.92	0.055	-.0025314	.2507446
_pd4	.9133204	.0657241	13.90	0.000	.7845135	1.042127
_pd5	.3751377	.0689438	5.44	0.000	.2399751	.5103003
_ad2	.2247588	.0626363	3.59	0.000	.1019938	.3475237
_ad3	.2671108	.0656561	4.07	0.000	.1384272	.3957943
_ad4	.5540466	.0701022	7.90	0.000	.4166488	.6914445
_ad5	.4628082	.0670387	6.91	0.000	.3313148	.5943017

DCE constant/intercept terms

"Original"

	Estimate	Std. Error
(Intercept)	0.21164586	0.0266645
incr_mo2_diff	0.44230775	0.0526517
incr_mo3_diff	-0.00271160	0.0602873
incr_mo4_diff	0.77762394	0.0592446
incr_mo5_diff	1.16562366	0.0653388
incr_sc2_diff	0.35921265	0.0576461
incr_sc3_diff	-0.00847601	0.0625683
incr_sc4_diff	0.67629896	0.0661428
incr_sc5_diff	0.39924359	0.0615881
incr_ua2_diff	0.27154495	0.0549798
incr_ua3_diff	0.04449129	0.0615065
incr_ua4_diff	0.53745959	0.0619102
incr_ua5_diff	0.57164557	0.0636063
incr_pd2_diff	0.25449076	0.0580883
incr_pd3_diff	0.03962980	0.0633029
incr_pd4_diff	0.80788425	0.0642941
incr_pd5_diff	0.30395110	0.0638620
incr_ad2_diff	0.27174584	0.0603230
incr_ad3_diff	0.19483877	0.0615435
incr_ad4_diff	0.88367616	0.0655376
incr_ad5_diff	0.52791784	0.0654440

Reversed

	Estimate	Std. Error
(Intercept)	-0.21164586	0.0266645
incr_mo2_diff	0.44230775	0.0526517
incr_mo3_diff	-0.00271160	0.0602873
incr_mo4_diff	0.77762394	0.0592446
incr_mo5_diff	1.16562366	0.0653388
incr_sc2_diff	0.35921265	0.0576461
incr_sc3_diff	-0.00847601	0.0625683
incr_sc4_diff	0.67629896	0.0661428
incr_sc5_diff	0.39924359	0.0615881
incr_ua2_diff	0.27154495	0.0549798
incr_ua3_diff	0.04449129	0.0615065
incr_ua4_diff	0.53745959	0.0619102
incr_ua5_diff	0.57164557	0.0636063
incr_pd2_diff	0.25449076	0.0580883
incr_pd3_diff	0.03962980	0.0633029
incr_pd4_diff	0.80788425	0.0642941
incr_pd5_diff	0.30395110	0.0638620
incr_ad2_diff	0.27174584	0.0603230
incr_ad3_diff	0.19483877	0.0615435
incr_ad4_diff	0.88367616	0.0655376
incr_ad5_diff	0.52791784	0.0654440

Non-linear (constrained) models

- While the wording is somewhat different, the levels for the 5 EQ-5D dimensions correspond roughly to:
  1. No problems
  2. Slight problems
  3. Moderate problems
  4. Severe problems
  5. Extreme problems ("unable" for mobility, self-care, and usual activities)
- With the standard 20-parameter model, each decrement is considered in isolation.
- An alternative approach is to assume that the levels describe similar relative decrements across dimensions, and that the dimensions differ primarily in scale.
- The simplest such model requires only 8 parameter: one for each dimension, and one for each of levels 2, 3, and 4

The 20-parameter model

$$tto = \alpha + \Sigma_l \Sigma_d \beta_{dl} x_{dl} + e$$

$$= \alpha + \beta_{MO2}x_{MO2} + \beta_{SC2}x_{SC2} + \beta_{UA2}x_{UA2} + \beta_{PD2}x_{PD2} + \beta_{AD2}x_{AD2} +$$

$$\beta_{MO3}x_{MO3} + \beta_{SC3}x_{SC3} + \beta_{UA3}x_{UA3} + \beta_{PD3}x_{PD3} + \beta_{AD3}x_{AD3} +$$

$$\beta_{MO4}x_{MO4} + \beta_{SC4}x_{SC4} + \beta_{UA4}x_{UA4} + \beta_{PD4}x_{PD4} + \beta_{AD4}x_{AD4} +$$

$$\beta_{MO5}x_{MO5} + \beta_{SC5}x_{SC5} + \beta_{UA5}x_{UA5} + \beta_{PD5}x_{PD5} + \beta_{AD5}x_{AD5} + e$$

EQ-5D

8-parameter model

- A single parameter per dimensjon (MO, SC, UA, PD, AD), and one for each of levels 2, 3, and 4 (L2, L3, L4):

$$tto = \alpha + \sum_l (\sum_d \beta_d \, x_{dl}) \, L_l + \, e$$

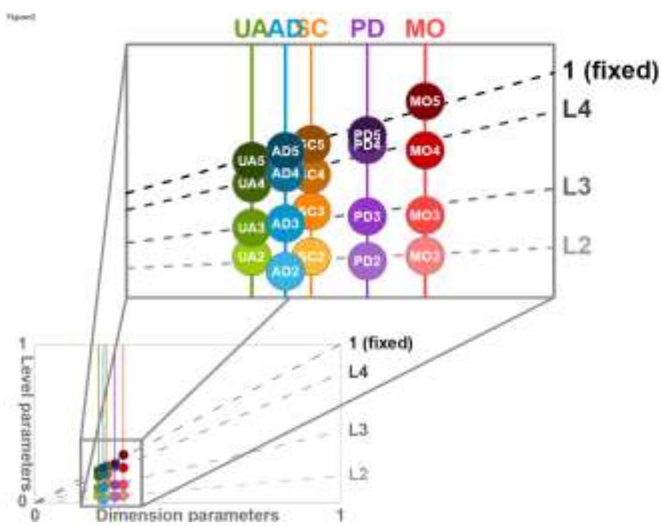
$$= \alpha + (\beta_{MO}x_{MO2} + \beta_{SC}x_{SC2} + \beta_{UA}x_{UA2} + \beta_{PD}x_{PD2} + \beta_{AD}x_{AD2})L_2 +$$

$$(\beta_{MO}x_{MO3} + \beta_{SC}x_{SC3} + \beta_{UA}x_{UA3} + \beta_{PD}x_{PD3} + \beta_{AD}x_{AD3})L_3 +$$

$$(\beta_{MO}x_{MO4} + \beta_{SC}x_{SC4} + \beta_{UA}x_{UA4} + \beta_{PD}x_{PD4} + \beta_{AD}x_{AD4})L_4 +$$

$$\beta_{MO}x_{MO5} + \beta_{SC}x_{SC5} + \beta_{UA}x_{UA5} + \beta_{PD}x_{PD5} + \beta_{AD}x_{AD5} + e$$

EQ-5D



Extensions

- The wording for level 5 on mobility, self-care and usual activities uses "unable to", while pain/discomfort and anxiety/depression uses "extreme".
- To handle this difference, a new parameter can be added for "unable to", for a total of 9 parameters
- Alternatively, the level structure can be estimated separately for the first three and the last two dimensions, for a total of 11 parameters
- These models are non-linear, and require specialized fitting functions. In STATA:

```
nl (value = {intercept} + /*
*/ (nl_mo2*{M=0} + nl_sc2*{S=0} + nl_ua2*{U=0} + nl_pd2*{P=0} + nl_ad2*{A=0}) * {L2=0.25} /*
*/ + (nl_mo3 * {M} + nl_sc3 * {S} + nl_ua3 * {U} + nl_pd3 * {P} + nl_ad3 * {A}) * {L3=0.5} /*
*/ + (nl_mo4 * {M} + nl_sc4 * {S} + nl_ua4 * {U} + nl_pd4 * {P} + nl_ad4 * {A}) * {L4=1} /*
*/ + (nl_mo5 * {M} + nl_sc5 * {S} + nl_ua5 * {U} + nl_pd5 * {P} + nl_ad5 * {A}))
*/ if ………………………………..,iter(1000)
```

www.euroqol.org

Non-linear models cont'd

- The code on the last slide requires the dummies to be coded such that mobility level 4 yields
  mo2 = 0, mo3 = 0, mo4 = 1, mo5 = 0
- For the DCE models, the coding is generally:
  mo2 = 1, mo3 = 1, mo4 = 1, mo5 = 0
- With this coding, the non-linear function previously presented will tend to become unstable and unreliable. It can be altered to accommodate this coding as follows:

```
nl (value = {intercept} + /*
*/ {MO = 0} * (_mo2 * {L2 = 0.25} + _mo3 * {L3 = 0.25} + _mo4 * {L4 = 0.25} + _mo5 * (1-L2-L3-L4)) + /*
*/ {SC = 0} * (_sc2 * L2 + _sc3 * L3 + _sc4 * L4 + _sc5 * (1-L2-L3-L4)) + /*
*/ {UA = 0} * (_ua2 * L2 + _ua3 * L3 + _ua4 * L4 + _ua5 * (1-L2-L3-L4)) + /*
*/ {PD = 0} * (_pd2 * L2 + _pd3 * L3 + _pd4 * L4 + _pd5 * (1-L2-L3-L4)) + /*
*/ {AD = 0} * (_ad2 * L2 + _ad3 * L3 + _ad4 * L4 + _ad5 * (1-L2-L3-L4)) /*
*/ if ………………………………..,iter(1000)
```

www.euroqol.org

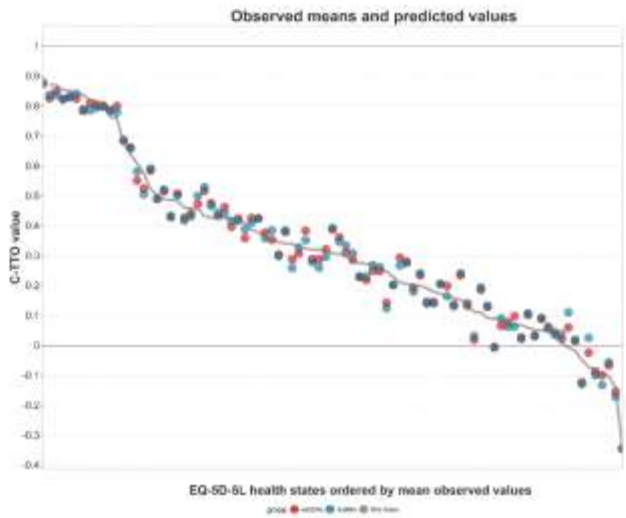


Table 2: Cross-validation results and model fit

			[illegible]	[illegible]	[illegible]	[illegible]
Mean absolute error	China	All data	[illegible]	[illegible]	[illegible]	[illegible]
		Method 1	0.0426	**0.0398**	**0.0399**	0.0434
		Method 2	0.0420	**0.0408**	0.0409	0.0416
	Spain	All data	[illegible]	[illegible]	[illegible]	[illegible]
		Method 1	0.0589	**0.0569**	0.0574	0.0576
		Method 2	0.0583	0.0574	[illegible]	[illegible]
	Singapore	All data	[illegible]	[illegible]	[illegible]	[illegible]
		Method 1	0.0944	**0.0869**	0.0871	0.0866
		Method 2	0.0888	**0.0851**	0.0857	0.0854
Lin's CCC	China	All data	[illegible]	[illegible]	[illegible]	[illegible]
		Method 1	0.9820	0.9811	**[illegible]**	**[illegible]**
		Method 2	0.9888	0.9882	**[illegible]**	0.9851
	Spain	All data	[illegible]	[illegible]	[illegible]	[illegible]
		Method 1	0.9547	**0.9578**	0.9576	0.9572
		Method 2	0.9668	[illegible]	[illegible]	[illegible]
	Singapore	All data	[illegible]	[illegible]	[illegible]	[illegible]
		Method 1	0.8377	**[illegible]**	[illegible]	[illegible]
		Method 2	0.9118	[illegible]	0.9116	[illegible]
Pearson's R	China	All data	[illegible]	[illegible]	[illegible]	[illegible]
		Method 1	0.9821	0.9812	**0.9839**	0.9838
		Method 2	0.9840	0.9840	**0.9856**	0.9848
	Spain	All data	[illegible]	[illegible]	[illegible]	[illegible]
		Method 1	0.9536	**0.9587**	0.9579	0.9581
		Method 2	[illegible]	**[illegible]**	[illegible]	[illegible]
	Singapore	All data	[illegible]	[illegible]	[illegible]	[illegible]
		Method 1	0.9382	**0.9400**	0.9390	0.9378
		Method 2	0.9527	**0.9603**	0.9557	0.9550

- The constrained ... lt in lower observed model fit than t...
- In cross-validat... from Spain, Singapore, and ... els displayed higher out-of-sample ... 20-parameter model
- Performance fo... under investigation

EQ-5D

Generalization of the non-linear models

- The *hyreg* command does not presently support non-linear models
- Built-in functions in STATA and R do not support non-linear interval regression
- A package allowing non-linear hybrid models (with and without intervals) has been developed for R, but is not documented to the point where it can be made publicly available.
- We have not found STATA functions that do not allow fitting of non-linear mixed effects models (i.e. random intercept/random slope models). This is possible using the *nlme* package in R.
- The upcoming hybrid function for R currently does not support mixed effects models.

www.euroqol.org

EQ-5D

Modelling TTO & DCE data using a hybrid model

Mark Oppe

oppe@euroqol.org

www.euroqol.org

TTO and DCE methods

- Individuals have a utility function which determines their preferences over health states
- TTO & DCE methods both try to measure the same utility function
- TTO & DCE each have their own weaknesses
  - e.g. scale compatibility (BTD vs WTD) for C-TTO
  - e.g. no anchors for use in QALY calculations for DCE
- Which method should we choose?

www.euroqol.org

TTO, DCE or both?

- TTO: trade-off between quality of life and length of life
  - How many years are you willing to give up to avoid being in impaired health?
- DCE: trade-off between quality of life and quality of life
  - Which health state is better?
- Both questions provide relevant information
- View TTO and DCE as complementary sources of information instead of competing
  - Include both types of information in a single model (Maximum Likelihood)

www.euroqol.org

Log likelihood of basic hybrid model (OLS & clogit)

$$lnL = -\frac{1}{2} * \sum_{j \in C} \left\{ \ln(2\pi\sigma^2) + \left(\frac{y_j - x\beta}{\sigma}\right)^2 \right\}$$

$$+ \sum_{j \in D} \left\{ ln \left(\frac{1}{1 + e^{(-x\beta')}}\right) * y_j + ln \left(\frac{e^{(-x\beta')}}{1 + e^{(-x\beta')}}\right) * (1 - y_j) \right\}$$

proportional rescaling parameter θ, such that
β' = β * θ

Log likelihood of basic hybrid model (OLS & clogit)

$$lnL = -\frac{1}{2} * \sum_{j \in C} \left\{ \ln(2\pi\sigma^2) + \left(\frac{y_j - x\beta}{\sigma}\right)^2 \right\}$$

$$+ \sum_{j \in D} \left\{ ln \left(\frac{1}{1 + e^{(-x\beta * \theta)}}\right) * y_j + ln \left(\frac{e^{(-x\beta * \theta)}}{1 + e^{(-x\beta * \theta)}}\right) * (1 - y_j) \right\}$$

EQ-5D

TTO data structure

BLOCK_ID	STATE_ID	ORDER_ID	MO	SC	UA	PD	AD	ORIGINAL_RESPONSE	Value	METHOD	TASK_TIME_sec	TTO_moves	TTO_resets	Profile	Severity
-1	0	1	0	0	0	0	0	10.5	0.05	TTO	441.3	43	0	00000	0
-1	-1	2	2	1	1	2	1	14	0.4	TTO	123.0	4	0	21121	7
-1	-2	3	3	5	5	5	4	10.5	0.05	TTO	222.4	6	0	35554	22
-1	-3	4	1	5	4	1	1	2	-0.8	TTO	221.4	7	0	15411	12
2	10	5	1	2	1	2	1	18.5	0.85	TTO	87.8	8	0	12121	7
2	12	6	3	4	1	5	5	10	0	TTO	46.9	2	0	34155	18
7	83	7	1	1	2	1	1	20	1	TTO	56.1	9	0	11211	6
2	16	8	2	3	5	1	4	2	-0.8	TTO	63.5	7	0	23514	15
2	15	9	3	2	4	4	3	2	-0.8	TTO	58.3	7	0	32443	16
2	13	10	5	2	2	1	5	10	0	TTO	27.5	2	0	52215	15
2	9	11	1	2	5	4	3	2	-0.8	TTO	60.6	7	0	12543	15
2	11	12	4	3	5	4	2	1	-0.9	TTO	42.3	8	0	43542	18
10	86	13	5	5	5	5	5	0	-1	TTO	42.6	10	0	55555	25
2	14	14	4	5	1	3	3	14	0.4	TTO	40.6	4	0	45133	16

EQ-5D

DCE data structure

STATE_ID	ORDER_ID	MO	SC	UA	PD	AD	ORIGINAL_RESPONSE	Value	METHOD	TASK_TIME_sec	Profile	Severity
18	4	1	4	5	3	3	B	0	DCE_A	19.6	14533	16
18	4	2	1	5	4	2	A	1	DCE_B	19.6	21542	14
42	6	2	3	1	3	4	A	1	DCE_A	18.2	23134	13
42	6	1	4	3	1	4	B	0	DCE_B	18.2	14314	13
75	5	5	3	4	3	1	B	0	DCE_A	9.4	53431	16
75	5	5	2	2	5	5	A	1	DCE_B	9.4	52255	19
76	2	5	1	5	2	2	B	0	DCE_A	21.4	51522	15
76	2	4	5	2	4	4	A	1	DCE_B	21.4	45244	19
127	7	1	4	2	2	4	A	1	DCE_A	33.2	14224	13
127	7	3	2	3	2	2	B	0	DCE_B	33.2	32322	12
180	3	4	4	1	4	5	A	1	DCE_A	7.2	44145	18
180	3	4	5	4	3	2	B	0	DCE_B	7.2	45432	18
196	1	1	1	2	2	1	A	1	DCE_A	52.1	11221	7
196	1	2	2	1	2	2	B	0	DCE_B	52.1	22122	9

STATE_ID	ORDER_ID	MO A	SC A	UA A	PD A	AD A	MO B	SC B	UA B	PD B	AD B	ORIGINAL_RESPONSE	Value	METHOD	TASK_TIME_sec	Profile A	Severity A	Profile B	Severity B
18	4	1	4	5	3	3	2	1	5	4	2	B	0	DCE_A	19.6	14533	16	21542	14
42	6	2	3	1	3	4	1	4	3	1	4	A	1	DCE_A	18.2	23134	13	14314	13
75	5	5	3	4	3	1	5	2	2	5	5	B	0	DCE_A	9.4	53431	16	52255	19
76	2	5	1	5	2	2	4	5	2	4	4	B	0	DCE_A	21.4	51522	15	45244	19
127	7	1	4	2	2	4	3	2	3	2	2	A	1	DCE_A	33.2	14224	13	32322	12
180	3	4	4	1	4	5	4	5	4	3	2	A	1	DCE_A	7.2	44145	18	45432	18
196	1	1	1	2	2	1	2	2	1	2	2	A	1	DCE_A	52.1	11221	7	22122	9

Data structure for HYREG command in Stata

state/pair	method	datatype	EQ state	value	mo2	mo3	mo4	mo5	sc2	sc3	…	ad5
1	TTO	1	22222	0.75	1	0	0	0	1	0	…	0
20	DCE_A	0	53121	0	0	0	0	1	0	1	…	0
20	DCE_B	0	32122	1	0	1	0	0	1	0	…	0
21	DCE_A	0	12345	1	0	0	0	0	1	0	…	1
21	DCE_B	0	23435	0	1	0	0	0	0	1	…	1

TTO & DCE: value* = 1- value

DCE: mo2* … ad5* = mo2_A – mo2_B … ad5_A – ad5_B
TTO: mo2* … ad5* = mo2 … ad5

state/pair	method	datatype	EQ state	value*	mo2*	mo3*	mo4*	mo5*	sc2*	sc3*	…	ad5*
1	TTO	1	22222	0.25	1	0	0	0	1	0	…	0
20	DCE_A-B	0	…	1	0	-1	0	1	0	-1	…	0
21	DCE_A-B	0	…	0	-1	0	0	0	1	-1	…	0

www.euroqol.org

Hybrid OLS combined with Conditional logistic model

- **Model assumptions**
  - Same as OLS for TTO and same as conditional logit for DCE
- **Data structure**

state/pair	method	datatype	EQ state	value*	mo2*	mo3*	mo4*	mo5*	sc2*	sc3*	…	ad5*
1	TTO	1	22222	0.25	1	0	0	0	1	0	…	0
20	DCE_A-B	0	…	1	0	-1	0	1	0	-1	…	0
21	DCE_A-B	0	…	0	-1	0	0	0	1	-1	…	0

- **Stata code**
  - **hyreg** value* mo2*-ad5*, datatype(datatype) contdist(normal) dichdist(logistic) noconstant

www.euroqol.org

Hybrid censored at -1 combined with conditional logistic model

- **Model assumptions**
  – Same as Tobit for TTO and same as conditional logit for DCE

- **Data structure**

state/pair	method	datatype	EQ state	value*	mo2*	mo3*	mo4*	mo5*	sc2*	sc3*	…	ad5*
1	TTO	1	22222	0.25	1	0	0	0	1	0	…	0
20	DCE_A-B	0	…	1	0	-1	0	1	0	-1	…	0
21	DCE_A-B	0	…	0	-1	0	0	0	1	-1	…	0

- **Stata code**
  – **hyreg** value* mo2*-ad5*, datatype(datatype) contdist(normal) dichdist(logistic) noconstant ll(-1)

Apples, oranges or a fruit salad?

- TTO: trade-off between quality of life and length of life
- DCE: trade-off between quality of life and quality of life
- Hybrid:
  – Uses all available information
  – Estimates are between those of TTO and those of DCE
  – DCE can help mitigate issues present in TTO and v.v.
- Since the "true" utilities are not known, ultimately the choice is a normative one:
  – Which (imperfect) utility theory?
  – Which (imperfect) data collection technique?
- Pragmatic basis for choice: data quality; value range; performance in applications (e.g. discriminative power)

Modelling TTO & DCE data using a advanced hybrid models

Juan M. Ramos-Goñi

jramos@euroqol.org

www.euroqol.org

We have seen:

Hybrid OLS combined with Conditional logistic model and sensoring at -1 for all responses

Basic

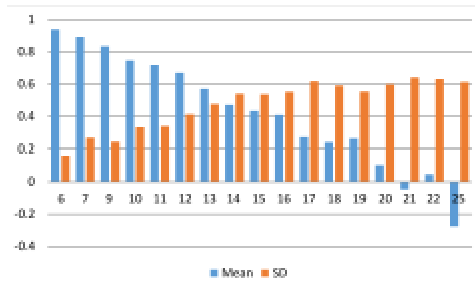
- Model assumptions
  - Same as OLS for C-TTO and same as conditional logit for DCE
  - Homoscedastic variance
- Stata code:
  - **hyreg** value dummy_variables, datatype(method_dummy)

Tobit

- Model assumptions
  - Same as tobit for C-TTO and same as conditional logit for DCE
  - Homoscedastic variance
- Stata code:
  - **hyreg** value dummy_variables, datatype(method_dummy) ll(-1)

www.euroqol.org

Does the homoscedasticity assumption fit in the case of EQ-VT data?



- As more severe is a state more discrepancy about its value

Hybrid models counting for heteroscedastic data

- In non heteroscedastic model, the variance is estimated as a single parameter. The idea behind heteroscedastic models is to model the variance as a function of the parameters which impacts on its value

- Model assumptions
  - Similar to OLS for C-TTO and similar to conditional logit for DCE
  - Heteroscedastic variance

- Stata code:
  - **hyreg** value dummy_variables, datatype(method_dummy) *hetcont(dummy_variables') hetdich(dummy_variables'')*

Hybrid intervals combined with Conditional logistic model

- Model assumptions
  - Similar to interval regression for C-TTO and similar to conditional logit for DCE
  - Heteroscedastic variance

- Stata code:
  - **hyreg** *ll_value ul_value* dummy_variables, datatype(method_dummy) *intervals*
    hetcont(dummy_variables') hetdich(dummy_variables'')

www.euroqol.org

Using intervals for censoring specific C- TTO responses on a hybrid model

- Model assumptions
  - Similar to interval regression for C-TTO and similar to conditional logit for DCE
  - Heteroscedastic variance

- Stata code:
  - **hyreg** *ll_value ul_value* dummy_variables, datatype(method_dummy) *intervals*

  hetcont(dummy_variables') hetdich(dummy_variables'')

Type of data		depvar1	depvar2
point data	a = [a,a]	a	a
interval data	[a,b]	a	b
left-censored data	(-inf,b]	.	b
right-censored data	[a,inf)	a	.

www.euroqol.org

Why we may be interested in censoring/truncating specific responses

- Lack of WTD explanation by the interviewer
- Specific interviewer/respondent behaviour during the interview leading to crude preferences rather than accurate stated preferences

state/pair	method	datatype	EQ state	value*	ll_value	ul_value	..............Dummies.....
1	TTO	1	22222	0	.	0	....
2	TTO	1	…	-1	.	-1	
3	TTO	0	…	0.6	0.5	0.7	

Hybrid models assuming different distributions for C-TTO or DCE data

- Assumptions
  - Similar to OLS for C-TTO and similar to conditional logit for DCE
  - Heteroscedastic variance
- Stata code:
  - **hyreg** ll_value ul_value dummy_variables, datatype(method_dummy) intervals

  contdist(normal|logistic) dichcont(normal|logistic) hetcont(dummy_variables') hetdich(dummy_variables'')

Questions

Thanks for attending to this workshop

www.euroqol.org