OCCUPATION CODING: DO'S AND DONT'S

With special reference to the International Standard Classification of Occupations ISCO-88

With an Extension on ISCO-08

Version 4, January 3 2010

Harry B.G. Ganzeboom[1]

Department of Social Research Methodology
Free University Amsterdam
E-mail: HBG.Ganzeboom@fsw.vu.nl

Contents

[1] An earlier version of these notes was presented before the General Assembly of the International Social Survey Programme, Chicago, April 30 2008 and was further developed for a Workshop on Occupation Coding at the Academia Sinica, Taipei, July 2008. I gratefully acknowledge useful discussions with Cinzia Meraviglia, Willem Saris and Tony Tam. **Additions for ESS are marked in RED**.

Synopsis: 20 rules for coding occupations

1. Assemble all the relevant verbatim information in a **coding file**. Beware of Excel's self-completion capacity!

2. If you have more than one occupation to code (current / last / father / mother), assemble all the information **in long format** in a coding file.

3. Include information on **industry** and **status in employment** with the information on occupations.

4. **Do not use information on criterion variables** such as education, actual earnings, gender.

5. Use **multiple coders.**

6. **Instruct and train** the coders, do not correct them. They have to make their own corrections.

7. Make sure that the **coders do not communicate** with each other.

8. Instruct the coders with respect to the **logic of the classification** and general principles, not with respect to specific examples.

9. Make sure that the coders have access to the **full documentation** (ISCO manual with definitions and indexes).

10. Also provide the coders with a **short list** of major and sub-major groups.

11. Have major, sub-major and minor group codes include **trailing zeroes**.

12. Make sure that coders can competently **sort the coding file**. Beware of Excel's capacity to sort separate columns!

13. First code the **first two digits**. There is little need to use the manual at this point (speed!).

14. Split the coding file **in random parts** between coders. Do not give A..M to one and N..Z to another one.

15. Have the coder portions **overlap** by about 10% XX.

16. Make sure that you have **all the information** before you start.

17. It is fine to match the coding file with **a database of already coded occupations** (even if this contains codes in another classification that can be converted into the target classification). However, make sure that these initial codes can be corrected (independently) by the coders.

18. Evaluate the quality of the coders afterwards using **an MTMM model**. Redo the work that has too low quality or use the information to guide you in a final review and adjudication.

19. Keep your work in a separate file and **archive it** for future use to match it with new coding files.

1. CHOOSING THE OCCUPATIONAL CLASSIFICATION

- Purchase an ISCO-88 manual (in English) and have your coders use it. Apart from the useful Introduction (11 pages) and the classification itself (10 pages) it contains over 400 pages of definitions, examples, indexes and guidance that are useful to coders.

- The full ISCO-88 manual is also available online at the ILO website: http://www.ilo.org/public/english/bureau/stat/isco/isco88/

- A frequently used alternative to official ISCO is ISCO(COM), which was created for the European Union (then: Community) by Peter Elias and associates at Warwick University. ISCO(COM) deviates from official ISCO only in minor ways (and there is nothing European about the deviations). However, the guide for users (http://www.warwick.ac.uk/ier/isco/intro.html) contains a lot of useful clarifying notes, some of which are reproduced here in Appendix D.

- **The full ISCO manual as well as the ISCO(com) interpretative notes are available from the ESS website. ISCO(com) is the official coding standard for ESS!**

- If an ISCO translation exists in your language, use it, but be critical about details that may get lost in translation.

- Many countries have indigenous detailed occupational classifications. It is acceptable to use this as a coding frame and transfer the codes later into ISCO. (This is the procedure I use myself and I find it very helpful to review the coding by looking at the ISCO after conversion.)

2. CODING FILE

- Retype all the narrative information from the questionnaire into a database. Recommended string length: A100. Retyping should be as verbatim as possible, but it is helpful to cut some useless information ("My father was a bookkeeper") or have it moved towards the end of the string. Spelling errors can also be corrected. Information that can be alphabetically sorted is vital to the speed of the coding process.

- **Occupational information can be a direct identifier: "Mayor of Amsterdam". Leaving directly identifying information in the coding file should be avoided. Occupational information can also be an indirect identifier: "Mayor" combined with a detailed location code is also an identifier. This is another reason why detailed occupational information should be stored in a separate file with special provisions on use and potential matches with publicly available data.**

- Retyping occupational information into a database is low-level clerical work, but of course errors can be made. If you use Excel, be aware of the dangers of its powerful string-completion facilities, which can lead to many errors. This is the way in my Excel: Tools > Options > Edit > Disable.

- Transfer all the string information into a coding file that has a 'long' format (see Appendix D, i.e. stack the information on respondent's occupation, spouse's, father's, mother's occupation, etc. At a minimum the coding files will contain:
    - CaseID
    - Varname
    - Occupational title (A100).

- Add other information on the occupation, preferably in numerical form, such as:
    - Industry
    - Employment status
    - Supervising status
    - Firm / farm size
    - Required qualifications (if asked in the interview!)

- DO NOT add information to the coding file other things than the occupation. In particular DO NOT add information on:
    - Education
    - Income
    - Age
    - Gender.

  Coders should not be allowed to peek at education to determine an occupation code!!

- If there was double measurement of occupation (e.g. by also asking a crude question, or in longitudinal surveys from earlier waves), DO NOT include this into the coding file. Independence of measurement errors is crucial for modelling and removing it. If you are coding multiple occupations from a panel file, enter these as separate records and keep the unit ID in a separate file.

- Make sure that you have all the occupations before you start coding. Adding in information at a later stage is usually a lot of work.

3. SEMI-AUTOMATIC AND AUTOMATIC CODING

- If you have coded occupations before and kept the results, it is fine to do semi-automatic coding by matching the new information with the old information. I maintain a source of around 50.000 previously coded titles and do exact matching on the first 50 characters of the string and transfer the old codes to the new coding file. This gets me around 40-50% of the new codes, but all of these need to be reviewed since many of these are wrong or imprecise. Appendix C gives a short spss syntax on how to make these matches.

- Increasingly the pro's (i.e. official statistical agencies) use knowledge databases to do in-field coding, i.e. providing interviewers and/or respondents with a potential occupation codes to choose from. Of course, such programs would be helpful to coders too. I have no experience here.

4. SELECTING AND TRAINING CODERS

- Recruitment of coders. Required qualities are: (A) Patience, (B) Interest in the world of work, (C) Willingness to look up information (library and internet skills) and correct initial judgments, (D) Ability to handle (sorting, searching, saving) the coding file, (E) honesty. In addition, it is helpful when coders are able to understand the (English language) ISCO-manual. Students can be a good choice, in particular when you know them personally and when there is chance they will encounter you again.

- Train the coders to understand ISCO-88, in particular its hierarchical digit structure. Looking at a form like Appendix A that shows the structure of ISCO using only the first two digits, is crucial. DO NOT train by showing them difficult examples.

- If you use multiple coders (DO!), it is fine that they have training sessions together. However, DO NOT allow them to consult during coding. It is crucial to have independent information.

5. MULTIPLE CODERS AND DIVIDING UP THE WORK

- It is a good idea to use multiple coders and not only because this makes it possible to speed up the work by division of labour. If coders make errors – and they always do – these will become more like random errors with multiple coders, but they may become systematic error if you employ only one coder. However, be aware that you can turn this advantage into a disadvantage if coders consult each other frequently, and you deprive yourself of the opportunity to check for coder error, be it random or systematic. So DO NOT let them work together.

- Dividing up the work: If you employ multiple coders to divide up the work, give each of them a random part of the coding file. DO NOT give one coder father's and mother's occupation to code and another coder respondent's and spouse's occupation. DO NOT give one coder all occupations starting with A..M and another one N..Z.

- It is a good idea to have the random parts of the coding file partly overlap. The double coded occupations will give information on the quality of the coders.

6. CODING

- Have your coders first code the first two digits (i.e. 10 major and 28 sub-major groups). This information fits on single sheet of A4 paper (see Appendix A) – at this stage DO NOT consult the manual elaborately, which is a major time-saver. This works best if the first two digits are followed by two trailing zeroes: 1100, 1200 etc.

- Frequently sorting the coding file is crucial for quick and accurate coding. So your coders should be able to sort the coding file themselves. Sorting by title helps much to group similar titles. Sorting by code helps much to check the results. [NB: if you use Excel, be aware of the pitfalls of sorting in this program that can sort columns separately!]

- If your coders have done the coding by the first two digits, this might be a natural point to review and discuss their results with them. Make sure you instruct them, DO NOT correct them. Then they can start filling in the details for minor and unit groups (3^{rd} and 4^{th} digit), using the manual. Tell them that it is not a problem changing the first two digits at this stage, if they think this is an improvement. At this stage it is most efficient to work from a file that is sorted by code and within codes by title, and to consult the manual frequently.

- Encourage your coders, in particular as they are confronted with vague and difficult information. Tell them this is not their fault, but yours when you designed the questionnaire and/or instructed the interviewers. Any occupation coding ends up with 10%-20% ambiguous pieces of information for which they have to make a best guess.

- An experienced and skilful coder can process on average around 250 occupations per hour.

7. CHECKING THE QUALITY OF THE CODERS

- Add ISCO-88 value labels to the codes to see whether coders have used illegitimate codes. These labels are available **on the ESS website** and on : http://home.fsw.vu.nl/hbg.ganzeboom/isko88/index.htm .

- Check the quality of the coders by reconstructing the coding file into a person file ('wide format') and calculate correlations between occupations (father, mother, respondent, spouse) after scoring the codes with a status scale (such as ISEI). The coder who obtains the highest correlation is the best, or at least the most consistent one. The one(s) with the lower correlations need(s) further scrutiny.

- ISEI conversion is available at: http://home.fsw.vu.nl/hbg.ganzeboom/isko88/iskoisei.sps.

- If you have employed multiple coders who have done double (=overlapping) coding, a MTMM model (see Figure E2) will allow you to determine (A) coder reliability (coefficients **a** and **b**), (B) coder unique systematic bias (coefficients **e** and **f**), (C) the true score correlation (coefficient **d**), (D) construct-reliability (coefficient **c**). You only need a correlation matrix and some high-school algebra to estimate this model. The coefficients of most interest in this context are **a** and **b**. It is to be expected that **e** and **f** are essentially zero. If they are not, you need to look at systematic differences between coders, which may be restricted to the systematic misclassification of a few numerous occupations (e.g. farmers).

- If you have used double coding and you are not a lisrelite, you need to harmonize the codes into a single one, using an adjudicator (this may be yourself). Make sure that you do not do this until all the independent double coding has been done. You can use the results of the MTMM model (a versus b) to give preference of the results of one coder above the other. Also make sure that you DO NOT change the original codes, but indeed produce a third one. If you are a lisrelite, the whole adjudication procedure is unnecessary and the original two codes are more valuable. DO NOT adjudicate with the initial coders being present, be independent!

8. ARCHIVING

- The coding file should be kept and archived with the original data file, for the following reasons:
  - Can be used as a coding frame at future occasions.
  - Can be used to upgrade or convert information to other or new classifications in the future. ISCO-08 is coming up!
  - It can be analysed by lisrelites. Coding difference may add between 5%-10% attenuation of structural coefficients and this can be corrected by using a multiple indicator model.

Appendix A: ISCO-88 Major and Sub-Major Groups

1000 LEGISLATORS, SENIOR OFFICIALS & MANAGERS

1100 LEGISLATORS & SENIOR OFFICIALS
1200 CORPORATE MANAGERS [LARGE ENTERPRISES]
1300 [SMALL ENTERPRISE] GENERAL MANAGERS

2000 PROFESSIONALS

2100 PHYSICAL, MATHEMATICAL & ENGINEERING SCIENCE PROFESSIONALS
2200 LIFE SCIENCE & HEALTH PROFESSIONALS
2300 TEACHING PROFESSIONALS
2400 OTHER PROFESSIONALS

3000 TECHNICIANS AND ASSOCIATE PROFESSIONALS

3100 PHYSICAL & ENGINEERING SCIENCE ASSOCIATE PROFESSIONALS
3200 LIFE SCIENCE & HEALTH ASSOCIATE PROFESSIONALS
3300 TEACHING ASSOCIATE PROFESSIONALS
3400 OTHER ASSOCIATE PROFESSIONALS

4000 CLERKS

4100 OFFICE CLERKS
4200 CUSTOMER SERVICES CLERKS

5000 SERVICE WORKERS & SHOP & MARKET SALES WORKERS

5100 PERSONAL & PROTECTIVE SERVICES WORKERS
5200 [SALESPERSONS, MODELS & DEMONSTRATORS]

6000 SKILLED AGRICULTURAL & FISHERY WORKERS

6100 MARKET-ORIENTED SKILLED AGRICULTURAL & FISHERY WORKERS
6200 SUBSISTENCE AGRICULTURAL & FISHERY WORKERS

7000 CRAFT ETC TRADES WORKERS

7100 EXTRACTION & BUILDING TRADES WORKERS
7200 METAL, MACHINERY ETC TRADES WORKERS
7300 PRECISION, HANDICRAFT, PRINTING ETC TRADES WORKERS
7400 OTHER CRAFT ETC TRADES WORKERS

8000 PLANT & MACHINE OPERATORS & ASSEMBLERS

8100 STATIONARY-PLANT ETC OPERATORS
8200 MACHINE OPERATORS & ASSEMBLERS
8300 DRIVERS & MOBILE-PLANT OPERATORS
8400 SEMI-SKILLED WORKERS NFS

9000 ELEMENTARY OCCUPATIONS

9100 SALES & SERVICES ELEMENTARY OCCUPATIONS
9200 AGRICULTURAL, FISHERY ETC LABOURERS
9300 LABOURERS IN MINING, CONSTRUCTION, MANUFACTURING & TRANSPORT

Appendix B: The Italian coding file

Variable Information

Variable	Position	Label	Print Format
INTNR	1	Interview number	F8
VARNAME	3	Father / Respondent	A8
TIPO_LAV	4	Job Title	A200
DES_LAV	5	Description	A200
POS_OCC	6	Status in Occupation	A54
SETTORE	7	Industry	A90
ISKO3	8	Fabiana	F8
ISKO4	9	Federica	F8
ISKO34	10	Bestcode	F8

Tabel B1: Part of the Italian coding file, unsorted

VARNAME	TIPO_LAV	DES_LAV	POS_OCC	SETTOE	ISKO3	ISKO4
isko	impresa di pulizie	puliva uffici della dogana	artigiano	altri servizi alle imprese	9132	9132
f1isko	imbiachino			altri servizi al consumatore finale	7141	7141
isko	ristorante	aiutante cuoco	contratto stagionale	ristoranti, bar e cantine	5122	5122
f1isko	costruzioni			costruzioni	1210	7122
isko	clinica privata	inserviente	contratto di altro tipo	altri servizi alle imprese	5132	5132
f1isko	barca a rema			trasporti marittimi e per vie d'acqua	8340	8340
isko	imbarcazione	operaio motorista	lavoro a chiamata	trasporti marittimi e per vie d'acqua	9312	8340
f1isko	costruzioni			costruzioni	1210	7122
isko	costruzione	autista	contratto di manodopera in leasing	costruzioni	1210	8330
f1isko	operaio muratore impresa edile			costruzioni	7129	7122
isko	salariato agricolo	trattorista	contratto di manodopera in leasing	agricoltura, caccia e relativi servizi	9211	8331
f1isko	tornitore meccanico			manutenzione e riparazione autoveicoli e motoveicoli, vendita componenti e accessori	7223	7223
isko	stiratrice	stiro per conto di persone di volta in volta mi contattano	lavoro a chiamata	79	8264	9133
f1isko	contadino			agricoltura, caccia e relativi servizi	6111	6111
isko	collaboratrice domestica	mi occupavo di fare i servizi	lavoro a chiamata	77	9131	9131
f1isko	operaio			costruzioni	7122	9313
isko	parrucchiere	parrucchire	artigiano	artigiano	5141	5141
f1isko	muratore,lavorava con lo zio era una piccola impresa			costruzioni	7122	7122
isko	lavoro come operaio in una azienda che produce gomma plastica,teloni in pvc,tende da sole	operaio	contratto di manodopera in leasing	fabbricazione articoli in gomma e materie plastiche	8232	8232
f1isko	operaio azienda privata			calzaturificio	8266	8266
isko	corrispondenza estera e traduzioni in una casa editrice		contratto di manodopera in leasing	settore grafico editoriale	2444	2444
isko	lavor contabile o durante le fiere gli davo una mano		libero professionista	non ricorda	3433	4190
f1isko	autista di camion			costruzioni	8324	8324
isko	autista di camion	guida camion, pale meccaniche	contratto di manodopera in leasing	costruzioni	8324	8330
f1isko	commerciante di mobili			vendita al dettaglio di altri beni nuovi	1314	1314
isko	bracciante agricola	raccolta olive	contratto stagionale	agricoltura, caccia e relativi servizi	9211	9211
f1isko	magazziniere			vendita al dettaglio di altri beni nuovi	4131	4131

Tabel B2: Part of the Italian coding file, sorted by TIPO_LAV

VARNAME	TIPO_LAV	DES_LAV	POS_OCC	SETTOE	ISKO3	ISKO4
f1isko	autista di camion			costruzioni	8324	8324
isko	autista di camion	guida camion, pale meccaniche	contratto di manodopera in leasing	costruzioni	8324	8330
f1isko	barca a rema			trasporti marittimi e per vie d'acqua	8340	8340
isko	bracciante agricola	raccolta olive	contratto stagionale	agricoltura, caccia e relativi servizi	9211	9211
isko	clinica privata	inserviente	contratto di altro tipo	altri servizi alle imprese	5132	5132
isko	collaboratrice domestica	mi occupavo di fare i servizi	lavoro a chiamata	77	9131	9131
f1isko	commerciante di mobili			vendita al dettaglio di altri beni nuovi	1314	1314
f1isko	contadino			agricoltura, caccia e relativi servizi	6111	6111
isko	corrispondenza estera e traduzioni in una casa editrice		contratto di manodopera in leasing	settore grafico editoriale	2444	2444
isko	costruzione	autista	contratto di manodopera in leasing	costruzioni	1210	8330
f1isko	costruzioni			costruzioni	1210	7122
f1isko	costruzioni			costruzioni	1210	7122
isko	imbarcazione	operaio motorista	lavoro a chiamata	trasporti marittimi e per vie d'acqua	9312	8340
f1isko	imbiachino			altri servizi al consumatore finale	7141	7141
isko	impresa di pulizie	puliva uffici della dogana	artigiano	altri servizi alle imprese	9132	9132
isko	lavor contabile o durante le fiere gli davo una mano		libero professionista	non ricorda	3433	4190
isko	lavoro come operaio in una azienda che produce gomma plastica,teloni in pvc,tende da sole	operaio	contratto di manodopera in leasing	fabbricazione articoli in gomma e materie plastiche	8232	8232
f1isko	magazziniere			vendita al dettaglio di altri beni nuovi	4131	4131
f1isko	muratore,lavorava con lo zio era una piccola impresa			costruzioni	7122	7122
f1isko	operaio			costruzioni	7122	9313
f1isko	operaio muratore impresa edile			costruzioni	7129	7122
f1isko	operaio azienda privata			calzaturificio	8266	8266
isko	parrucchiere	parrucchire	artigiano	artigiano	5141	5141
isko	ristorante	aiutante cuoco	contratto stagionale	ristoranti, bar e cantine	5122	5122
isko	salariato agricolo	trattorista	contratto di manodopera in leasing	agricoltura, caccia e relativi servizi	9211	8331
isko	stiratrice	stiro per conto di persone di volta in volta mi contattano	lavoro a chiamata	79	8264	9133
f1isko	tornitore meccanico			manutenzione e riparazione autoveicoli e motoveicoli, vendita componenti e accessori	7223	7223

Tabel B3: Part of the Italian coding file, sorted by ISKO4

VARNAME	TIPO_LAV	DES_LAV	POS_OCC	SETTOE	ISKO3	ISKO4
f1isko	commerciante di mobili			vendita al dettaglio di altri beni nuovi	1314	1314
isko	corrispondenza estera e traduzioni in una casa editrice		contratto di manodopera in leasing	settore grafico editoriale	2444	2444
f1isko	magazziniere			vendita al dettaglio di altri beni nuovi	4131	4131
isko	lavor contabile o durante le fiere gli davo una mano		libero professionista	non ricorda	3433	4190
isko	ristorante	aiutante cuoco	contratto stagionale	ristoranti, bar e cantine	5122	5122
isko	clinica privata	inserviente	contratto di altro tipo	altri servizi alle imprese	5132	5132
isko	parrucchiere	parrucchire	artigiano	artigiano	5141	5141
f1isko	contadino			agricoltura, caccia e relativi servizi	6111	6111
f1isko	costruzioni			costruzioni	1210	7122
f1isko	costruzioni			costruzioni	1210	7122
f1isko	operaio muratore impresa edile			costruzioni	7129	7122
f1isko	muratore,lavorava con lo zio era una piccola impresa			costruzioni	7122	7122
f1isko	imbiachino			altri servizi al consumatore finale	7141	7141
f1isko	tornitore meccanico			manutenzione e riparazione autoveicoli e motoveicoli, vendita componenti e accessori	7223	7223
isko	lavoro come operaio in una azienda che produce gomma plastica,teloni in pvc,tende da sole	operaio	contratto di manodopera in leasing	fabbricazione articoli in gomma e materie plastiche	8232	8232
f1isko	operaio azienda privata			calzaturificio	8266	8266
f1isko	autista di camion			costruzioni	8324	8324
isko	costruzione	autista	contratto di manodopera in leasing	costruzioni	1210	8330
isko	autista di camion	guida camion, pale meccaniche	contratto di manodopera in leasing	costruzioni	8324	8330
isko	salariato agricolo	trattorista	contratto di manodopera in leasing	agricoltura, caccia e relativi servizi	9211	8331
f1isko	barca a rema			trasporti marittimi e per vie d'acqua	8340	8340
isko	imbarcazione	operaio motorista	lavoro a chiamata	trasporti marittimi e per vie d'acqua	9312	8340
isko	collaboratrice domestica	mi occupavo di fare i servizi	lavoro a chiamata	77	9131	9131
isko	impresa di pulizie	puliva uffici della dogana	artigiano	altri servizi alle imprese	9132	9132
isko	stiratrice	stiro per conto di persone di volta in volta mi contattano	lavoro a chiamata	79	8264	9133
isko	bracciante agricola	raccolta olive	contratto stagionale	agricoltura, caccia e relativi servizi	9211	9211
f1isko	operaio			costruzioni	7122	9313

Appendix C: SPSS syntaxes

From wide to long format.

```
** OCCUPATION STRINGS ARE LOCATED IN FOCC ROCC **.

string title (A100) / varname (A8).

compute title=substr(focc,1,100).
compute varname="FATH".
save outfile="fff.sav"/keep=respnr varname title.

compute title=substr(rocc,1,100).
compute varname="RESP".
save outfile="rrr.sav"/keep=respnr varname title.

add files /file="fff.sav" /file="rrr.sav".

sort cases by title.
```

Matching with source file that has coded occupations

```
** MATCHING WITH SOURCE FILE THAT HAS CODED OCCUPATIONS **.
** VARIABLES OF THE SOURCE FILE ARE: OCCTITLE, ISKO **.

get file="codingfile.sav".

string key (a50).

comp key=substr(occtitle,1,50).
sort cases by key.

add files /file="source.sav" /file=* in=new.

recode new (sysmiss=0).

do if (new eq 0).
comp key=substr(occtitle,1,50).
end if.

sort cases by key new.

do if (sysmis(isko)).
if (lag(key) eq key) isko=lag(isko).
end if.

temp.
select if (new eq 1).
freq isko.

select if (new eq 1).

save outfile="codingfile_coded.sav".
```

From long to wide coding file.

```
get file="codingfile_coded.sav".

if (varname eq "FATH") fisko=isko.
if (varname eq "RESP") risko=isko.

Sort cases by respnr.

temp.
select if (varname eq "FATH").
save outfile="fff_coded.sav" /keep=respnr fisko.

temp.
select if (varname eq "RESP").
save outfile="rrr_coded.sav" /keep=respnr risko.

match files /file="fff_coded.sav" /file="rrr_coded.sav" /by respnr.

define @isko()
  fisko risko
!enddefine.

define @isei()
  fisei risei
!enddefine.

include "iskoisei.sps".

corr fisei risei.
```

Appendix D: SOME NOTORIOUS PROBLEMS IN CODING ISCO-88 AND THEIR SOLUTIONS

Managers, supervisors, proprietors

- Managers: there are two sub-major groups of managers: [1200] Corporate Managers and [1300] General Managers. Corporate Managers work in large – multi-department -- firms that have at least 2 other managers. General Managers work in small firms with at most one other manager. In practice, the distinction will have to be made by title and/or firm size.
- At the minor group level, there are two kinds of department managers: [1220] Production and Operations Department Managers and [1230] Other Department Managers. So, [1220] manage department that do the core business, [1230] head support departments. One will often have to look at industry to make the distinction.
- Working proprietors [small firm] go into general managers.
- Work supervisors [Foreman] go into [1200], if supervising is their only task; they go into [1300] if it is their dominant task, but when they still work along with their subordinates, and they are coded with their subordinates if they are only a 'lead worker'.

Professionals and associate professionals in teaching and nursing

[3230] Associate Professionals in Nursing and [3300] Associate Professionals on Teaching are associate [=assistant] to [2230] Nursing Professionals, and [2330]-[2340] Teaching Professionals. The classification assumes that they are at least two levels in these fields of work and leaves it to the coders to make the distinction. If there is no such differentiation in your country, choose the highest (professional) alternative.

Note that primary and pre-primary teachers always have to be coded as Professionals, even if you feel that they are not at the same skill level as, say, university professors.

Farmers and farm workers

Farmers and farm workers can be coded in 5 places in ISCO-88
[1211] Production Department Manager in Agriculture
[1311] General Managers in Agriculture
[6100] Market Oriented Skilled Agricultural Workers
[6200] Subsistence Agricultural Workers
[9200] Agricultural Labourers

[1211] would be rare – it can only occur in very large farms that have other (support) departments. The choice between [6100] and [1311] is hard to make, and the ISCO manual does not help very much. Most often self-employed farmers are coded as [1311] and [6100] is left empty. However, I think that users would be better helped with coding small farmers in [6100], as [1311] is a four-digit code and easily vanishes in processing.

Craft workers and machinery workers

At the major group level, ISCO distinguishes between [7000] Craft-Trades Workers and [8000] Machine Operators. While related, the distinction is not the same as between skilled

workers and semi-skilled workers. Coding problems arise because many related occupations have parallels in the two major groups, e.g.:

[7110] and [8110]	Miners
[7340] and [8250]	Printers
[7410] and [8270]	Food Producers
[7420] and [8140]	Wood Treaters
[7430] and [8260]	Textile Producers

Mostly, interview materials do not give very much of a clue whether someone is/was a machine worker or not. If no further evidence is available, I advice to use the [7000] variant [skill-level rule, see below].

Crude descriptions

Very often respondents will give only crude descriptions of occupations, which is not saying that these descriptions are unclear. The major solution is to code these descriptions on the 1- or 2-digit level (and use trailing zeroes). Some examples:

Managers	[1200]
Shop Owner	[1300]
Independent	[1300]
Entrepreneur	[1210]
Foreman	[1319]
Skilled Worker	[7000]
Semi-Skilled Worker	[8000]
Unskilled Worker	[9000]

In practice, this works quite well, in particular when one is primarily interested in occupational status.

Ambiguous and multiple descriptions

Very often respondents give information that can be interpreted in multiple ways, because a job has multiple components, or someone has multiple jobs. The Introduction to the ISCO provides a set of rules that can be adapted to this situation. Code the information using the following rules in sequence:

- ***Numerical dominance rule:*** when one activity dominates, or one interpretation is the far more plausible one given the distribution of activities in the population, code accordingly,
- ***Skill level rule:*** if a set of activities involves a mix of skill levels, choose the most skilled one.
- ***Production rule:*** if a set of activities involves production next to sales and/or management, choose the production occupation.

A list of important minor and unit groups distinctions

Not all distinctions between minor (3-digit) and unit (4-digit) groups are equally relevant in a sociological sense, as (sub-)major groups are fairly homogeneous in status. However, there

are important exceptions to this rule, and these are worth special checking when reviewing the codes:

[1221] [1311]	Farmers are very different from other Managers
[2220] [2230]	Doctors versus Nurses
[2310] – [2390]	Various levels of Teachers

However, the shortness of this list should not prevent you from using the full four-digit code. Adding in the 3rd and 4th digit is actually not so much work, once you have done the first two.

Appendix E: MTMM models for multiple coders

If you employ multiple coders (to divide up and speed up the coding process) you can check the quality of their work by having them overlap for a piece of for the entire coding file. You can learn about the quality of their work not only by comparing their choice of the overlapping occupations, but if you have multiple occupations to code (e.g. respondent and spouse), you can also compare them using the between-occupations correlation, provided that you have divided up the coding file in random parts. I recommend creating overlap next to randomised separation of the coding file. Independent of the size of the coding file, the overlap should be at least 200-400 (XX?) randomly chosen occupations to get to sufficiently small confidence intervals.

I present two examples where 2 coders were used for father-respondent data, one for Italy[2] (2 coders coded 1800 pairs, and one for Taiwan, where two groups[3] of coders coded 375 pairs. In both cases there was complete overlap of the occupations between the coders. After transferring the coding file into a wide format and scoring the ISCO codes by the ISEI metric, we obtain the following correlation matrices.

Figure E1: Between-coder correlations in two coders – two occupations situation, two countries

	Italy, 2005, N=1800 occupations				Taiwan, 2007, N=375 occupations			
	Fisei1	Fisei2	Risei1	Risei2	Fisei1	Fisei2	Risei1	Risei2
Fisei1	1.000				1.000			
Fisei2	0.772	1.000			0.865	1.000		
Risei1	0.352	0.332	1.000		0.389	0.369	1.000	
Risei2	0.321	0.322	0.811	1.000	0.448	0.442	0.821	1.000

First we look at the correlations that directly measure the between-coder agreement. These correlations should be high and they are. On detailed inspection, we can make the following further observations:

- Taiwanese coders are on average a bit better than the Italians; or: Taiwanese occupations are a bit easier to code than Italian.
- The Italians are slightly more in agreement about respondents than about fathers, but for the Taiwanese this is the other way around. In my experience the Taiwanese pattern is the more common one.
- No one is perfect, coders introduce considerable error in the measurement process. Otherwise, there correlations would be 1.00. Note that the average between-coder correlation for the same occupation (say, 0.81) suggests that for each statistical relationship with occupation, 10% gets lost ('attenuated') by coding error alone!

[2] The data were kindly provided by Cinzia Meraviglia. The data were collected in 2005 by Luca Ricolfi and the coders were Fabiana and Federica.

[3] One group consisted of the local fieldwork supervisors who initially coded the data in-field, the other group were 49 attendants to a Workshop in Occupation Coding in Taipei, each of whom coded 25 occupations. I averaged the code of male and female coders.

Note that we do not learn anything yet about the relative quality of the two coders at this stage. However, we can learn about that by inspecting the four intergenerational correlations. These correlations are produced by three components:

1. The true intergenerational correlation between fathers and respondents. This does not differ between coders, and is not important for evaluating the differences between them, but it is often what we are most interested in.
2. Random error brought in by coders. If one coder brings in more random error than the other, this will decrease both his own intergenerational correlation, and the intergenerational correlation with the other coder.
3. Systematic (correlated) error, i.e. ways in which coders systematically (i.e. identical for fathers and respondents) misclassify occupations. This may increase the within-coder intergenerational correlation, but not the between-coder one (unless they have brought in the same biases, which is indistinguishable from the true intergenerational correlation).

If there is only random error and this is the same for all coders, the four intergenerational correlations will be identical (but they may all be attenuated = downwardly biased). This is pretty much the case in both matrices, so there are only minor differences between coders in random error and their systematic differences must be small.

A formal structural equation model can be used to estimate the parameters precisely (see Figure E2) in LISREL, AMOS or similar software. However, with some high school algebra you get pretty good results using only paper and pencil, in particular if you assume that there is negligible systematic bias (effects **e** and **f**). It involves solving the following four equations:

$$\mathbf{c} = \sqrt{(r12/r34)}$$

$$\mathbf{a}*\mathbf{b} = r12*\mathbf{c} = r34*1$$

$$\mathbf{a}/\mathbf{b} = \sqrt{(r13/r24)}$$

$$\mathbf{d} = r13/\mathbf{a}*\mathbf{a} = r24/\mathbf{b}*\mathbf{b}$$

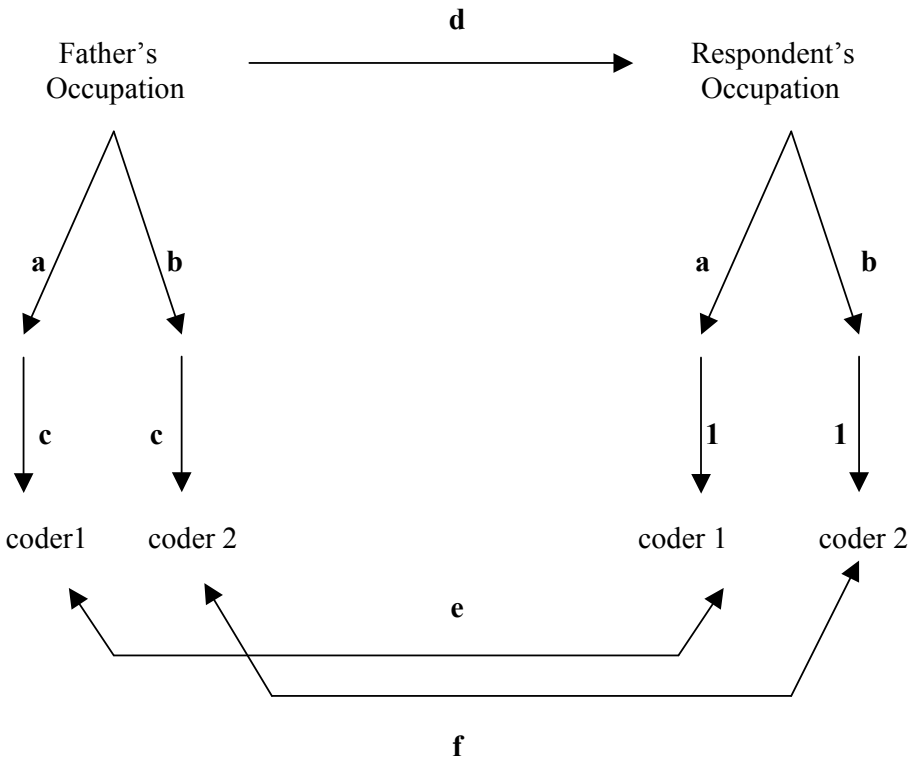
As the algebraic system is overdetermined (4 unknowns with 6 independent correlations), the result of the last step may be inconsistent. This would suggest that effect **e** and **f** matter. Unfortunately, in this system one cannot independently identify **a/b** relative to **e/f**. If **e/f** is of great concern to you, you should bring in more variables (e.g. father's and/or respondent's education) and you can identify these ratio's.

Using Lisrel, we obtain the following estimates

	Italy	Taiwan	
a	0.925	0.902	coder 1 reliability
b	0.877	0.954	coder 2 reliability
c	0.975	0.982	construct reliability
d	0.419	0.492	true score correlation

If you have more coders, or – even better – more occupations, it is easier to identify coding error, both random and systematic. If tested on these two examples, the differences between coders are not statistically significant.

Figure E2: A simple MTMM model to identify random and systematic coding error.



If r12/r34 > 1, interchange **c** and **1**.

EXTENSION ON ISCO-08

QUESTIONS AND ANSWERS ABOUT ISCO-08

This document contains questions and answers on the new International Standard Classification of Occupations 2008 [ISCO-08] and its relationship to previous versions of ISCO and derived occupational status measures such as ISEI, SIOPS, EGP and ESEC.

Last revision: January 3 2010 [HG]

What is ISCO-08 and what can you do with it?

ISCO-08 stands for the International Standard Classification of Occupations 2008 and is published and maintained by the International Labour Organization on behalf of the International Conference of Labour Statisticians, who decided on the final form of ISCO-08 in its resolution of December 6 2007. ISCO-08 in its official form is documented on the ILO/ISCO website (http://www.ilo.org/public/english/bureau/stat/isco/index.htm), that also documents the previous versions: ISCO-58, ISCO-68 and ISCO-88.

ISCO is an international classification aimed for producing comparative official statistics and social research. Many countries maintain national occupational classifications that are often more detailed and may have quite different logics of classifications than ISCO. Such national classifications need to be converted to ISCO for comparative purposes (which will imply some loss of distortion of the information). Increasingly in the past, countries have chosen to adopt ISCO or a somewhat refined version as its national classification. International social research projects, such as the ISSP, ESS, EVS, PISA and IALS have chosen ISCO as their preferred occupational classification.

Note that ISCO is used to classify occupations and does not constitute a sociologically meaningful scale as such. In statistical analysis, the ISCO classification is used as an intermediary step to derive various measures of occupational status, such as ISEI, SIOPS or EGP/ESEC.

While ISCO-08 was adopted by ILO in December 2007, it has hardly been used as a classification tool by projects conducted in 2008 or 2009. However, in the near future, international research projects such as PISA and PIAAC, and most likely the 2010 round of national population censuses will use ISCO-08 as their occupational classification tool.

What is on the ILO website?

The ILO website shows the official version of ISCO-08 using one-line titles of the 570, occupations, 120 minor groups, 34 sub-major groups and 10 major groups. There is also:

- A 583-page document (dated July 9 2009) with *draft* definitions of all the groups. This document also lists examples of occupations to be coded or not to be coded into a certain category.
- Two correspondence tables between ISCO-88 and ISCO-08, one showing how ISCO-88 maps into ISCO-08, and another how ISCO-08 can be mapped back into ISCO-88. Note that these documents imply a many to many mapping of the two classifications.
- Finally, there is a set of discussion documents that comment on definitions by industry or main tasks. These documents were/are used to develop the official definitions.

Note that there appears to be no such thing as an official ISCO-08 manual as of yet. There are many details in the documents at the ILO site that make clear that the final manual is still under construction and that some details have not been settled yet.

The ILO/ISCO website also contains documentation on the three previous versions of ISCO. While there is a pdf version of the complete manuals for 1958 and 1968, ISCO-88 is only documented in an abbreviated form using the definition parts of the manual.

Note that the following document lists a large number of occupations (English language) with their ISCO-68 and ISCO-88 codes:
http://www.ilo.org/public/english/bureau/stat/isco/isco88/alpha.htm.

How do ISCO-88 and ISCO-08 relate?

The official aim of ISCO-08 (stated in the preamble to the ILO-resolution) was to produce ISCO-08 as a "minor revision" of the previous version. One reading of that statement could be that the revision basically amounts to how unit and minor groups are organized, as most of the changes indeed occur at these 4- and 3-digit levels. At the major and sub-major level, fewer modifications have been made and they are mostly in the form of further refinement. However, some of the changes at the more detailed level involve that minor groups are shifted between sub-major groups and occasionally between major groups: despite nominal equivalence at the major and sub-major level, this implies changes in underlying contents.

But however minor or major the changes have been, adopting the new classification involves that researchers familiarize themselves with the new classification. It is most important to understand the hierarchical digit system. For any user, it is best to learn the major and sub-major groups system by heart.

For users who are familiar with ISCO-88 it may be helpful to have a list of the most important changes at the 1-digit (major) and 2-digit (sub-major) level.

Major groups (1-digit level):

- There are small changes in the names of major groups.
- 0000 (Military) is now formally added as a major group.
- In the underlying contents the most dramatic change involves that some occupations that were formerly classified in 8000, are now classified in 3000. These involve Manual Supervisors and part of Stationary Plant Operators.
- Another major change is that two large groups that used to be part of major group 1000, are now classified elsewhere: Shopkeepers go into 5200 and Farmers into 6100.

Minor groups (2-digit level):

- There are now 34 sub-major groups instead of 28. The major changes are:
  - Three new sub-major groups were introduced inside 0000 to distinguish three levels of military ranks.
  - [1200], [1300], [1400)] Managers have been fundamentally revised, involving reordering between sub-major groups and demotion of shop-keepers to 5000.
  - [9400] Food Preparation Assistants have been added a sub-major group. They have no well-defined counterpart in ISCO-88.
  - [3500] Information and Communications Technicians have been upgraded to the sub-major groups level. This also happened to at least 15 other sub-major groups, that are somehow upgraded or downgraded, split of combined.
- In other words: the changes in sub-major groups is certainly not minor. For this reason recoding 2-digits ISCO-88 into 2-digit ISCO-08, without paying attention to the underlying 3- and 4-digit contents is not a good idea.

Two other general observations are important to keep in mind:

- Despite its stated *raison-d'etre* that ISCO-08 adapts to new divisions of labor as they arise in modernizing societies around the world, many changes relative to ISCO-88 are better interpreted as attempts to repair faults of the previous classification and reintroduce elements that were abolished in the transition from ISCO-68 to ISCO-88. Instances are: (A) reintroduction of manual supervisors, (B) separation of shopkeepers from other managers and working proprietors, (C) reintroduction of occupations such as bicycle repairman, kitchen helper, service station attendant and payroll clerk.
- Over-all ISCO-08 appears to be a bit more detailed than its predecessor.

How does the hierarchical digit system work and how many digits do we need?

ISCO-08 is conveniently organized in a system of four hierarchical digits that can be displayed as follows:

[1000] Major group 1
    [1100] Sub-major group 11
    [1200] Sub-major group 12
        [1210] Minor groups 121
            [1211] Unit group 1211
            [1212] Unit group 1212
etc.

(Note that we adopt the convention to list all groups using trailing zeroes. The official ISCO documentation does not. The particular convenience of this is that string sorting and numerical sorting amount to the same thing.)

The hierarchical digit system implies that one can move up levels by simply truncation of trailing digits (and replacing them by 0). This system was first implemented in developing ISCO-88 and was not in place in ISCO-68 and ISCO-58, which creates much trouble in manipulating these codes in computer programs.

Note that the introduction of a major group 0000 may create confusion with missing values codes. Missings are preferably codes as 9999 or negative (-1).

The hierarchical system is a deliberate attempt to group similar occupations in a single branch of the classification in the same way as a hierarchical cluster analysis would do. How similar different members of a branch are, depends of course much on the purpose of an analysis. One approach to answering this question is to adopt a certain scale for all occupations and analyze how much variance is explained at each level. When we take ISEI as our criterion, the following numbers are found for the database of ISSP 2002-2007:

Unit groups	100%
Minor groups	95.5%
Sub-major groups	89.1%
Major groups	82.8%

This result would suggest that any association with occupation is attenuated with the square root of the numbers listed. It needs to be emphasized that these numbers could be smaller or larger when another criterion (such as earnings) would be used.

Still using the ISEI scoring, we can more qualitatively list instances where most heterogeneity would be incurred in only 2 digits were used:

- 1100: Legislators and Corporate Managers
- 2200: Doctors and Nurses
- 2300: University Professors and (Pre-)primary Teachers
- 2600: Lawyers, Librarians and Sociologists
- 3100: Engineering Technicians and Manufacturing Supervisors, Aircraft Pilots
- 3300: Government Officials and Sales Representatives
- 3400: Chefs (Cooks) and Professional Sportsmen
- 5100: Cooks, Waiters and Hairdressers
- 5200: Shop Salesmen and Market Vendors
- 7300: Handicraft Workers and Printers
- 7500: Food Production Workers and Textiles Workers
- 8300: Train drivers, Truck drivers, Sailors
- 9100: Office cleaners and Window Washers
- 0000: Army Officers and Other Army Ranks

Idem when we use only 3 digits:

- 1110: Legislators and Union Officials
- 1210: Finance Managers and Human Resource Managers
- 1220: Sales Managers and R&D Managers
- 1340: Child Care Services Managers and Financial Services Managers
- 2160: Building Architects and Graphic Designers
- 2210: General Medical Practitioners and Specialist Medical Practitioners
- 2220: Nurses and Midwives

- 2260: Dentists and Dieticians
- 2340: Primary and Preprimary Teachers
- 2510: System Analysts and Web Designers
- 2610: Judges and Other Legal Professionals
- 2630: Economists and Religious Professionals
- 2650: Dancers and Sculptors
- 3150: Ship Captains and Aircraft Pilots
- 3250: Dental Assistants and Ambulance Workers
- 3320: Sales Representatives and Buyers
- 3330: Conference Planners and Real Estate Agents
- 3430: Photographers and Chefs
- 4210: Bank Teller and Bookmaker
- 4220: Hotel Receptionists and Survey Research Interviewers
- 5110: Train Conductors and Travel Guides
- 5160: Astrologers, Undertakers and Driving Instructors
- 5220: Shop Keepers and Shop Sales Assistants
- 5240: Fashion Models and Door-to-door Salesmen
- 7110: Bricklayers and Carpenters
- 7120: Roofers and Plumbers
- 7210: Jewellery Workers and Basket Weavers
- 7510: Bakers and Butchers
- 7530: Tailors and Shoemakers
- 8110: Train Drivers and Taxi Driver
- 9333: Freight Handler and Shelf Filler

At the same time it is true that many of the finer distinctions made will be of no practical importance to whatever analysis is done with the occupations.

Many researchers will be tempted to settle for the use of only 2 of 3 digit codes instead of the full version. In doing so they are indeed likely to capture most of the variations in social outcomes brought about by occupations. At the same time, they will bereave (XX) themselves and other users of the data of some extra variance explained, opportunities to compare different ways to scale occupation data and flexibility in uses of the data.

One other element should be brought to the considerations. The novel user of detailed occupation codes may have the impression that coding only 2 digits is a relatively easy task, whereas adding the last 1 of 2 digits is an enormous burden upon the coding process. This is not so, or should not be so. Whereas the coding of major and sub-major groups can (and should) be done initially without consulting the documentation extensively, in the end it is detailed occupations that need to be coded and the choices should be justified using all the available documentation. If one operates coding as a two-step process, in which in a first round the data are classified in a crude way, and in a second round refinement and revision are added, coding the third and fourth digits is not a tremendous amount of work and will in fact add much to the proper coding of the first and second digit.

How is ISCO related to skill level?

Like its predecessor ISCO-08's primary logic of classification is the skill levels of occupations. However, this point of view is not consistently maintained, and this is a good thing. Major groups 2000, 3000, 4000, 7000, 8000 and 9000 are in clear order by skill levels,

but the skill levels of major groups 1000, 5000 and 6000 is ambiguous. There is no connection between skill levels and the ordering of sub-major groups within major groups.

How has the classification of managers changed from ISCO-88 and ISCO-08?

ISCO-88 presented a rather clear definition of how to classify managers and seemed to resolve long-standing issues with this in ISCO-68. First, ISCO-88 distinguished organizations in large and small by using the idea that large organizations have departments and small organizations have not. This distinction was operationalized in a footnote to the classification as small organization having only 1 or 2 managers, while large organizations have at least 2 department managers and at least one 'corporate' manager. The managers of small organizations were called 'general manager', a rather confusing term. The department managers in large organization were divided into two kinds: those who oversaw the production or operation of the organization itself (mending the 'core business') and those were in charge of 'support' departments, such as finance, sales, research and development, etc. The operations department managers were then subdivided according to industry, which made that the coder was faced with the choices like between "Operations Department Manager in Transport" and "Transport Department Manager". While these distinctions are actually conceptually clear, this system was in practice hard to implement as no data source contains the number of managers to begin with, and needed an industry classification to be coded properly. Another problem was that the management categories so defined were rather heterogeneous in occupational status.

ISCO-08 has reordered the information in a dramatically different way:

- Corporate managers ('CEO') are no longer classified with department managers, but with leading government officials.
- Managers are now distinguished primarily by level of activity and industry:
  [1200] Administrative and commercial managers
  [1300] Production and specialized services managers
  [1400] Hospitality, retail and other services managers.
- Shop Keepers (and Shop Supervisors) who used to be included among managers, are now moved to sales workers (5200).
- Farm Managers (& Proprietors) have now been relegated to Agricultural Workers (6000), unless they run very large agricultural enterprises.
- Additional categories have been created elsewhere for some occupations that used to be classified as Managers:
  [0110] and [0210] contain Army Officers and Non-Commissioned Officers,
  [3120] contains three kinds of Mining, Manufacturing and Construction Supervisors,
  [3341] contains Office Supervisors.
  Notice that these additions all arise at the 3- of 4-digit level.

It remains to be seen whether these changes amount to an improvement in practice.

Has the classification of farmers changed between ISCO-88 and ISCO-88?

In ISCO-88 farmers could be coded in five different places in the classification:

[1211] Production Department Manager in Agriculture
[1311] General Managers in Agriculture
[6100] Skilled Farm Workers

[6200] Subsistence Farmers
[9200] [Unskilled] Agricultural Labourer

In ISCO-88 it was particularly unclear whether self-employed farmers, running a small farm were to be coded as [1311] or [6100]. Different coders, different researchers and different countries made different decisions on this. Unfortunately, this can be a very consequential choice, as farmers can be an extremely numerous group, in particular among parents and in developing economies, and are very different in socio-economic status (and social mobility) from other 'General Managers'. Note also that [1311] is a 4-digit code and disappears as a unit when one uses only two or three digits. For this reason it was recommended to avoid [1311] altogether.

In ISCO-08 [1211] has been maintained as [1310], but unfortunately the reference to 'department' has been removed. However, the ILO draft definitions explicitly refer to "large scale agricultural, horticultural and forestry operations such as plantations, large ranches, collective farms and agricultural co-operatives", and so this code would be inappropriate for the average farmer.

This leaves only [6100] and its minor and unit groups as an option to code farmers.

Can ISCO-08 accommodate crude answers to occupation questions, as often arise in survey responses?

Crude responses often occur in surveys and occupational classifications should have facilities to accommodate them (crude responses are not necessarily bad of uninformative responses). Unfortunately, ISCO-08 is not very good at this. However, there are some possibilities. These arise because it is not necessary to code every response to the same levels of detail and a higher-level code may be quite appropriate and preserve the information. Here are some examples:

Managers	[1000] or [1300]
Shop Owner	[5221]
Independent	[1400]
Entrepreneur	[1100]
Foreman	[3120]
Skilled Worker	[7000]
Semi-Skilled Worker	[8000]
Unskilled Worker	[9000]

How do I handle ambiguous and multiple descriptions?

Very often respondents give information that can be interpreted in multiple ways, because a job has multiple components, or someone has multiple jobs. The Introduction to the ISCO-88 manual provides a set of rules that can be adapted to this situation. Code the information using the following rules in sequence:

- ***Numerical dominance rule:*** when one activity dominates, or one interpretation is the far more plausible one given the distribution of activities in the population, code accordingly,
- ***Skill level rule:*** if a set of activities involves a mix of skill levels, choose the most skilled one.

- ***Production rule:*** if a set of activities involves production next to sales and/or management, choose the production occupation.

Do I need / can I use additional information (on industry or status in employment) to code data in ISCO-08?

Strictly speaking coding ISCO-08 requires only a description of the tasks and responsibilities to code the occupation, next to the title of the occupation. However, other information may be available and can be useful to determine the most appropriate code. This information may include:

- Industry.
- Self-employment
- Supervisory status.

It is suggested that you provide your coders with this information to expedite their work.

- Industry is not only useful because ISCO often refers to industrial distinctions ('health', 'government') but also because respondents often mix up occupation and industry.
- Self-employment is useful to delineate shop-keepers from shop-supervisors.
- Supervisory status is useful to categorize responses into one or the other managerial or supervisory codes.

However, it is suggested that you do not apply this information mechanically, but judge its application on a case-by-case basis.

It is also recommended that you do not refer to education, earnings, gender or age while coding, as the relationships between these variables and occupational status are often subject to research and they should not be mixed up during coding.

How can I upgrade ISCO-88 data to ISCO-08 data?

It is not possible to upgrade ISCO-88 data to ISCO-08 data without loss or distortion of information. This would only be possible if ISCO-08 at some level would be a many-to-one mapping of ISCO-88, but it is not. Given the greater detail of ISCO-08, the reverse strategy (downgrading ISCO-08 into ISCO-88) might be the more commendable strategy for researchers who need to harmonize data that are coded in the two classifications, but that remains to be seen.

If one is willing to accept the potential loss and distortions, a straight one-to-one recode is available at isco8808.sps. This (spss) recode chooses the 'best' possible alternative among options. It also indicates whether there were more options to choose from and how many. The one-to-one recodes were established using the many-to-many conversions that ILO provides on its website.

It is sometimes possible to improve the conversion by taking other data into account. In particular it is sometimes helpful to have additional data on industry, or status in employment to choose a more appropriate ISCO-08 category.

The spss conversion tool will work at all levels of detail of the originating ISCO-88 distribution (if coded using trailing zeroes), but note that the quality of conversion remains detail-sensitive. More detail helps to make it better. Converting ISCO-88 into ISCO-08 using

only 2-digit codes on both sides is a different thing as converting at a 4-digit level and then aggregating to two digits!!

Is it useful to use ISCO-88 as a bridge to obtain ISCO-08 codes?

If you have new data to code, nothing is gained by first coding into ISCO-88 and then upgrade to ISCO-08. If your aim would be to obtain a double coded dataset, it would be more useful to work the other way around: first code ISCO-08 and then downgrade to ISCO-88.

A different situation arises when you have already coded ISCO-88 but have also additional data (e.g. the original strings or a national classification) available. In these cases it is useful to use the one-to-one recode as a starter and then use the number of options suggested by the recoding module and the many-to-many correspondence tables provided by ILO to decide what the best fitting option is.

A similar situation arises when you use an existing database coded in ISCO-88 to 'automatically' code occupations.

Suppose I would have these ISCO-08 codes, what can I do with them?

At the moment: very little. Most social research involving ISCO codes works with derived occupational status scales, such as ISEI, SIOPS or EGP/ESEC. These scales are not yet available in final format for ISCO-08. The construction and validation of such occupational status scales awaits the availability of large datasets that are coded in detailed ISCO-08.

To fill the gap, at least for the time being, Ganzeboom has created to provisional occupational status scales, one for ISEI and one for ISEC (which incorporates EGP and ESEC). Recodes are available on: http://home.fsw.vu.nl/hbg.ganzeboom/isco08.

The construction of these two provisional scales is discussed in a companion document.

ISCO-08 MAJOR AND SUB-MAJOR GROUPS

ISCO	sub-major	TITLE	ISEI-08	ISSP-N
0000		**Armed forces occupations**	53	1207
1000		**Managers**	**62**	**14458**
2000		**Professionals**	**65**	**32064**
3000		**Technicians and associate professionals**	**51**	**25797**
4000		**Clerical support workers**	**41**	**21591**
5000		**Service and sales workers**	**31**	**34316**
6000		**Skilled agricultural, forestry and fishery workers**	**18**	**8245**
7000		**Craft and related trades workers**	**35**	**25931**
8000		**Plant and machine operators, and assemblers**	**32**	**15376**
9000		**Elementary occupations**	**20**	**20003**
	0100	Commissioned armed forces officers	65	79
	0200	Non-commissioned armed forces officers	53	1090
	0300	Armed forces occupations, other ranks	30	38
	1100	Chief executives, senior officials and legislators	69	2361
	1200	Administrative and commercial managers	68	3281
	1300	Production and specialised services managers	60	5031
	1400	Hospitality, retail and other services managers	53	3477
	2100	Science and engineering professionals	69	4871
	2200	Health professionals	66	4030
	2300	Teaching professionals	63	12622
	2400	Business and administration professionals	64	3491
	2500	Information and communications technology professionals	69	1662
	2600	Legal, social and cultural professionals	66	4730
	3100	Science and engineering associate professionals	51	5510
	3200	Health associate professionals	46	4848
	3300	Business and administration associate professionals	53	11919
	3400	Legal, social, cultural and related associate professionals	45	2199
	3500	Information and communications technicians	57	944
	4100	General and keyboard clerks	41	5787
	4200	Customer services clerks	40	3931
	4400	Other clerical support workers	40	4916
	4300	Numerical and material recording clerks	43	5756
	5100	Personal service workers	30	9995
	5200	Sales workers	33	14600
	5300	Personal care workers	26	5989

	5400	Protective services workers	40	3081
	6100	Market-oriented skilled agricultural workers	18	6911
	6200	Market-oriented skilled forestry, fishery and hunting workers	24	812
	6300	Subsistence farmers, fishers, hunters and gatherers	10	332
	7100	Building and related trades workers, excluding electricians	34	7455
	7200	Metal, machinery and related trades workers	38	7001
	7300	Handicraft and printing workers	33	1956
	7400	Electrical and electronic trades workers	43	2848
	7500	Food processing, wood working, garment and other craft and related trades workers	27	5386
	8100	Stationary plant and machine operators	29	6127
	8200	Assemblers	29	1100
	8300	Drivers and mobile plant operators	36	7381
	9100	Cleaners and helpers	17	7865
	9200	Agricultural, forestry and fishery labourers	14	2463
	9300	Labourers in mining, construction, manufacturing and transport	24	5641
	9400	Food preparation assistants	15	
	9500	Street and related sales and service workers	25	547
	9600	Refuse workers and other elementary workers	26	1848