Coding and Scaling Occupations in ESS R1-R5

Harry B.G. Ganzeboom

University of Lausanne

Lausanne, October 23 2013

Occupations in ESS R1-R5

- Occupations are measured for respondent (current/last, partner (current), father, mother (at resp. age 16).

- R&P occupations are coded into ISCO-88 (ISCO-com) by local agencies (on behalf of national coordinator): ISCOCO, ISCOCOP.

- No further checks on procedure or testing of quality.

Parental Occupations in ESS

- Parental occupations are NOT coded.

- Verbatim strings are published on ESS website.

- For parental (F&M)(occupations there is also a showcard (precoded – crude).

- Total number of strings to be coded: ca. 360.000 in 25 languages.

Policy changes

- Parental showcard was changed to ISSP 1987 format in R4-R5.

- As of R4 ESS requested countries to deposit 'raw data' in NSD archive.

- As of R6, occupations (R&P) are to be coded in ISCO-08.

Goals

- Code all the parental occupations (into ISCO-88).

- Develop status measures [ISEI-08] for ISCO-08.

- Develop tools for assessing the quality of coding and scaling the occupations.

Results and conclusions (1)

- All parental occupations are now coded in ISCO-88. Codes are freely available on www.harryganzeboom.nl.

- These codes are good!

- Coding was done by expert and amateur coders; this does not make a difference.

- On average coding variability leads to 15% loss of information.

- The implementation of the ISSP 1987 showcard was a clear improvement.

- The ISSP 1987 showcard has about the quality of measurement as the detailed ISCO codes.

Results and conclusions (2)

- For ISCO-08 there is now a ISEI-08 available. ISEI-08 can also be applied to ISCO-88. (ISEI-88 can not be applied to ISCO-08!)

- ISCO-88 can be (automatically) converted into ISCO-08 without much loss of information.

- Using coders to check and improve the automated conversion of ISCO-88 into ISCO-08 helps, but only very marginally.

- ISCO-08/ISEI-08 creates somewhat better measurement quality (4%) than ISCO-88/ISEI-88.

Coding project

- Verbatim strings were organized in a 'long' coding file. Coders do not see the association between fathers and mothers!

- Other variables added: self-employment and supervising status. NOT ADDED: education, earnings, gender.

- If possible, strings were matched with previously coded occupation in the same language and codes transferred.

- Coders received a 2 hour training, concentrating on the logic of ISCO-88.

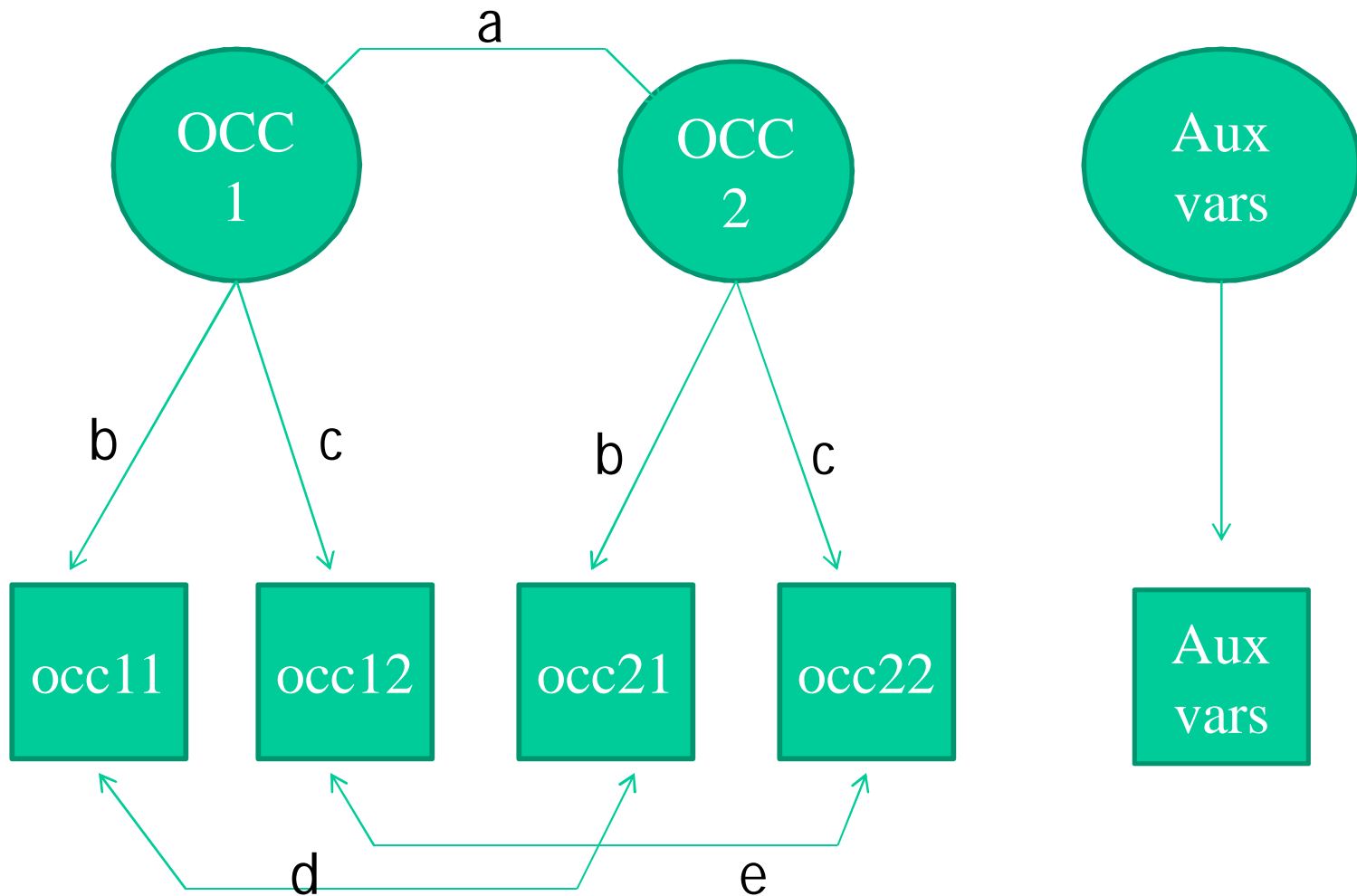- In some instances we had occupations coded by 2 coders.

Quality (1)

- ***How do you examine the quality of coders in languages that you cannot read?***

- The standard answer is:
  - Independent, double coding
    - However, what do you learn from a low inter-coder correlation?
    - What do you do when two coders disagree?
      - How do you choose the best?
      - Or use a third coder (the "adjudicator")?

Quality (2)

- Compare quality between coders by examining association with criterium ('validation') variables.

  – Such as: education, income and … other occupations!

- If you have more than one occupation coded (father & mother), you can apply MTMM modeling.

MTMM

MTMM

- MTMM coefficients:
  - **a**: true correlation
  - **b, c:** Measurement coefficients for reliability of coder 1 and coder 2.
  - **d, e**: Systematic measurement error (??) by coder 1 and coder 2.
- The elementary MTMM model is not identified (despite having 6 correlations with 5 coefficients).
- The model becomes identified by using auxiliary variables (such as education, income).
- You do not need much overlap between the two coders to estimate the model using ML for missing values.

The use of showcards

- In ESS we have also the showcard for both parents: these are independent measures of occupational status.

- The showcard is a rather bad measure of occupational status, in particular in R1-R3 (as I will show).

- However, the bad quality of the showcard measure does not really damage its power as second indicator in the MTMM model.

Showcards

- The showcard in 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.

Results of MTMM parental occs.

- Reliability coders: 0.84
- Reliability showcards: 0.70 (old), 0.86 (new)
- Systematic error coders: 0.13
- Systematic error showcard: 0.32 (old), 0.12 (new)
- Using novice or experienced coders makes no difference to the final quality of measurement.

ISEI-08

- ISEI-08 was developed using converted ISCO-88 codes from ISSP 2002-2008.

- Conversion: straigt cross-walk, but using self-employment and supervising status into account to define supervising occupation (new in ISCO-88).

- ISEI is defined as an optimal scaling of detailed occupations in an elementary status attainment model: occupational status translates education into earnings.

- Due to the construction method, ISEI-08 can also be applied to ISCO-88.

- New: the scaling now refers to men and women. Previously only men.

Coding ISCO-08 in ESS

- Due to the new archiving policy it is now possible to have access to the strings of respondent / partner in ESS.

- For a limited set of countries, all occupation in R5 were coded into ISCO-08 in two steps:

  - Automatic conversion
  - Check by coder.

Results ISCO-08 / ISEI-08

- New ISEI-08 is a clear improvement (about 3%)
- New ISCO-08 is only marginal improvement ($< 1\%$)
- Having a coder check the automated conversion is only marginal improvement ($< 1\%$).

Hindsight

- It would have been much better to ask the coders to translate the occupation files before coding. Google translater has become a big help.

- I do not have a good recommendation on the optimal overlap between coders.

	ESS R1-R3	ESS R4-R5
	Traditionele professionele beroepen *zoals:* accountant – advocaat – medicus – wetenschapsbeoefenaar – civiel/werktuigbouwkundig ingenieur	**Wetenschappelijke en vakspecialistische beroepen** b.v. dokter, leraar, ingenieur, kunstenaar, accountant

ESS R1-R3

Traditionele professionele beroepen
zoals: accountant – advocaat – medicus – wetenschapsbeoefenaar – civiel/werktuigbouwkundig ingenieur

Moderne professionele beroepen
zoals: leerkracht – verple(e)g(st)er – fysiotherapeut – maatschappelijk werker – welzijnswerker – artiest – musicus – politieagent (brigadier of hoger) – software ontwerper

Administratieve en intermediaire beroepen
zoals: secretaresse – persoonlijk assistent – administratief medewerker – kantoorbeambte – medewerker in een call centre – assistent in de verpleging – medewerker in een kinderdagverblijf

Senior manager of bestuurders
(gewoonlijk verantwoordelijk voor planning, organisatie en coördinatie van werk en financiën)
zoals: financieel manager – algemeen directeur

Technische en ambachtelijke beroepen
zoals: monteur – bankwerker – opzichter – loodgieter – drukker – gereedschapmaker – elektricien – boer – tuinman – treinmachinist

Deels routinematige handarbeid en dienstverlenende beroepen
zoals: postbeambte – machinebediener – beveiligingsbeambte – conciërge – landarbeider – assistent in een cateringbedrijf – receptionist – winkelbediende

Routinematige handarbeid en dienstverlenende beroepen
zoals: vrachtwagenchauffeur – bestelwagenchauffeur – schoonmaker – portier – inpakker –machinenaai(st)er – koerier– arbeider – ober/serveerster – barpersoneel

Midden of junior managers
zoals: kantoormanager – winkelmanager – bankmanager – manager landbouwbedrijf – restaurantmanager – magazijnchef – caféhouder

ESS R4-R5

Wetenschappelijke en vakspecialistische beroepen
b.v. dokter, leraar, ingenieur, kunstenaar, accountant

Hoger leidinggevend beroep
b.v. bankier, directeur in groot bedrijf, hogere ambtenaar, vakbondsvertegenwoordiger

Administratieve beroepen
b.v. secretariaatsmedewerker, kantoorbediende, office manager, boekhouder

Commerciële beroepen
b.v. sales manager, winkeleigenaar, winkelbediende, verzekeringsagent

Dienstverlenende beroepen
b.v. eigenaar van een restaurant, politieagent, kelner, verpleegkundige, verzorgende, kapper

Geschoolde arbeider
b.v. ploegbaas, automonteur, drukker, gereedschapsmaker, elektricien

Half geschoolde arbeider
b.v. metselaar, buschauffeur, timmerman, metaalbewerker, bakker

Ongeschoolde arbeider
b.v. handarbeider, fabrieksarbeider

Landbouwberoep
b.v. zelfstandige boer, landarbeider, tractor bestuurder, visser

Table 3: Measutement quality using two models. Model 1: no auxiliary variables, Model 2: with education and family income as auxiliary variables.

Completely standardized estimates (with standard errors)

	Model 1	Model 1a	Model 1b	Model 2	Model 2a	Model 2b
	R1-R5	R1-R3	R4-R5	R1-R5	R1-R3	R4-R5
[b] FOCC → fisco88	0.847 (0.0025)	0.850 (0,0039)	0,845 (0,0031)	0,827 (0,0012)	0,826 (0,0018)	0,829 (0,0016)
[c] FOCC → fcrude	0,776 (0,0023)	0,703 (0,0036)	0,862 (0,0031)	0,794 (0,0012)	0,724 (0,0019)	0,877 (0,0014)
[d] MOCC → misco88	0,853 (0,0031)	0,862 (0,0049)	0,841 (0,0036)	0,816 (0,0016)	0,815 (0,0024)	0,817 (0,0020)
[e] MOCC → mcrude	0,761 (0,0030)	0,688 (0,0042)	0,849 (0,0036)	0,781 (0,0015)	0,707 (0,0025)	0,861 (0,0018)
[a] FOCC ←→ MOCC	0,664 (0,0024)	0,651 (0,0037)	0,673 (0,0031)	0,655 (0,0023)	0,645 (0,0037)	0,662 (0,0031)
[f] fisco88 ←→ misco88	0,091 (0,0096)	0,100 (0,0156)	0,081 (0,011)	0,129 (0,0049)	0,141 (0,0076)	0,123 (0,0062)
[g] fcrude ←→ mcrude	0,278 (0,0048)	0,321 (0,0037)	0,167 (0,011)	0,267 (0,0040)	0,318 (0,0048)	0,124 (0,0079)
L2	10,1	5,4	11,3	493,4	480,7	93,0
Ndf	1	1	1	9	9	9
N	247452	135839	111613	251324	137805	113519

Stata commands:

Model 1: sem (FOCC → zfisei) (FOCC → zfosei) (MOCC → zmisei) (MOCC →zmosei) (OCC→zisei@1) if essround > 0, standardized var(e,zisei@0) method(mlmv) cov(e,zfosei*e,zmosei) cov(e,zfisei*e,zmisei)

Model 2: sem (FOCC → zfisei) (FOCC → zfosei) (MOCC → zmisei) (MOCC → zmosei) (OCC→zisei@1) (EDUC → zeducyr@1) (HINC → zlnhinc) (FED → zfeducyr) (MED → zmeducyr) if essround > 0, standardized method(mlmv) cov(e,zfosei*e,zmosei) cov(e,zfisei*e,zmisei) var(e,zisei@0) var(e,zeducyr@0) var(e,zlnhinc@0) var(e,zfeducyr@0) var(e,zmeducyr@0)

Isei scales detailed occupations (isco-88), osei showcard occupations (crude).

Table 4: Determinants of measurement quality of coded occupation and showcard-occupations.

		CODED OCCUPATIONS				SHOWCARD			
		FATHER		MOTHER		FATHER		MOTHER	
		[b]	Se	[d]	se	[c]	se	[e]	se
AT	Oostenrijk	,791	,054	,801	,060	,708	,111	,687	,110
BE	Belgie	,770	,054	,791	,060	,711	,111	,677	,110
BG	Bulgarije	,897	,039	,900	,043	,782	,080	,788	,079
CH	Zwitserland	,796	,054	,760	,060	,821	,111	,739	,110
CZ	Cyprus	,768	,039	,753	,043	,705	,080	,680	,079
CZ	Tsjechie	,872	,035	,887	,039	,698	,073	,672	,072
DE	Duitsland	,814	,054	,803	,060	,797	,111	,751	,110
DK	Denemarken	,798	,054	,755	,060	,702	,111	,682	,110
EE	Estland	,842	,056	,860	,062	,736	,115	,743	,114
ES	Spanje	,807	,050	,821	,055	,621	,102	,651	,101
FI	Finland	,821	,050	,800	,055	,706	,102	,638	,101
FR	Frankrijk	,746	,034	,687	,037	,687	,069	,594	,069
GB	Groot-Brittanie	,730	,050	,706	,055	,725	,102	,756	,101
GR	Griekenland	,871	,035	,851	,039	,757	,073	,720	,072
HR	Kroatie	,827	,044	,861	,049	,769	,090	,748	,090
HU	Hongarije	,792	,050	,815	,055	,820	,102	,812	,101
IE	Ierland	,775	,054	,764	,060	,713	,111	,703	,110
IL	Israel	,837	,038	,772	,042	,734	,078	,692	,077
IS	Ijsland	,815	,053	,802	,059	,812	,110	,880	,109
IT	Italie	,768	,058	,766	,065	,822	,120	,812	,119
LT	Litouwen	,786	,044	,868	,049	,800	,090	,761	,090
LU	Luxemburg	,783	,058	,740	,065	,839	,120	,749	,119
LV	Letland	,817	,042	,859	,047	,789	,087	,797	,086
NL	Nederland	,802	,054	,769	,060	,545	,111	,525	,110
NO	Noorwegen	,814	,034	,763	,037	,783	,069	,758	,069
PL	Polen	,859	,050	,908	,055	,738	,102	,702	,101
PT	Portugal	,885	,054	,900	,060	,679	,111	,687	,110
RO	Roemenie	,808	,042	,845	,047	,671	,087	,609	,086
RU	Rusland	,810	,039	,850	,043	,802	,080	,763	,079
SE	Zweden	,830	,034	,771	,037	,787	,069	,760	,069
SI	Slovenie	,759	,035	,813	,039	,774	,073	,772	,072
SK	Slowakije	,789	,036	,814	,040	,807	,074	,793	,074
TR	Turkije	,886	,059	,851	,076	,567	,122	,584	,121
UA	Oekraine	,816	,050	,898	,056	,802	,104	,775	,103
ESS-ROUND									
	ESS-R1	0	ref	0	ref	0	ref	0	ref
	ESS-R2	,006	,014	,008	,015	-,064	,028	-,044	,028
	ESS-R3	,006	,014	,016	,016	-,047	,029	-,035	,029
	ESS-R4	-,002	,013	,000	,015	,118	,028	,136	,027
	ESS-R5	-,004	,015	-,006	,017	,122	,031	,134	,030
CODER BEGINNER/PRO									
	Pro	,010	,024	,000	,027	,016	,049	,010	,049

Table 6: Measurement quality of coding in ISCO-88 and ISCO-08. Unstandardized coefficient. Estimation pooled over father, mother, respondent and partner.

	(a)	(b)	(c)	(d)
	Reference	New scaling	Automated conversion	With correction by coder
	ISCO-88 ISEI-88	**ISCO-88 ISEI-08**	**ISCO-08 ISEI-08**	**ISCO-08 ISEI-08**
BE	1	1,028	1,030	1,057
CY	1	1,051	1,045	1,076
DE	1	1,024	1,022	1,022
DK	1	1,009	0,999	,999
HU	1	1,003	1,008	1,003
IE	1	1,021	1,039	1,027
NL	1	1,021	1,004	1,027
PL	1	1,010	1,012	0,992