Coding and Scaling Occupations in ESS R1-R5

Harry B.G. Ganzeboom

DANS ESS-workhop

Den Haag, 27 september 2012

ESS DEVO project

- Education: harmonize all country specific indicators and develop scaling
  - → ISLED International Standard Level of Education
- → Occupations
  - → Code all parental occupation into ISCO-88.
  - → Convert / code all occupations into ISCO-08
  - → Develop international standard measures for occupational status in ISCO-08.

Today

- Part 1: Coding parental occupations into ISCO-88
- Part 2: Convert / code all occupations into ISCO-08

Part 1: Coding parental occupations into ISCO-88

Occupations in ESS

- ESS collects data on:
  - Respondent's current / last occupation
  - Partner's current occupation
  - Father's * Mother's occupation when Resp 16
- Resp's & Partner's occupations are coded in ISCO-88. As of Round 6 this is going to change to ISCO-08.
- Father's and mother's occupation are NOT coded, but verbatim strings are available.
- Also, for father's and mother's occupation ESS collects information using a (crude) showcard.

Showcard R1-R3

Showcard R4-R5

Showcards

- The showcard R1-R3 is a particularly bad one to collect parental occupations:
  - Unclear labels
  - Categories unordered
  - No category for farm (!!).
- Showcard R4-R5 (taken from ISSP 1987) is better – this is basically the first digit of ISCO-88.

The sunny side

- We have a file with social mobility information in strictly comparable format for 34 countries.
- The social mobility information also covers mothers.
- We have double measurement for at least part of the SAT model.
- We have even access to occupation strings and can check on coding problems.

The dark side

- Over 360.000 occupations to code in some 25 languages.

	codingstatus					codingstatus			
	-1 No info	0 Tobecoded	1 Coded	Total		-1 No info	0 Tobecoded	1 Coded	Total
AT	917	3	8780	9700	IL	2218	0	9224	11442
BE	2255	810	10587	13652	IS	17	0	858	875
BG	452	69	10184	10705	IT	271	7	2832	3110
CY	743	16	4395	5154	LT	282	3	2902	3187
CZ	2086	0	10008	12094	LU	205	10	3782	3997
DE	4366	0	19448	23814	LV	291	0	5672	5963
DK	993	13	9147	10153	NL	1961	0	11295	13256
EE	1193	739	9069	11001	NO	208	0	12738	12946
ES	1637	835	10458	12930	PL	171	133	13781	14085
FI	285	0	15939	16224	PT	3943	13	13491	17447
FR	165	67	12365	12597	RO	3917	1	4652	8570
GB	2077	0	14711	16788	RU	1985	0	11259	13244
GR	37	1	9485	9523	SE	3861	0	13588	17449
HR	16	0	2924	2940	SI	1185	0	4760	5945
HU	1999	0	10862	12861	SK	54	1	7283	7338
CH	5712	17	11887	17616	TR	48	4	3244	3296
IE	3317	0	11109	14426	UA	175	2	8775	8952
					Total	49042	2744	311494	363280

The coding project (ISCO-88)

- All parental occupation strings are collected in “coding files”.
- These coding files are matched with existing codes for previous rounds and (where available) string information on resp and partner.
- We then recruited coders familiar with (mostly: native in) the language and trained them in the use of ISCO-88.

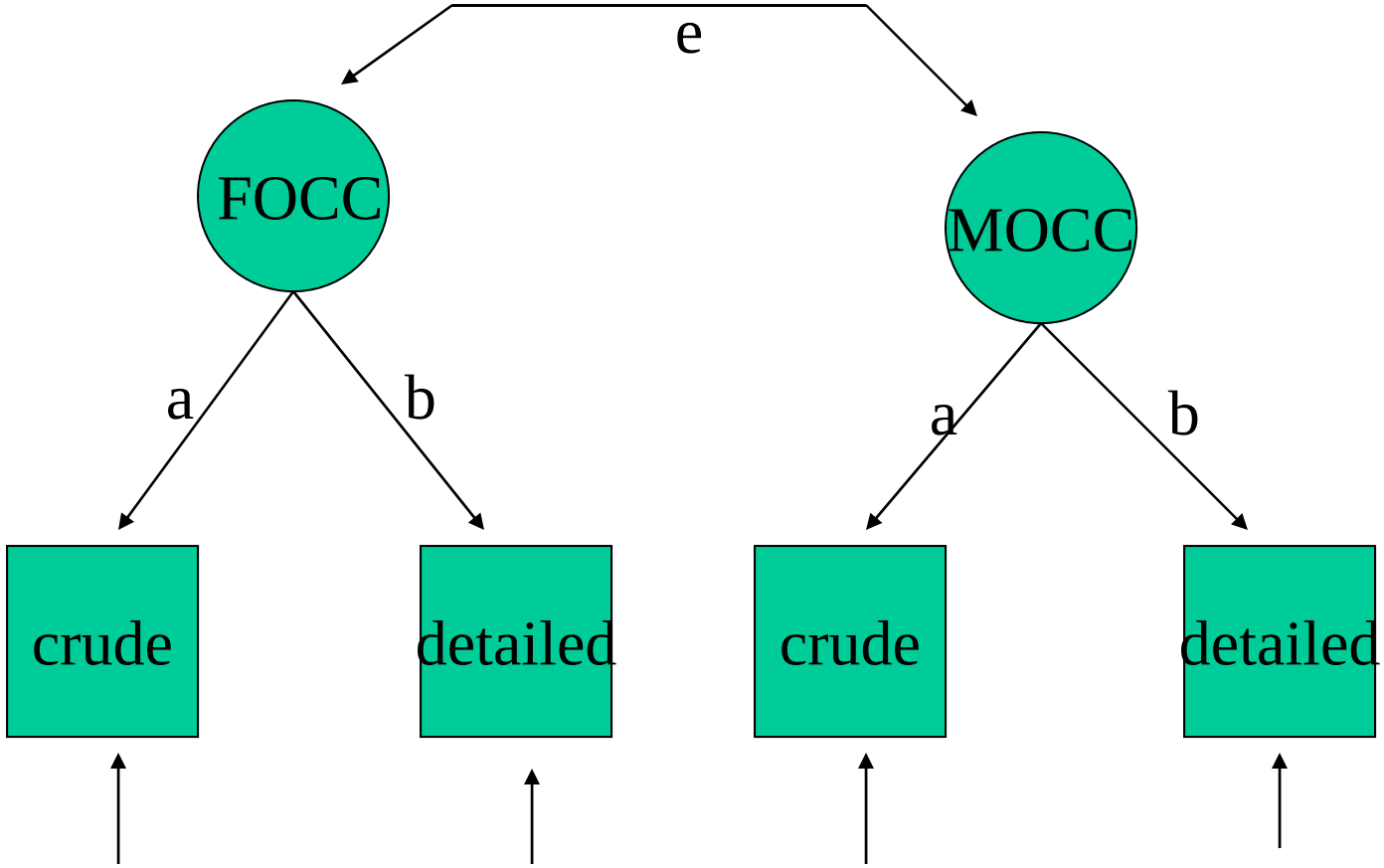
Coders

- Trained in 2 hour session.
- Materials: ISCO-88 overview, ILO website.
- Coders were encouraged to look for local versions of ISCO-88.
- Payment: 200 occupations per hour is the standard rate.
- In a few instances we had multiple coders per file, so that we can compare dtectly between coders.

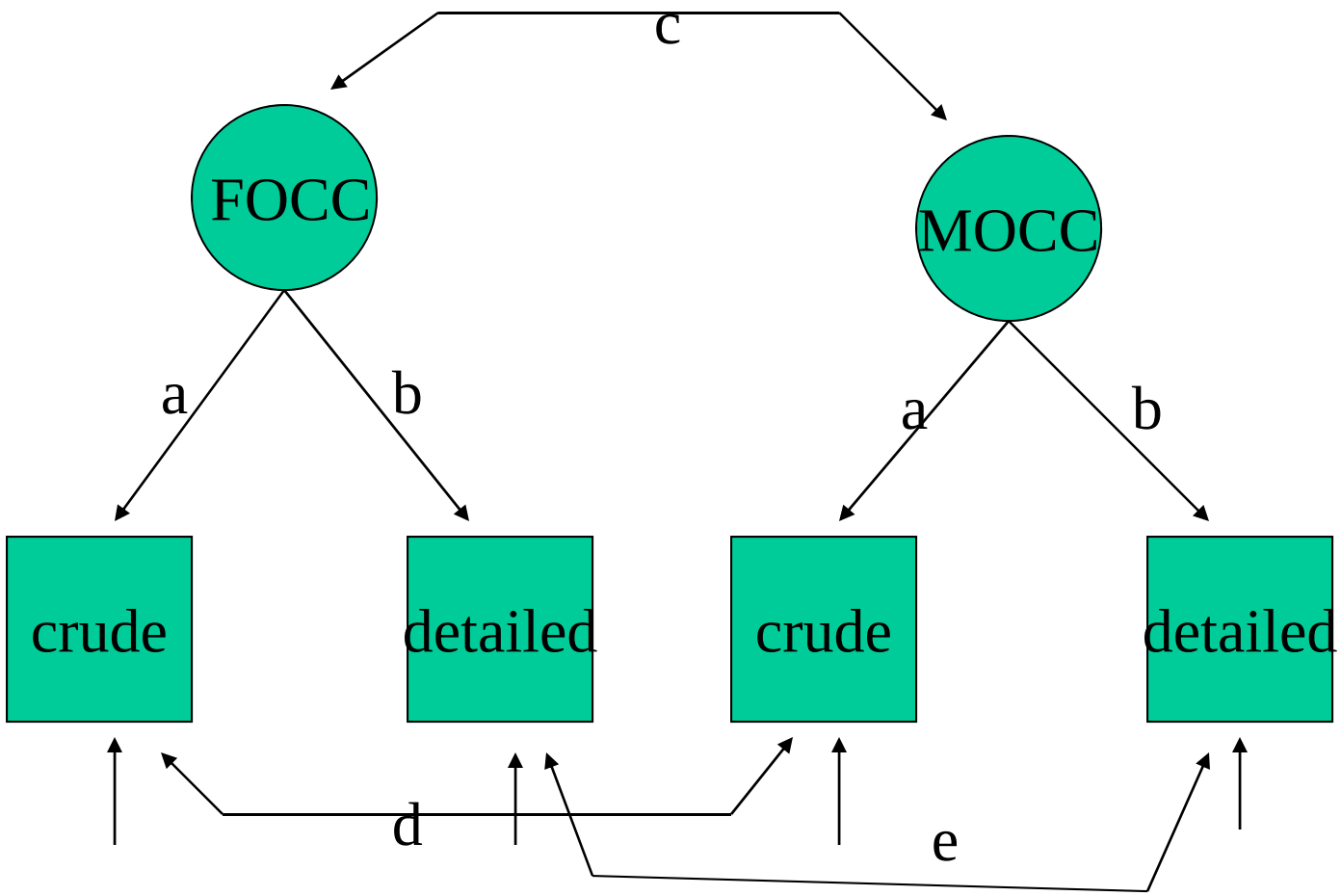
How can you check the quality of coders (if they coded different parts of the file)?

- (Double) coders should have very high correlation (or agreement?) for overlapping parts of the file.
- Coders should produce the same correlation with criteria they do not have access to: the other occupations, educations (status attainment structure).
- In ESS, we have also the correlations with the showcard on parental occupations.

Latent variable model

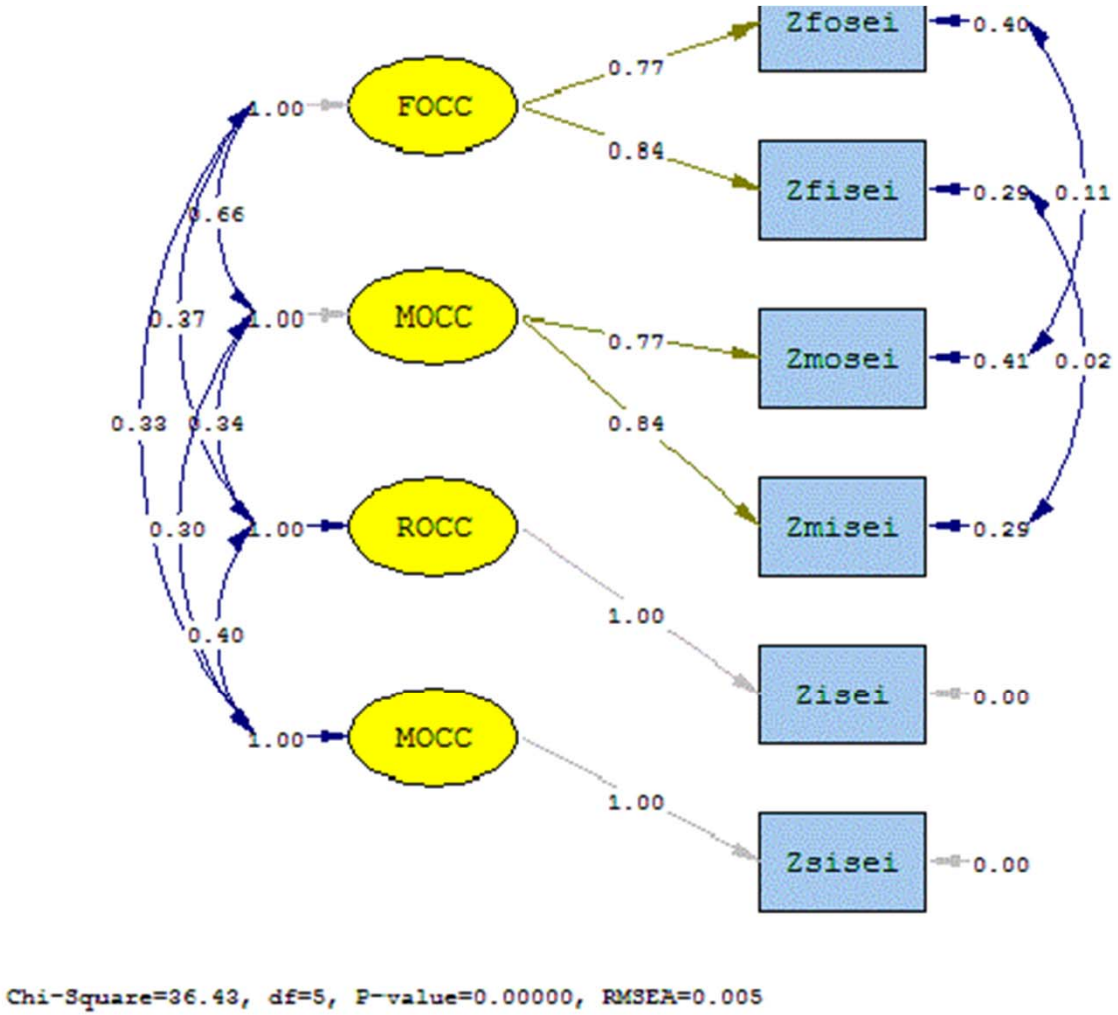


MTMM model



MTMM model

- The simple MTMM model with correlated residuals is not identified in 2x2 design.
- However all effects are identified when we add more variables to the model, in particular:
  - Occupations of respondent and spouse
  - Educations of respondent, spouse, father & mother.



Chi-Square=36.43, df=5, P-value=0.00000, RMSEA=0.005

Part 2: Converting / coding all occupations into ISCO-08

ISCO-08

- ISCO-08 is a “minor” revision of ISCO-88.
- Logic stays more or less the same, but all numbers change.
- Sociologically important revisions:
  - Manual supervisors
  - Managers
  - Shopkeepers

ISEI-08

- ISEI-08 has been developed with “mock” ISCO-08 categories constructed in ISSP 2000-2008 (using conversion and taking into account employment status).
- ESS R5 can / will be double coded (including respondents and partners).
- This is ideal testing ground for ISEI-08.

Steps

- ISCO-88 is converted into ISCO-08x.
- Coder get information about the number of alternative splits.
- Coders revises ISCO-08x into ISCO-08.
- We can evaluate loss / gain of information in a latent-variable model.
- Finished for: CY, DE, DK, IE, NL

ISCO-88 vs ISCO-08: NL

NL	Father	Mother	Resp	Partner	All
Old isko-88	0.971	0.914	0.941	0.957	0.950
Isqo-08 conversion	0.974	0.943	0.971	0.983	0.966
Isqo-08 conversion	0.947	0.972	0.937	0.959	0.953
Isqo-08 revised	0.963	0.967	0.978	0.983	0.970
Old isko-88	0.930	0.891	0.883	0.909	0.906
Isqo-08 revised	0.945	0.913	0.954	0.960	0.943

Cyprus

CY	Father	Mother	Resp	Partner	All
Old isko-88	0.958	0.726	0.922	0.991	0.927
Isqo-08 conversion	0.973	0.916	0.977	1.016	0.972
Isqo-08 conversion	0.907	0.876	0.969	1.001	0.943
Isqo-08 revised	0.988	0.911	0.947	1.008	0.965
Old isko-88	0.894	0.689	0.901	0.961	0.883
Isqo-08 revised	0.973	0.905	0.932	0.989	0.955

Denmark

DK	Father	Mother	Resp	Partner	All
Old isko-88	0.962	0.903	0.916	0.966	0.945
Isqo-08 conversion	1.003	1.012	1.013	1.062	1.009
Isqo-08 conversion	0.98	0.962	0.997	1.057	0.994
Isqo-08 revised	0.97	0.985	0.99	1.032	0.988
Old isko-88	0.947	0.867	0.914	0.968	0.925
Isqo-08 revised	0.974	1.001	1.035	1.035	0.995

Germany

DE	Father	Mother	Resp	Partner	All
Old isko-88	0.967	0.928	0.932	0.936	0.944
Isqo-08 conversion	0.981	0.923	0.976	0.957	0.966
Isqo-08 conversion	0.985	0.984	0.99	0.984	0.987
Isqo-08 revised	0.984	0.991	0.984	0.98	0.994
Old isko-88	0.961	0.923	0.927	0.933	0.939
Isqo-08 revised	0.97	0.923	0.964	0.951	0.957

Ireland

IE	Father	Mother	Resp	Partner	All
Old isko-88	0.919	0.843	0.9	0.903	0.914
Isqo-08 conversion	0.986	0.938	0.989	0.951	0.985
Isqo-08 conversion	0.974	0.95	0.985	1.003	0.975
Isqo-08 revised	0.986	0.964	0.993	0.988	0.986
Old isko-88	0.922	0.842	0.901	0.919	0.903
Isqo-08 revised	0.987	0.938	0.99	0.986	0.98

Conclusions

- ISEI-08 is a bigger improvement (relative to ISEI-88) than ISCO-08 relative to ISCO-88:
  - ISEI-08 converted / ISEI-88 +2%
  - ISEI-08 revised / ISEI-88 +3%
  - ISEI-08 revised / converted +1%
- It is OK to convert ISCO-88 into ISCO-08.
- And apply ISEI-08.

Further steps

- Complete ISCO-88 coding for R1-R5.
- Complete ISCO-08 double coding for R5.
- Produce class categories for ISCO-08 and compare using double coded data.