PAIRED COMPARISONS

MODELS AND APPLICATIONS

Regina Dittrich
Reinhold Hatzinger

REAL PAIRED COMPARISONS (PC)

a method of data collection where individuals are asked to judge a number of different pairs of objects, taken from a larger set of J objects.

for each comparison between two objects j and k, the individual can respond

j preferred to k

k preferred to j

aim is to rank objects into a preference order – obtain an overall ranking of the objects

common in food tasting, and in social and management surveys, …

THE BASIC BRADLEY TERRY MODEL (BT- MODEL)

for the comparison (jk) of object j to object k:

we observe $N_{(jk)j}$, the number of times where j is preferred to k
we observe $N_{(jk)k}$, the number of times where k is preferred to j

$n_{(jk)}$ is the number of times this comparison was performed, i.e.,
$N_{(jk)j} + N_{(jk)k}$

let $p_{(jk)j}$ be the probability that j is preferred to k in comparison (jk)

the BT-model is

$$p_{(jk)j} = \frac{\pi_j}{\pi_j + \pi_k}$$

π_j are a set of so-called worth parameters
the worth parameters are constrained to sum to one

THE LOG-LINEAR REPRESENTATION OF THE BT–MODEL LLBT

The Model can be formulated as a log-linear model following the usual Multinomial / Poisson - equivalence.

the expected value $m_{(jk)j}$ of $N_{(jk)j}$ is given as $m_{(jk)j} = n_{(jk)} p_{(jk)j}$

$$p_{(jk)j} = \frac{\pi_j}{\pi_j + \pi_k} = \frac{\sqrt{\pi_j/\pi_k}}{\sqrt{\pi_k/\pi_j} + \sqrt{\pi_j/\pi_k}}$$

then our basic paired comparison model (PC Model) for one comparison is

$$\ln m_{(jk)j} = \mu_{(jk)} + \lambda_j - \lambda_k$$
$$\ln m_{(jk)k} = \mu_{(jk)} - \lambda_j + \lambda_k$$

λ_j are the object parameters
$\mu_{(jk)}$ are nuisance parameters

this model formulation is feasible for further extensions

DESIGN STRUCTURE LLBT - 2 RESPONSES

- for 3 objects we have 3 comparisons

comparison	*decision*	*counts*	μ	λ_1^O	λ_2^O	λ_3^O
(12)	O_1	$N_{(12)1}$	1	1	-1	0
(12)	O_2	$N_{(12)2}$	1	-1	1	0
(13)	O_1	$N_{(13)1}$	2	1	0	-1
(13)	O_3	$N_{(13)3}$	2	-1	0	1
(23)	O_2	$N_{(23)2}$	3	0	1	-1
(23)	O_3	$N_{(23)3}$	3	0	-1	1

EXTENSIONS TO THE BASIC PC MODEL - RESPONSES

allow for ties (trichotomous responses, nominal)
undecided

for each comparison between two objects j and k , the response can be

j preferred to k

no preference
undecided

k preferred to j

Using the respecification of the probabilities suggested by Davidson and Beaver (1977):

the paired comparison model (PC Model) is now

$$\ln m_{(jk)j} = \mu_{(ik)} + \lambda_j - \lambda_k$$
$$\ln m_{(jk)0} = \mu_{(ik)} + \gamma$$
$$\ln m_{(jk)k} = \mu_{(ik)} - \lambda_j + \lambda_k$$

where γ is the undecided parameter

λ´s are the object parameters
$\mu_{(jk)}$ are nuisance parameters

DESIGN STRUCTURE LLBT- UNDECIDED

comparison	*decision*	*counts*	μ	γ	λ_1^O	λ_2^O	λ_3^O
(12)	O_1	$N_{(12)1}$	1	0	1	-1	0
(12)	*no*	$N_{(12)0}$	1	1	0	0	0
(12)	O_2	$N_{(12)2}$	1	0	-1	1	0
(13)	O_1	$N_{(13)1}$	2	0	1	0	-1
(13)	*no*	$N_{(13)0}$	2	1	0	0	0
(13)	O_3	$N_{(13)3}$	2	0	-1	0	1
(23)	O_2	$N_{(23)2}$	3	0	0	1	-1
(23)	*no*	$N_{(23)0}$	3	1	0	0	0
(23)	O_3	$N_{(23)3}$	3	0	0	-1	1

EXTENSIONS TO THE BASIC PC MODEL – STRUCTURE

position effects

Is there an effect due to the ordering of the presentation of objects?
Does it make a difference ?

MANCHESTER UNITED – INTER MILAN
(playing in Manchester)

INTER MILAN – MANCHESTER UNITED
(playing in Milan)

introduce additional parameter(s) for position effects

EXTENSIONS TO THE BASIC PC MODEL – COVARIATES

OBJECT COVARIATES

(Dittrich, Hatzinger, Katzenbeisser, J. Royal Statistical Society, C, 1998)

to model the objects by a few characteristics

$$\lambda_j^O = \sum_{q=1}^{Q} \beta_q x_{jq}$$

x_{jq} covariate for characteristic q of object j
β_q effect of characteristic q

CATEGORICAL SUBJECT COVARIATES

(Dittrich, Hatzinger, Katzenbeisser, J. Royal Statistical Society, C, 1998)

CONTINOUS SUBJECT COVARIATES

one contingency table for each subject

SMOOTHED SUBJECT COVARIATES (GAMS)

(Francis, Dittrich, Hatzinger, Penn, J. Royal Statistical Society, C, 2002)

CATEGORICAL SUBJECT COVARIATES

(Dittrich, Hatzinger, Katzenbeisser, J. Royal Statistical Society, C, 1998)

Are the preference orderings different for different groups of subjects?

For one categorical subject covariate we now have

$$\ln m_{(jk)j|s} = \mu_{(jk)s} + \lambda_s^S + (\lambda_j^{O_j} + \lambda_{js}^{O_j S}) - (\lambda_k^{O_k} + \lambda_{ks}^{O_k S})$$

where

$\lambda_j^{O_j}$ object parameter
$\lambda_{js}^{O_j S}$ interaction parameter between object j and subject category s
λ_s^S fixing the margin for category s of covariate S (nuisance)

$\mu_{(jk)s}$ nuisance parameters

DESIGN-STRUCTURE LLBT- 3 RESPONSES (UNDECIDED)

one subject covariate *S* on two levels

comparison	*decision*	*counts*	μ	λ^S	γ	λ_1^O	λ_2^O	λ_3^O	λ_{12}^{OS}	λ_{22}^{OS}	λ_{32}^{OS}
(12)	O_1	$N_{(12)1\|1}$	1	0	0	1	-1	0	0	0	0
(12)	no	$N_{(12)0\|1}$	1	0	1	0	0	0	0	0	0
(12)	O_2	$N_{(12)2\|1}$	1	0	0	-1	1	0	0	0	0
⋮	⋮	⋮	⋮					⋮			
(12)	O_1	$N_{(12)1\|2}$	4	1	0	1	-1	0	1	-1	0
(12)	no	$N_{(12)0\|2}$	4	1	1	0	0	0	0	0	0
(12)	O_2	$N_{(12)2\|2}$	4	1	0	-1	1	0	-1	1	0
⋮	⋮	⋮	⋮					⋮			

EXAMPLE: UNIVERSITY PREFERENCES

CEMS – exchange programme

students of the WU can study abroad visiting one of currently 17 CEMS universities

aim of the study:

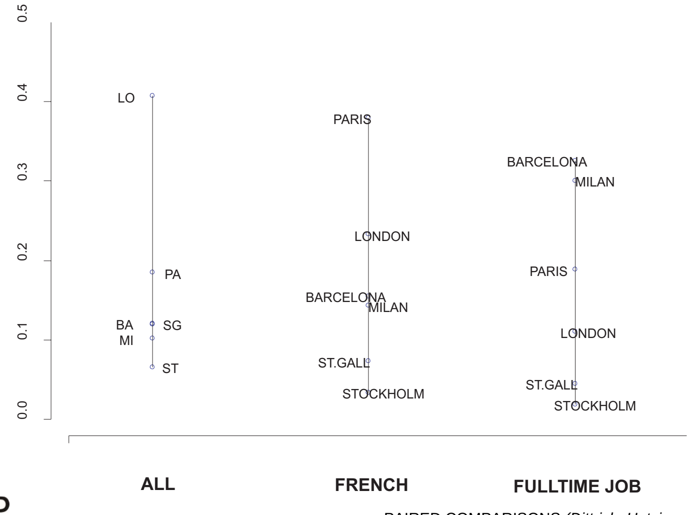
- preference orderings of students for different locations
- identify reasons for these preferences

data:

- PC-responses allowing for ties about their choices of 6 selected CEMS universities for the semester abroad (London, Paris, Milan, Barcelona, St.Gall, Stockholm)
- several covariates (e.g., gender, working status, language abilities, …)

Main results:

- in general London by far the most popular place
- depending on language abilities according places move up
- full-time working students prefer Latin universities



PATTERN MODELS:

(Dittrich, Hatzinger, Katzenbeisser, Comp.Stat.&Data Analysis, 2002)

- one important feature of the pattern models:
  - o dependencies between decisions can be taken into account
- in general: we look at response pattern vectors

$$\boldsymbol{Y} = (Y_{(12)}, Y_{(13)}, \dots\dots, Y_{(1J)}, \dots\dots\dots, Y_{(J-1,J)})$$

- we assume that dependencies between responses come from repeated evaluation of the same objects in paired comparisons
  - o comparing (object j with object k) and (object j with object l)
  - o the assessment of object j might be similar in both comparisons

PATTERN MODELS (NO UNDECIDED)

$$Y_{(jk)} = \begin{cases} -1 & \text{if object } O_k \text{ is preferred over } O_j \\ 1 & \text{if object } O_j \text{ is preferred over } O_k \end{cases}$$

$$P(Y_{(jk)} = 1) = \frac{\pi_j}{\pi_j + \pi_k} = \frac{\sqrt{\pi_j/\pi_k}}{\sqrt{\pi_k/\pi_j} + \sqrt{\pi_j/\pi_k}} = c\,\frac{\sqrt{\pi_j}}{\sqrt{\pi_k}}$$

We model the probability for a response pattern:

$$P(\boldsymbol{Y} = \boldsymbol{y}) = \Delta \prod_{j<k} \left(\frac{\sqrt{\pi_j}}{\sqrt{\pi_k}} \right)^{y_{(jk)}}$$

where Δ is a normalizing constant

ONE PATTERN (3 OBJECTS)

The probability for a pattern is given by:

$$p(y_{12}, y_{13}, y_{23}) = C^* \left(\frac{\sqrt{\pi_1}}{\sqrt{\pi_2}}\right)^{y_{12}} \left(\frac{\sqrt{\pi_1}}{\sqrt{\pi_3}}\right)^{y_{13}} \left(\frac{\sqrt{\pi_2}}{\sqrt{\pi_3}}\right)^{y_{23}}$$

the log-linear pattern model can be written as:

$$\ln m_\ell(y_{12}, y_{13}, y_{23}) = \gamma + (y_{12} + y_{13})\lambda_1 + (y_{23} - y_{12})\lambda_2 + (-y_{13} - y_{23})\lambda_3$$

For a certain response pattern (1,1,1) it is:

$$\ln m(1,1,1) = \gamma + 2\lambda_1 - 2\lambda_3$$

to model dependencies we include terms of the form

$$\exp(\theta_{(jk),(jl)} \ \ y_{(jk)} y_{(jl)})$$

where pairs of comparisons have one object in common

the pattern model including dependencies for one pattern

$$\ln m(y_{12}, y_{13}, y_{23}) = \gamma + (y_{12} + y_{13})\lambda_1 + (y_{23} - y_{12})\lambda_2 + (-y_{13} - y_{23})\lambda_3 \\ + \theta_{1,23}\, y_{12} y_{13} + \theta_{2,13}\, y_{12} y_{23} + \theta_{3,12}\, y_{13} y_{23},$$

DESIGN STRUCTURE

					γ	λ_1	λ_2	λ_3	$\theta_{12.13}$	$\theta_{12.23}$	$\theta_{13.23}$
pattern	y_{12}	y_{13}	y_{23}	counts	const	x_1	x_2	x_3	$y_{12}y_{13}$	$y_{12}y_{23}$	$y_{13}y_{23}$
s_1	1	1	1	n_1	1	2	0	-2	1	1	1
s_2	1	1	−1	n_2	1	2	-2	0	1	-1	-1
s_3	1	−1	1	n_3	1	0	0	0	-1	1	-1
s_4	1	−1	−1	n_4	1	0	-2	2	-1	-1	1
s_5	−1	1	1	n_5	1	0	2	-2	-1	-1	1
s_6	−1	1	−1	n_6	1	0	0	0	-1	1	-1
s_7	−1	−1	1	n_7	1	-2	2	0	1	-1	-1
s_8	−1	−1	−1	n_8	1	-2	0	2	1	1	1

PATTERN MODELS - PC-RESPONSES (WITH UNDECIDED)

- This extension is straightforward
- We have a third response category and therefore the number of patterns is (3 ^ number of comparisons)

FOR ALL PATTERN MODELS - PC-RESPONSES

extensions mentioned for the LLBT Model also possible

- **OBJECT COVARIATES**
- **POSITION EFFECT**
- **CATEGORIAL SUBJECT COVARIATES**
- **CONTINOUS SUBJECT COVARIATES**
- **SMOOTHED SUBJECT COVARIATES**

(Contingency tables might become very large!)

OTHER RESPONSE FORMATS:

DERIVED PAIRED COMPARISONS (PC)

Ordinal response formats are transformed into paired comparisons

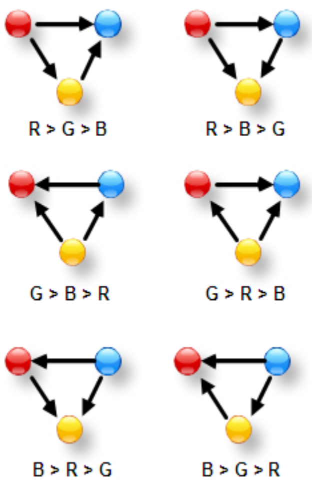
FULL RANKINGS :

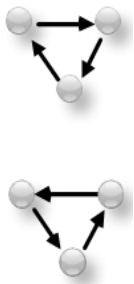
- o people are asked to rank objects (items) regarding a certain aspect
- o all possible pairs are constructed
- o the one with higher rank will be judged as preferred
- o no undecided category !

Comparison Pattern Models: Rankings - PC

Data				comparison		
R	G	B	Response	RG	RB	GB
1	2	3	R>G>B	1	1	1
1	3	2	R>B>G	1	1	-1
-	-	-	-	1	-1	1
2	3	1	B>R>G	1	-1	-1
2	1	3	G>R>B	-1	1	1
-	-	-	-	-1	1	-1
3	1	2	G>B>R	-1	-1	1
3	2	1	B>G>R	-1	-1	-1

Transitive Patterns





Intransitive Patterns

bei Rankings nicht möglich

PATTERNMODELS - RANKINGS

DESIGN STRUCTURE

	ranking-patterns			paired comparison-patterns			design structure				
	B	C	N		Y			X			
r_k	O_1	O_2	O_3	(12)	(13)	(23)	counts	const	x_1	x_2	x_3
r_1	1	2	3	1	1	1	n_1	1	2	0	-2
r_2	1	3	2	1	1	-1	n_2	1	2	-2	0
r_3	2	1	3	-1	1	1	n_3	1	0	2	-2
r_4	2	3	1	1	-1	-1	n_4	1	0	-2	2
r_5	3	1	2	-1	-1	1	n_5	1	-2	2	0
r_6	3	2	1	-1	-1	-1	n_6	1	-2	0	2

we have only 6 possible patterns - compared to 8 when using real PC !

OTHER RESPONSE FORMATS:

DERIVED PAIRED COMPARISONS (PC)

RATINGS

- o respondents are ask to rate items (objects) where the responses are typically defined by endpoints like ‘very important´ - ‘very unimportant´
- o all possible pairs are constructed
- o the one with higher rating will be judged as preferred
- o undecided category is always included !

- o meaningful, if relative ordering (preference) is of primary interest

EXAMPLE: RATINGS

We used a data set collected by the British Household Panel Study in 1996 where we have chosen three Likert items which ask respondents about the concern about:

- the destruction of the ozone layer (OZ)
- the high rate of unemployment (UN)
- declining moral standards (MO)

The possible answers are:

- A great deal 1
- A fair amount 2
- Not very much 3
- Not at all 4

Low numbers mean a high concern and higher number lower concern!

EXAMPLE (CONT.)

Converting the data to paired comparison form

We compare each pair of items and

- if 1st of a pair gets the lower score, the 1st item is preferred – coded 1
- if 1st of a pair gets the higher score, the 2nd item is preferred – coded -1
- if they have equal scores, there is no preference (undecided) – coded 0

PATTERN MODELS - RATING

Assuming independence the probability for one comparison is

$$p_{ij} = P\{y_{ij}\} = P\{Y_{ij} = y_{ij}\} = \begin{cases} c\left(\frac{\sqrt{\pi_i}}{\sqrt{\pi_j}}\right)^{y_{ij}} = c \quad \frac{\sqrt{\pi_i}}{\sqrt{\pi_j}}, & \text{if } y_{ij} = 1 \\ c\left(\frac{\sqrt{\pi_i}}{\sqrt{\pi_j}}\right)^{y_{ij}} = c \quad \frac{\sqrt{\pi_j}}{\sqrt{\pi_i}}, & \text{if } y_{ij} = -1 \\ c \quad u_{ij}\,, & \text{if } y_{ij} = 0\,, \end{cases}$$

We model the probability for a response pattern:

$$P(\boldsymbol{Y} = \boldsymbol{y}) = \Delta \prod_{j<k} \left(\frac{\sqrt{\pi_j}}{\sqrt{\pi_k}}\right)^{y_{(jk)}}$$

where Δ is a normalizing constant

PATTERN MODELS - RATING

DESIGN STRUCTURE

Rating patterns			derived PC-patterns			unique PC-patterns			counts	design structure X						
										γ	λ_1	λ_2	λ_3	δ_{12}	δ_{13}	δ_{23}
i_1	i_2	i_3				y_{12}	y_{13}	y_{23}			x_1	x_2	x_3	u12	u13	u23
1	1	1	0	0	0	0	0	0	n_1	1	0	0	0	1	1	1
1	1	2	0	1	1	0	1	1	n_2	1	1	1	-2	1	0	0
1	2	1	1	0	-1	1	0	-1	n_3	1	1	-2	1	0	1	0
1	2	2	1	1	0	1	1	0	n_4	1	2	-1	-1	0	0	1
2	1	1	-1	-1	0	-1	-1	0	n_5	1	-2	1	1	0	0	1
2	1	2	-1	0	1	-1	0	1	n_6	1	-1	2	-1	0	1	0
2	2	1	0	-1	-1	0	-1	-1	n_6	1	-1	-1	2	1	0	0
2	2	2	0	0	0											

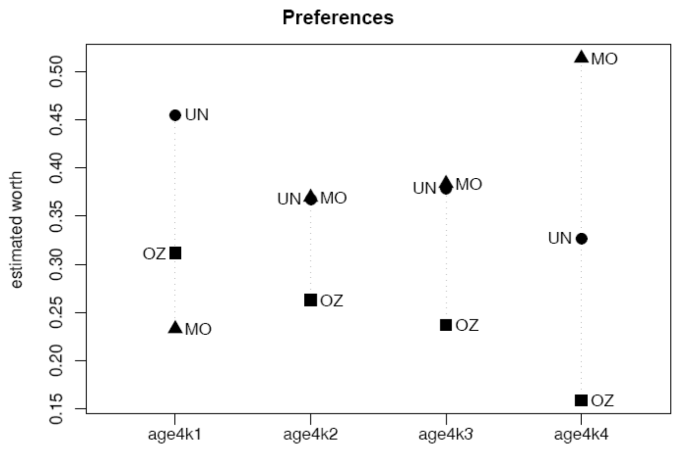
- we have only 7 possible patterns - compared to 8 when using real PC !
- additionally we have undecided parameters (here for each comparison)

PATTERN MODELS - RATING

DESIGN STRUCTURE
including dependencies

unique PC-patterns			counts	design structure X									
				γ	λ_1	λ_2	λ_3	δ_{12}	δ_{13}	δ_{23}	$\theta_{12.13}$	$\theta_{12.23}$	$\theta_{13.23}$
y_{12}	y_{13}	y_{23}			x_1	x_2	x_3	u12	u13	u23	$y_{12}y_{13}$	$y_{12}y_{23}$	$y_{13}y_{23}$
0	0	0	n_1	1	0	0	0	1	1	1	0	0	0
0	1	1	n_2	1	1	1	-2	1	0	0	0	0	1
1	0	-1	n_3	1	1	-2	1	0	1	0	0	-1	0
1	1	0	n_4	1	2	-1	-1	0	0	1	1	0	0
-1	-1	0	n_5	1	-2	1	1	0	0	1	1	0	0
-1	0	1	n_6	1	-1	2	-1	0	1	0	0	-1	0
0	-1	-1	n_6	1	-1	-1	2	1	0	0	0	0	1

Results, when including subject covariate AGE (4 levels)



A PAIRED COMPARISON APPROACH FOR THE ANALYSIS OF ORDINAL PC DATA

(Dittrich, Francis, Hatzinger, Katzenbeisser,J. Statistical Modelling, 2004)

- *Motives of students to obtain a doctoral degree at the WU:*
  title, scientific research, career,
  state of transition to work, further education

  in each comparisons the students had to decide if:
  motive i is much more important, more important, slightly more important or
  motive j is much more important, more important, slightly more important

- *Importance of motives for various groups of students:*
  faculty members (yes/no), gender, break after master (yes/no)

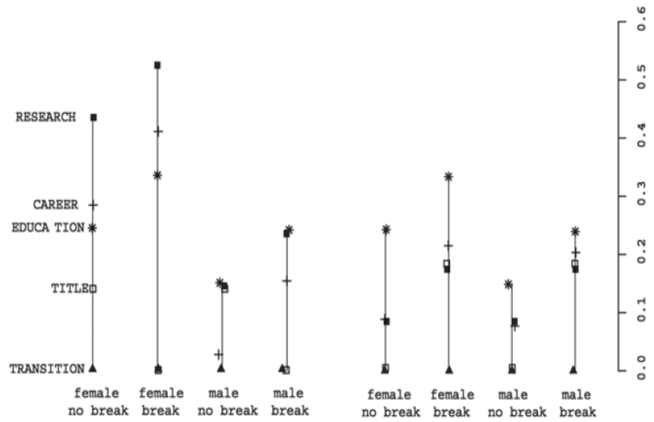


Figure 1 Estimated parameters for WU faculty (left part) and for not WU faculty (right part)

A LOG LINEAR PATTERN MODEL FOR MULTIVARIATE PAIRED COMPARISONS - (MULTIDIMENSIONAL)

(Dittrich, Francis, Hatzinger, Katzenbeisser, Mathematical Social Sciences 2006)

266 first year students at the WU were asked to compare 4 Austrian party leaders on more than one attribute (2001) – paired comparisons

attributes: competence in social issues
competence in economic issues

- dependencies between the decisions of the judges
- association structures between the attributes

party leaders: Social democrat
Conservative
Green
Freedom party

location of Austrian party leaders in two dimensional preference space (social and economic competence)

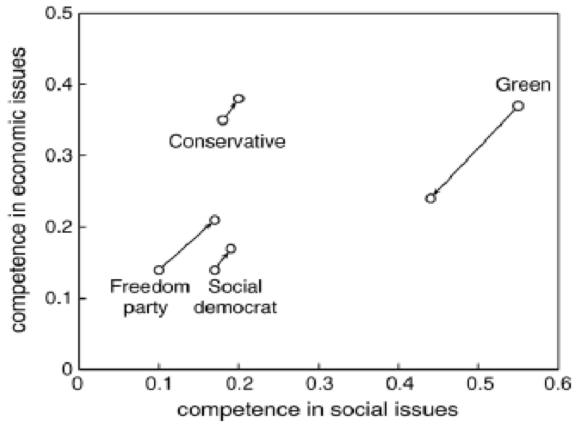


Fig. 2. Ascribed competence in social and economic issues.

A PAIRED COMPARISON APPROACH FOR THE ANALYSIS OF LIKERT SCALE DATA

(Dittrich, Francis, Hatzinger, Katzenbeisser,J. Statistical Modelling, 2007)

Data from ISSP 2000 (International Social Survey)

6 Likert-items:	air pollution caused by car (CAR)
	air pollution caused by industry (INDUSTRY)
	pesticides and chemicals used in farming (FARM)
	pollution of rivers, lakes (WATER)
	rise in world's temperature (TEMP)
	modifying genes of certain crops (GENE)

Responses to each of item: 5 points

(1) extremely dangerous for environment
(2) very dangerous
(3) somewhat dangerous
(4) not very dangerous
(5) not dangerous

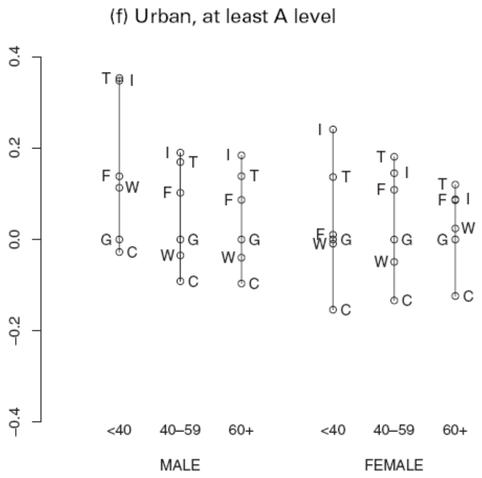
Subject covariates:

country: Austria, Great Britain
age: 1 = <40, 2 = 41-59 , 3 = 60+ years
sex
edu (level of education): 1 = below A-level/matrice, 2 = A-level/matrice or higher)

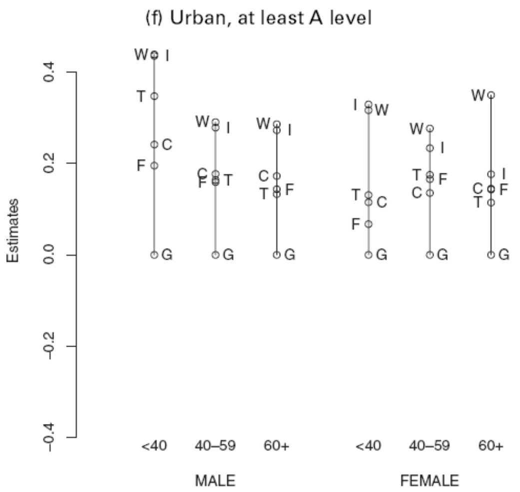
urb (locality of residence):
1 = urban area,
2 = suburbs of large cities, small town, county seat,
3 = rural area

Sample size: 782 Austria, 813 Great Britain

Austria



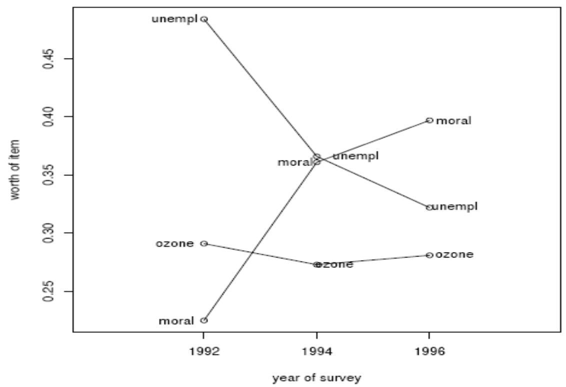
Great Britain



MODELLING REPEATED RATINGS

(Dittrich, Francis, Katzenbeisser (2008) Research Report Series, 67)

comparisons of the same objects by the same judges are made on more than one occasion
British Household Panel Study: follow-up of concerns about topical issues (unemployment, ozone layer, decline in moral standards)



MISSING DATA IN PC EXPERIMENTS

(Dittrich, Francis, Hatzinger, Katzenbeisser (2007) Proceedings of 22nd IWSM)

239 pupils and students in Vienna were asked to judge the importance of qualities of a good teacher in a complete PC-experiment

items: Clarity and structure of instructions
Activity Success un getting students to participate
Support Looking after every single pupil
Flexibility Waits until all have finished task

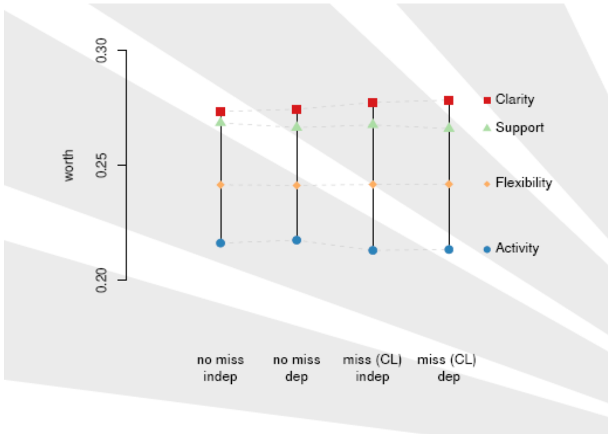
4 different models had been fitted:

- independence model – leaving out the missing responses
- dependence model – leaving out the missing responses

Composite Link models (CL) – including missing responses

- CL – independence model
- CL – dependence model

Comparison of the results when using different models:



MIXTURE MODELS IN RANKED DATA

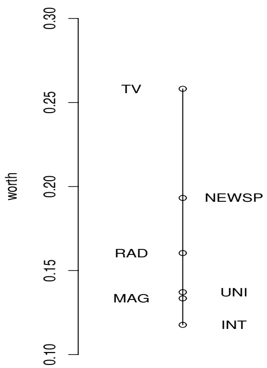
(Francis, Dittrich, Hatzinger (2005) 3rd World Conf. Comp.Stat.&Data Analysis)

nonparametric ML estimation of subject effects

Eurobarometer (2001) N=12,000 respondents

sources of information about scientific developments
ratings (likert items)

TV
Radio
Newspapers
Magazines
Scientific magazines
The internet
School/University



LATENT CLASS PC MODELS

(Francis, Dittrich, Hatzinger, Mair (2006) Int. Conf. on Latent Variable Models in Health Science)

nonparametric ML estimation of subject effects

aims of social work ***ranked*** *by first year social work students from Australia, Canada, UK, and USA*

68. Please order the following possible aims of social work in order, from 1 to 6 (1 being the most important and 6 the least important)

a) to work towards a more fair and just society .....
b) to provide care for people .....
c) to help people maintain relationships.....
d) to ensure that the law is upheld .....
e) to prevent harm to vulnerable people .....
f) to ensure that people's rights are exercised ..