

Codering en schaling van beroepen in de European Social Survey¹

Harry B.G. Ganzeboom en David Nikoloski

We report on the coding of the variables for occupation of fathers and mothers, which are only available as strings in the European Social Survey (ESS) round 1 to round 5. With the cooperation of more than 30 coders these strings were coded in the international standard classification ISCO-88. Through a Multi-Trait Multi-Method (MTMM) measurement model, which also makes use of the ESS showcards for parental occupations, we show that with the coding we achieve a good level of reliability (approximately 0.85). The showcard for occupations also achieves a good level of reliability, especially after the revision in round 4. The quality of the coding is quite homogeneous between the different countries or language regions and does not depend on the degree of expertise of the coder. Finally, we report on the conversion of occupations in ESS round 5 to the new classification ISCO-08 and the corresponding ISEI-08 scale. This new procedure brings a slight improvement (approximately 3%) of quality in measurement.

Inleiding

Hoewel de European Social Survey [ESS] in eerste instantie zijn waarde ontleent aan de periodieke peiling van het opinieklimaat in een zeer groot aantal Europese landen, is het bestand ook uitgedoemd tot een van de grootste en rijkste waarnemingen in Europa van sociaal-demografische gegevens. Een bijzonder waardevol onderdeel hiervan vormt de ondervraging naar opleidingen en beroepen van beide ouders van de respondent. Gecombineerd met gegevens over opleiding en beroep van de respondent zelf, vormen deze een database over intergenerationele mobiliteit die geen gelijke kent, zowel in bereik (heel Europa, ca. 33 landen) en omvang (tezamen circa 250.000 cases in vijf tweejaarlijkse ronden [R1-R5]), als in – potentiële – kwaliteit. De nadruk moet hier gelegd worden op de kwalificatie ‘potentieel’: in de praktijk is het gebruik van deze gegevens namelijk omgeven met een groot aantal problemen, dat de volledige benutting ervan tot op heden onmogelijk maakt. Het belangrijkste probleem is dat de beroepen van ouders in de meeste landen niet gecodeerd zijn, maar slechts als *string* ter beschikking staan. Het gaat om bijna 400.000 beroepsaanduidingen in meer dan 25 verschillende talen. Het eerste doel van dit artikel is verslag te doen hoe deze beroepsaanduidingen binnen het kader van het door NWO gefinancierde project ‘Improving the Measurement of Social Background in the European Social Survey’ gecodeerd zijn in de International Standard Classification of Occupations

ISCO-88. Naast problemen doen zich in de ESS echter ook tamelijk onverwachte en onbedoelde mogelijkheden voor om de kwaliteit van beroepsmetingen te bepalen en te verbeteren. Het tweede doel van het artikel is om een meetmodel te ontwikkelen en toe te passen waarmee de kwaliteit van de beroepsgegevens, zowel gecodeerd als op de toonkaart, kan worden berekend. Het derde doel is instrumenten te ontwikkelen en toe te passen waarmee beroepsgegevens uit de ESS – en andere projecten – kunnen worden geconverteerd naar de nieuwe standaard van internationale beroepenclassificatie, ISCO-08.

Achtergronden

De beroepen van ouders en respondenten worden in de ESS verschillend gemeten. De meting van het beroep van de respondent (en ook van diens partner) is het eenvoudigst. ESS schrijft hier een open vraagstelling voor, waarop de antwoorden onder verantwoordelijkheid van de Nationaal Coördinator [NC] dienen te worden gecodeerd naar de *International Standard Classification of Occupations 1988* [ISCO-88], of meer precies: de EU-versie daarvan, ISCO-88(com). Dit levert viercijferige beroepencodes op die als ISCOCO en ISCO COP in de databestanden zijn opgenomen. Er doet zich op dit punt een opmerkelijke situatie voor. Terwijl in het ESS-project alle stappen van dataverzameling (bijvoorbeeld het formuleren van de vraagstellingen, vertaling naar nationale talen, steekproeftrekking en begeleiding van interviewers) in zeer hoge mate geprotocolleerd zijn en vanaf centraal niveau gemonitord worden, ontbreekt elke vorm van standaardisatie en controle ingeval van de beroepen. NC's zijn niet alleen betrekkelijk vrij in hoe zij de beroepenvragen formuleren, maar over hoe zij van (open) antwoorden tot de ISCO-coderingen komen ontbreekt elke rapportage. Ervaringen van Ganzeboom als participant in het ESS-netwerk als Nederlandse NC in R3-R4 hebben geleerd dat het probleem vaak zelfs nog dieper zit: veel NC's hebben geen benul hoe hun beroepencodes tot stand komen. Zij hebben deze verantwoordelijkheid overgedragen aan hun veldwerkorganisaties en zien nooit de strings die tot de uiteindelijke codes leiden. Zij weten bijvoorbeeld niet hoeveel en welke codeurs hiermee aan het werk zijn geweest, hoe de codeurs getraind zijn en wat de kwaliteit van hun werk is. In ten minste één land (Noorwegen) is de toegang tot de beroepenstrings voor de NC op basis van privacywetgeving zelfs verhinderd.

Tot op heden heeft de ESS weinig nader beleid uitgezet ten aanzien van de verzameling en codering van beroepsgegevens. Toch is er een belangrijke ontwikkeling te melden, die voor het onderstaande zeer relevant is. Vanaf R4 heeft men aan de deelnemende landen gevraagd om niet alleen de finale gegevens in hun gecodeerde vorm te deponeren, maar ook de data te archiveren in hun meest ruwe vorm: data vóórdat variabelen zijn geconstrueerd en met name voordat de beroepen zijn gecodeerd. Deze 'ruwe data' zijn niet direct beschikbaar voor gebruikers, maar worden wel bewaard in het NSD-archief in Bergen (NO), en kunnen op verzoek – en onder strikte voorwaarden – *on site* geanalyseerd worden. De gegevens zijn in meerdere opzichten 'ruw' te noemen. Zo is het moeilijk om de verschillende datasets aan elkaar te koppelen en ontbreken in veel gevallen juist de beroepenstrings van de respondent en diens

partner. Niettemin is voor vijftien landen in R5 deze unieke informatie wel beschikbaar. Wij kennen geen enkele andere database voor vergelijkend onderzoek waarin de oorspronkelijke beroepenstrings, gecombineerd met beroepencodes, op een dergelijke schaal beschikbaar zijn.

De situatie rondom de beroepen van de vader en de moeder van de respondent in de ESS is een geheel andere. Op dit punt is, al voordat in 2002 R1 van start ging, het besluit genomen dat de NC's *niet* tot codering van de gegevens over hoefden te gaan. In plaats daarvan konden zij ook de feitelijke beroepenstrings in het NSD-archief deponeren. In de praktijk hebben de meeste landen van deze optie gebruikgemaakt, met als gevolg dat de gepubliceerde ESS-data een totaal unieke collectie van ongecodeerde beroepen van vaders en moeders bevatten. In tegenstelling tot de strings van respondenten en hun partner zijn deze strings vrij beschikbaar op de ESS data-website. Ook op dit punt bestaat geen enkel ander internationaal dataverzamelingsproject waarin zoiets aanwezig is. De achtergrond van de beslissing om beroepen van ouders niet te coderen schijnt te zijn geweest dat men veronderstelde zo een belangrijke tijdsbesparing te verkrijgen. Gegeven dat men toch de beroepen van de respondent en diens partner moest coderen, lijkt dit een argument van betrekkelijk geringe waarde.

Samen met het besluit om de strings van vader en moeder niet te coderen, en om hiervoor te compenseren, werd besloten om ook naar beroepen van vader en moeder te informeren via een toonkaart (een gesloten vraag). Deze toonkaart (zie Tabel 1, linker kolom) vertoonde in eerste instantie opmerkelijke gebreken. Terwijl veruit het meest voorkomende ouderlijke beroep in Europa 'boer' is en elke elementaire analyse van sociale mobiliteit het inzicht biedt dat intergenerationale overdracht dramatisch verschilt tussen boeren en alle andere beroepen (praktisch iedereen die boer wordt, is van boerenkomst – dit komt in andere beroepen niet op deze schaal voor), staat de in ESS R1-R3 gebruikte toonkaart niet toe dit antwoord te kiezen. Ook andere vertrouwde en zeer belangrijke onderscheidingen in beroepen zijn niet terug te vinden in de toonkaart: er is geen helder onderscheid tussen hand- en hoofdarbeid, of tussen beroepen met hoog en laag aanzien. Verder is de betekenis van gebruikte aanduidingen als 'senior' en 'junior', 'modern' en 'traditioneel' in deze context onduidelijk. Ten slotte is er geen duidelijk rangorde in de categorieën, hetgeen het uitkiezen van het best passende antwoord bemoeilijkt. Hoewel daarmee het vermoeden rijst dat de toonkaart een slechte meting van ouderlijk beroep is, heeft de gehele situatie ertoe geleid dat we in de ESS wat betreft de ouders niet alleen de beschikking hebben over een unieke verzameling beroepenstrings, maar dat we ook nog over een onafhankelijke tweede meting van die beroepen beschikken. In het onderstaande zal berekend worden hoe slecht de toonkaart is, maar ook geïllustreerd worden dat een slechte tweede meting toch heel waardevol kan zijn.

De toonkaartmeting is niet onveranderd gebleven. Na aandringen van Ganzeboom, in zijn rol als NC voor Nederland, en een overtuigende presentatie dat de samenhang tussen de toonkaartgegevens en de beroepen die de respondenten in de open vragen hadden aangegeven over hun ouders wel heel gering was, althans in Nederland, is met ingang van R4 de toonkaart veranderd in een formaat dat sterk geïnspireerd is op een soortgelijke vraag in de ISSP 1987 (zie Ganzeboom (2005) en De Vries & Ganzeboom

Tabel 1 *Meting van beroepen van ouders via gesloten vraagstelling in ESS, R1-3 en R4-5, met daarbij gebruikte ISEI-schalingen*

R1-R3	ISEI#	R4-R5	ISEI#
Traditionele professionele beroepen zoals: accountant – advocaat – medicus – wetenschapsbeoefenaar – civiel / werktuigbouwkundig ingenieur	82	Wetenschappelijke en vakspecialistische beroepen zoals: dokter – leraar – ingenieur – kunstenaar – accountant	84
Moderne professionele beroepen zoals: leerkracht – verple(e)g(st)er – fysiotherapeut – maatschappelijk werker – welzijnswerker – artiest – musicus – politieagent (brigadier of hoger) – software ontwerper	65	Hoger leidinggevend beroep zoals: bankier – directeur in groot bedrijf – hogere ambtenaar – vakbondsvertegenwoordiger	67/69
Administratieve en intermediaire beroepen zoals: secretaresse – persoonlijk assistent – administratief medewerker – kantoorbeambte – medewerker in een call center – assistent in de verpleging – medewerker in een kinderdagverblijf	48	Administratieve beroepen zoals: secretariaatsmedewerker – kantoorbediende – office manager – boekhouder	48
Senior manager of bestuurders (gewoonlijk verantwoordelijk voor planning, organisatie en coördinatie van werk en financiën) zoals: financieel manager – algemeen directeur	67/69	Commerciële beroepen zoals: sales manager – winkeleigenaar – winkelbediende – verzekeringsagent	42/53
Technische en ambachtelijke beroepen zoals: monteur – bankwerker – opzichter – loodgieter – drukker – gereedschapsmaker – elektriciens – boer – tuinman – treinmachinist	43	Dienstverlenende beroepen zoals: eigenaar van een restaurant – politieagent – kelner – verpleegkundige – verzorgende – kapper	41/48
Deels routinematige handarbeid en dienstverlenende beroepen zoals: postbeambte – machinebediener – beveiligingsbeambte – conciërge – landarbeider – assistent in een cateringbedrijf – receptionist – winkelbediende	28	Geschoolde arbeider zoals: ploegbaas – automonteur – drukker – gereedschapsmaker – elektriciens	43/44
Routinematige handarbeid en dienstverlenende beroepen zoals: vrachtwagenchauffeur – bestelwagenchauffeur – schoonmaker – portier – inpakker – machinenaai(st)er – koerier – arbeider – ober/serveerster – barpersoneel	23	Half geschoolde arbeider zoals: metselaar – buschauffeur – timmerman – metaalbewerker – bakker	28
Midden of junior managers zoals: kantoomanager – winkelmanager – bankmanager – manager landbouwbedrijf – restaurantmanager – magazijnchef – caféhouder	67	Ongeschoolde arbeider zoals: handarbeider – fabrieksarbeider	23
		Landbouwberoep zoals: zelfstandige boer – landarbeider – tractorbestuurder – visser	17/26

Het eerstgenoemde cijfer is voor loondienst, het tweede indien men voor eigen rekening werkzaam is.

(2008) en tabel 1, rechter kolom). In het onderstaande zal worden geanalyseerd in hoeverre deze verandering een verbetering is geweest.

Ten slotte dient nog aangegeven te worden dat de ESS, net als veel andere surveyprojecten, voor een belangrijke vernieuwing staat aangaande de codering van beroe-

pen: de invoering van de nieuwe versie van de *International Standard Classification of Occupations* 2008 [ISCO-08]. Deze nieuwe classificatie is in de praktijk sinds 2010 beschikbaar, en de documentatie ervan is in 2012 afgerond met de publicatie van een uitvoerige manual (ILO, 2012). De invoering van deze nieuwe codering is aanstaande in een reeks van internationale onderzoeksprojecten en wordt in ESS ook vanaf R6 geëffectueerd. Een voor de hand liggende vraag is nu welke gevolgen deze verandering zal hebben voor de kwaliteit van de beroepsgegevens: zullen de nieuwe beroepsgegevens beter of slechter zijn dan de oude, en in hoeverre zal er sprake zijn van een trendbreuk?

Het ESS-DEVO project *Improving Social Background Measures in the European Social Survey*

Bovenstaande achtergronden zijn in 2010 voor NWO aanleiding geweest om als onderdeel van de ESS Developmental Projects [ESS-DEVO] financiering te geven aan het onderzoek *Improving Social Background Measures in the European Social Survey*. Wat de hierboven aangegeven situatie van de beroepen² aangaat, heeft dit project als doelstellingen:

- Het coderen van alle in ESS R1-R5 beschikbare beroepenstrings van vaders en moeders in ISCO-88 en de kwaliteit van de coderingen te toetsen via een meetmodel.
- Het ontwikkelen van instrumenten om de overgang van ISCO-88 naar ISCO-08 te faciliteren, in het bijzonder om beroepen gecodeerd in ISCO-88 om te zetten in ISCO-08 en schalen te ontwikkelen die in samenhang met ISCO-08 gebruikt kunnen worden om de status van beroepen te meten.

Het onderzoeksproject, dat nog loopt tot eind 2013, heeft goede vorderingen gemaakt: de bovenstaande doelstellingen zijn inmiddels grotendeels gerealiseerd. In dit artikel wordt over de behaalde resultaten verslag gedaan en worden daarover onderzoeksvragen beantwoord. Eerst wordt uit de doeken gedaan welke procedures zijn gevolgd om de beroepenstrings van vaders en moeder in ISCO-88 in R1-R5 te coderen, en hoe de conversie van ISCO-88 gegevens naar ISCO-08 in R5 is uitgevoerd. Daarna lichten we toe hoe de kwaliteit van deze coderingen onderzocht kan worden met een factor-analytisch meetmodel, dat het mogelijk maakt random en systematische meetfouten te onderscheiden. Met het gebruik van dit model geven we vervolgens antwoord op de volgende vragen:

- Hoe goed is de kwaliteit van de nieuwe ISCO-88 coderingen van de ouderlijke beroepen in ESS R1-R5?
- In hoeverre zijn verschillen in meetkwaliteit van de beroepencoderingen terug te voeren op de gevolgde procedures, in het bijzonder de selectie en training van de codeurs?
- Levert de overgang van ISCO-88 naar ISCO-08 coderingen verlies of toename van meetkwaliteit op?

De ESS beroepenstrings

De beroepenstrings van vader en moeder kunnen eenvoudig vanaf de ESS data-website worden gedownload als SPSS-datafiles. De strings zijn hier per land en per ronde georganiseerd, en verschijnen in lokale scripts: Cyrillisch, Grieks, Hebreeuws, maar ook in romaans alfabet met alle diakritische tekens waaraan bijvoorbeeld het Deens, Noors, Tsjechisch, Frans en Duits zo rijk zijn. Het heeft tamelijk veel moeite gekost om deze gegevens samen te voegen en in toegankelijke vorm aan de beroepencodeurs ter beschikking te stellen. De moeilijkheden komen hoofdzakelijk voort uit de beperkte en eigenaardige manier waarop SPSS met strings omgaat – met name waar het de behandeling van lokale scripts betreft. Met enige moeite zijn niettemin per land de gegevens van alle beschikbare ronden samengevoegd in een zogenaamde codeerfile. Een dergelijk bestand verzamelt alle te coderen (en eventueel reeds gecodeerde) strings in een ‘lang formaat’, waarin elk te coderen beroep een *record* vormt. Codeerfiles verbreken het familieverband (je ziet niet welke vader bij welke moeder hoort) en maken het eenvoudig om de informatie te sorteren en erin te zoeken (zie verder Ganzeboom (2010)).

Alvorens de codeurs aan het werk te zetten, werden de strings alfabetisch geordend en gematcht met bronnen waarin reeds beroepenstrings met ISCO-88 codes voorkwamen. In een aantal gevallen (Nederlands, Duits, Fins, Frans) waren zulke bronnen beschikbaar, maar ze werden ook in de loop van het project opgebouwd. Zo kwamen de strings van R5 pas ter beschikking op het moment dat al een aantal landen afgerond was wat betreft R1-R4: de eerder gecodeerde gegevens werden dan een bron voor de automatische codeerslag. In andere gevallen werd gematcht met andere bronnen, in het bijzonder ook de coderingen van beroepen van respondenten en partners uit de in Bergen (NO) gedeponeerde ruwe data van respondenten en partners. Ook werden gegevens uit hetzelfde taalgebied (DE – AT– CH; BE – NL; BE – FR – CH; IE – GB; CY – GR) wederzijds als bron gebruikt. De ISCO-88 coderingen uit de bronbestanden werden alleen overgedragen wanneer een beroepsstring volledig identiek was met die van een reeds gecodeerd record. Het succes van deze automatische codeerslag hangt sterk af van de kwaliteit van de bron en de aard van de te coderen strings: onze ervaringen leren dat rond de 40% van alle beroepstitels op deze manier een min of meer correcte code toegekend kunnen krijgen. Bij sterk gestandaardiseerde en repetitieve gegevens kan dit echter wel oplopen tot 75%.

De aldus half gecodeerde verzameling strings werd aan de codeurs ter beschikking gesteld, doorgaans als een Excel-bestand. Aan elke record werd numerieke informatie toegevoegd over twee andere aspecten van het uitgeoefende beroep: (a) al dan niet zelfstandig (= in eigen bedrijf) werkzaam zijn, en (b) het aantal personen waarover men leiding geeft. Hoewel ISCO-88 in strikte zin niet gevoelig is voor deze twee kenmerken, helpt het vaak toch te bepalen wat een best passende codering is.

Voor fase twee van het project, het coderen van de beroepen uit R5 in de nieuwe classificatie ISCO-08, werden de strings van beroepen van respondent en partner op locatie uit het NSD data-archief verkregen. In veel landen is naar deze beroepen geïnfomeerd via meerdere vragen, die evenzovele strings opleveren: beroepstitel, taken,

voor het beroep vereiste opleiding, en de bedrijfstak waarin men werkzaam is. Al deze informatie werd tezamen met de ouderlijke beroepen uit R5 opgenomen in de codeerfile waarmee de beroepencodeur aan de slag ging.

ISCO-88: de bestaande beroepenclassificatie en haar afleidingen

ISCO-88 (ILO, 1990) is de tot op heden veelgebruikte internationale standaardclassificatie van beroepen. Anders dan haar voorgangers (ISCO-58 en ISCO-68) is ISCO-88 het standaard meetinstrument voor beroep geworden in belangrijke internationale surveys, zoals ESS, ISSP, SHARE, EVS en PISA. In toenemende mate wordt de classificatie ook als nationale beroepencode gebruikt.

ISCO-88 is opgebouwd als een viercijferig hiërarchisch systeem, bijvoorbeeld als:

```
2000 Professionals
    2100 Science and Engineering Professionals
        2110 Physical and Earth Science Professionals
            2111 Physicists and Astronomers
            2112 Meteorologists
```

De onderscheiden detailleringniveaus staan bekend als *major groups* (9x), *sub-major groups* (28x), *minor groups* (116x) en *unit groups* (390x). De detailleringniveaus kunnen door elkaar gebruikt worden. Tezamen zijn er dus 543 codes om uit te kiezen. Sociologisch gezien zijn de eerste twee cijfers van de code veruit het belangrijkste; de onderscheidingen op de twee meest gedetailleerde niveaus hebben weinig sociologische relevantie. Er zijn echter een paar zeer belangrijke uitzonderingen op deze regel. Zo kan men boeren (1311) alleen maar onderscheiden van andere kleine zelfstandigen door de volle vier cijfers te gebruiken, en heeft men het derde cijfer nodig om verpleegsters (2230) en medici (2220) uit elkaar te halen.

Met beroepencodes alleen kun je in sociologisch onderzoek nog niet zo veel doen: het zijn de afleidingen naar beroepsstatus die in feite in statistische modellen geanalyseerd worden. Beroepsstatus dient hier breed te worden opgevat, omdat er veel verschillende overtuigingen leven over welke afleiding het beste en meest relevant is. Globaal genomen bestaan er drie typen afleidingen:

- Het beroepsprestige: dit is hoe men een beroep evalueert. Een internationale prestigeschaal is voor het eerst opgesteld door Treiman (1977) op basis van ISCO-68, en deze is naderhand door Ganzeboom & Treiman (1996, 2003) gegeneraliseerd naar ISCO-88.
- Sociaal-economische index [SEI] van beroepsstatus: net als bij prestige worden in een SEI-index beroepen naar een continue score geschaald, maar nu op basis van samenhang met criteriumvariabelen, in het bijzonder de opleiding en het inkomen van beroepsbeoefenaren. Een SEI-index kan geïnterpreteerd worden als het aspect van beroepen dat ervoor zorgt dat er inkomensverschillen tussen lager en hoger opgeleiden bestaan. Een internationale SEI-index [ISEI] is voor ISCO-68 ontwikkeld door Ganzeboom e.a. (1992) en vervolgens opnieuw voor ISCO-88 door Ganzeboom & Treiman (1996, 2003).

- Sociale klassen: indelingen van beroep naar discrete categorieën, meestal op basis van een combinatie van beroepstitel met de kenmerken zelfstandigheid en leidinggevendheid. Het meest gebruikte (EGP-)schema is ontwikkeld door Erikson, Goldthorpe & Portocarero (1979), en Ganzeboom e.a. (1992, 1996, 2003) geven aan hoe de afleiding daarvan geconstrueerd kan worden uit ISCO. Recent is door Rose & Harrison (2010) een beperktere indeling voorgesteld (European Socio-Economic Classes: ESEC) die min of meer hetzelfde doet.

ISCO-08: de nieuwe beroepenclassificatie

ISCO-08 is door de International Conference of Labour Statisticians en de International Labour Organisation, die de internationale beroepenclassificatie in beheer hebben, aangekondigd als een *minor revision* van ISCO-88. Achter deze wat bedrieglijke formulering gaan allerlei zaken schuil. Op het eerste gezicht lijkt ISCO-08 identiek aan ISCO-88 wat betreft de hoofdingeling (*major groups*) en stemmen ook zeer veel *sub-major groups* overeen met de eerdere indeling. Nadere bestudering leert dat de nieuwe classificatie gedetailleerder is dan de vorige (ca. 10% meer aanduidingen) en ook dat buiten de major groups nagenoeg alle codegetallen zijn veranderd (Ganzeboom & Treiman, 2013). Ongeveer 50% van alle beroepsgroepen is van inhoud of systematiek veranderd. Veel veranderingen zijn vanuit sociologisch gezichtspunt betrekkelijk marginaal. Onder alle veranderingen zitten echter ook een paar substantiële, die de nieuwe classificatie onvergelijkbaar maken met de vorige. De belangrijkste substantiële veranderingen zijn:

- Op een aantal plaatsen in de indeling is ruimte gegeven aan beroepen die in belangrijke mate of uitsluitend leidinggevend zijn, ook op een lager statusniveau. Voorbeelden zijn de komst van *Production Supervisors*, *Office Supervisors* en *Shop Supervisors*, en het onderscheid tussen officiersrangen bij militaire beroepen. ISCO-08 maakt het weer mogelijk veel voorkomende beroepen als ‘ploegbaas’, ‘voorman’ en ‘uitvoerder’ (en vele andere) eenduidig te coderen. Dat was niet mogelijk in ISCO-88 (wel in ISCO-68).
- De indeling van leidinggevende (*management*) beroepen is ingrijpend veranderd. ISCO-88 maakt alleen onderscheid tussen managers die leiding geven aan een afdeling van een organisatie en managers die leiding geven aan een gehele organisatie, dan wel aan een kleine organisatie zonder afdelingen. In ISCO-08 is management opgedeeld naar de aard van de verantwoordelijkheden: er zijn hogere leidinggevendenden en lagere, en er is een aparte plaats ingebouwd voor management in de horeca. De nieuwe logica is evenmin dwingend als de oude, maar de nieuwe indeling lijkt wel beter aan te sluiten bij de termen waarin ondervraagden over beroepen rapporteren.
- De codering van zelfstandige boeren is eenduidiger geworden. Door de ISCO-08 indeling is specifiek het probleem opgelost dat boeren niet onderscheiden kunnen worden van andere kleine zelfstandigen en leidinggevendenden in het kleinbedrijf wanneer men de codering beperkt tot de eerste twee cijfers.

- De codering van de zelfstandige winkeliers is daarentegen nog ambiguër geworden dan deze al was. Winkeliers zijn nu in een enkele minor group samengevoegd met toezichthoudend personeel in winkels en winkelbedienden, slechts onderscheiden in het vierde codecijfer.

Opmerkenswaardige, maar minder belangrijke, algemene trends in de herziening zijn verder:

- Er zijn talloze nieuwe beroepen op ICT-gebied onderscheiden, zoals bijvoorbeeld *webdesigner* en *ICT teacher*.
- De hoeveelheid onderscheidingen in de ambachtelijke industrie is daarentegen drastisch verminderd.
- Bij de elementaire beroepen (schoonmaak, bewaking, eenvoudige productie-arbeid) zijn er juist meer onderscheidingen ingevoerd.
- In een opmerkelijk aantal gevallen zijn beroepen ingevoegd die in ISCO-88 verwijderd waren ten opzichte van haar voorganger ISCO-68: fietsenmaker, pompbediende, keukenhulp, et cetera. Hoewel het nergens aangekondigd is, is het duidelijk dat men bij de *minor revision* ook fouten in de vorige herziening heeft willen goedmaken, soms tot op het overdrevene af. Zo onderscheidt men nu vier categorieën van *subsistence* (= op zelfvoorziening gerichte) *farmers*, hoewel het idee van specialisatie toch intrinsiek strijdig lijkt te zijn met de aard van dit beroep.

Zoals ook bij ISCO-88 het geval was, hebben de viercijferige beroepencodes van ISCO-08 geen directe betekenis bij analyses. Afgeleide schalen voor prestige (SIOPS-08), sociaal-economische status (ISEI-08) en sociale klasse (ISEC-08) zijn inmiddels geconstrueerd door Ganzeboom & Treiman (2013). Omdat deze afleidingen hebben plaatsgevonden met behulp van gegevens (afkomstig uit de ISSP 2000-2008) die zelf nog niet in ISCO-08 waren gecodeerd, maar waarin substitueert ISCO-08 codes geconstrueerd werden via conversie, is het dubbel coderen van de ESS R5 gegevens in zowel ISCO-88 als ISCO-08 een eerste grootschalige toets voor de kwaliteit van de afgeleide schalen. Tabel 2 geeft een overzicht van de beide ISEI-schalings, waar het de major groups van de beide classificaties aangaat. De verhuizing van de winkeliers van 1000 naar 5000 wordt hierin zichtbaar door de stijging van het gemiddelde ISEI van hoofdgroep 1000.

Tabel 2 Major groups van de International Standard Classification of Occupations, met ISEI-88 en ISEI-08 schaling

		ISEI-88	ISEI-08
1000	Leidinggevende functies	55	64
2000	Intellectuele, wetenschappelijke en artistieke beroepen	70	78
3000	Technici en vakspecialisten	54	57
4000	Administratief personeel	45	43
5000	Dienstverlenend personeel en verkopers	40	30
6000	Geschoolde landbouwers, bosbouwers en vissers	23	20
7000	Ambachtlieden	34	29
8000	Bedieningspersoneel van machines en installaties, assembleurs	31	26
9000	Elementaire beroepen	20	18

De categorietitels zijn ontleend aan de ISCO-08 classificatie (EU-vertaling)

Werving, training en begeleiding van codeurs

Codeurs werden geworven via open sollicitatie en netwerkcontacten. De grootste groep bestond uit buitenlandse studenten in Amsterdam, maar ook een aantal in het buitenland woonachtige studenten en collega-onderzoekers verleenden medewerking aan het codeerwerk. Er werden geen andere ingangseisen gesteld dan ambitie, competentie in de betreffende taal, vaardigheid in het manipuleren van Excel-bestanden en beschikking over een internetverbinding. In het hierna volgende zullen we analyseren of de achtergronden van de codeurs van invloed zijn geweest op de verkregen resultaten. We verdelen de codeurs daartoe in twee groepen:

- *Naïeve codeurs*: dit zijn in doorsnee buitenlandse studenten in Amsterdam die geen enkele sociaal-wetenschappelijke ervaring hadden.
- *Professionele codeurs*: personen die al eerder beroepen gecodeerd hadden of als sociaal-wetenschappelijk onderzoeker werkzaam zijn. Hieronder worden ook de hoofdonderzoeker zelf en de bij het onderzoek betrokken assistenten gerekend.

De groep van naïeve Amsterdamse codeurs werd opgeroepen voor een gezamenlijke training van twee-drie uur, de professionele codeurs verzorgden deze training of werden, voor zover het personen in het buitenland betrof, via Skype en e-mail individueel geïnstrueerd. De training voor de naïeve codeurs had de volgende elementen:

- Een algemene introductie in het project en de aard van de gegevens.
- Een gedetailleerde behandeling van de structuur van ISCO-88, met een stipulering van bekende moeilijke problemen, in het bijzonder: management, boeren, onderscheid tussen professionals en semi-professionals.
- Een instructie in de logistiek van het project, in het bijzonder hoe contact te houden met het projectmanagement, betaling, et cetera.

De codeurs werkten voor een standaardtarief van 12 euro per uur, waarbij we ervan uitgingen dat zij gemiddeld tweehonderd beroepen per uur zouden kunnen beoordelen. Of deze gemiddelde tariefstelling redelijk zou overeenstemmen met de feitelijke voortgang was vooraf moeilijk te voorspellen, omdat dit afhing van twee voor ons onbekende grootheden: de aard van de te coderen strings en de kwaliteiten van de codeur. Veruit het moeilijkste bestand bleek uiteindelijk het Franse te zijn: hierop beten meerdere Franstalige codeurs hun tanden stuk en het werd uiteindelijk afgerond door een Nederlandse student Frans. De codeur zelf bleek een factor van belang: de Poolse strings leken bijzonder ingewikkeld en de aanvankelijk aangestelde codeur maakte maar moeizaam voortgang. Een nieuw aangestelde codeur kon het echter wel binnen het standaard tijdschema. We vroegen de codeurs ook zelf hun tijd bij te houden. In een enkel geval hebben we de beloning wat ruimer bemeten.

Codeurs verrichtten hun werkzaamheden onafhankelijk van elkaar, volgden hun eigen tijdschema en werkten op hun eigen computer. Zij werden verzocht moeilijke punten terug te rapporteren; hiervan werd spaarzaam gebruik gemaakt. Bij inlevering van de gecodeerde gegevens werd een controle uitgevoerd op compleetheid en geldigheid van alle codes. Vervolgens werden de gegevens in een correlatieel model in

verband gebracht met opleidingen van beide ouders en de beroepen en opleidingen van respondenten en partners. Het eerste punt van controle in deze fase was dan om de opleiding-beroepscorrelaties van beide ouders te vergelijken met die van de respondent en diens partner. Deze moesten in dezelfde orde van grootte zijn, anders was er reden tot zorg. Deze controles leidden tot feedback aan de codeur, waarbij een aantal moeilijke punten nog eens werd doorgenomen. In andere gevallen zochten we een tweede codeur voor correctie of een tweede codering.

Bij deze feedbackwerkzaamheden werden we erg ondersteund door de gegroeide mogelijkheden van automatische vertaling: het is tamelijk eenvoudig geworden om via Google Translate alle strings te vertalen. Dit levert weliswaar in een aantal gevallen oninterpreteerbare resultaten op, maar heel vaak is deze vertaling toch een bruikbaar handvat gebleken voor communicatie met de codeur. In de loop van het project zijn we ons zelfs gaan realiseren dat we ook een andere lijn van werken hadden kunnen volgen, namelijk om de codeurs te vragen om een automatische vertaling in het Engels te herzien. In het begin van het project leek dit ons een onhaalbare manier van werken, maar we zijn er anders over gaan denken. Als we alles hadden laten vertalen, waren de kosten niet veel hoger uitgevallen en was controle achteraf voor iedereen veel gemakkelijker geweest.

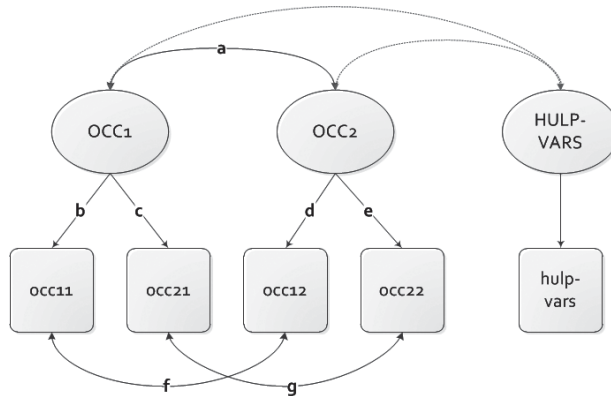
Doordat (a) de werkzaamheden over zoveel codeurs en langere tijd gespreid waren, (b) niet alle codeurs hun werk afmaakten en (c) een aantal mensen *offshore* en om niet werkte, kunnen we geen exacte rapportage geven hoeveel tijd de codeurs in totaal besteed hebben. Onze beste schatting is dat de codeurwerkzaamheden zelf (inclusief training en contacttijd) tussen de 1.600 en 2.000 uur gekost hebben. Daarnaast is het project bemenst geweest door een halftime assistent en heeft ook de hoofdonderzoeker veel tijd besteed aan instructie, datapreparatie en -verwerking. Tabel 2 geeft een overzicht van de gecodeerde gegevens per ronde per land.

Kwaliteit van de gegevens

De kwaliteit van de aldus verkregen coderingen wordt beoordeeld op basis van een factor-analytisch meetmodel waarvan de globale structuur is weergegeven in Figuur 1. In de afgebeelde vorm heeft het model betrekking op twee beroepen, die beide via twee indicatoren worden gemeten. De invulling hiervan wordt in eerste instantie gevormd door de beroepen van de vader en de moeder, die beide zowel worden gemeten via de gesloten vraagstelling als door de door ons gecodeerde strings. Een andere invulling van de indicatoren is codering in ISCO-88 dan wel ISCO-08 (en de bijbehorende schalingen ISEI-88 en ISEI-08), of codering via twee parallel werkende codeurs. Het model is nu weergegeven voor twee beroepen, maar kan gemakkelijk worden uitgebreid naar meerdere beroepen.

In het elementaire model met twee beroepen en dubbele indicatoren zijn zeven te schatten parameters weergegeven: [a] is de latente correlatie tussen de beide beroepen, [b] en [d] zijn de meetcoëfficiënten van de eerste indicator, [c] en [e] zijn de meetcoëfficiënten van de tweede indicator, en [f] en [g] zijn gecorreleerde residuen tussen eenzelfde indicator van verschillende beroepen. Door de aanwezigheid van residuele

Figuur 1 Het elementaire MTMM meetmodel



Latente variabelen (OCC1, OCC2) in ovaal, geobserveerde variabelen (occ11 - occ22) in vierkant.

correlaties [f] en [g] is het meetmodel equivalent met een MTMM (Multi Trait Multi Method) model (Saris & Gallhofer, 2007). Omdat tussen vier geobserveerde indicatoren maar zes correlaties bestaan, is het duidelijk dat het model in deze vorm niet geïdentificeerd is. Er zijn voor de hand liggende manieren om het aantal te schatten parameters te beperken. Zo geven de restricties $[b]=[d]$ en $[c]=[e]$ weer dat het meetmodel voor beide beroepen gelijk is: we veronderstellen dan effectief dat de open en gesloten manieren van meten hetzelfde werken voor vaders en moeders. Hoewel we op die manier vijf te schatten parameters overhouden, maakt dat het model nog niet geïdentificeerd indien we ook gecorreleerde residuen [f] en [g] willen schatten. Als we daarentegen veronderstellen dat de gecorreleerde residuen verwaarloosbaar zijn ($[f] = [g] = 0$), zijn de overblijvende vijf parameters wel geïdentificeerd. Deze residuen zijn echter inhoudelijk relevant. Een andere manier om identificatie te verkrijgen is het model uit te breiden met hulpvariabelen: kenmerken die met beroepen gecorreleerd zijn. Daarvoor staan ons veel mogelijkheden ter beschikking: van al deze betrokken personen hebben we een opleidingsmeting, en ook het huishoudinkomen is een goede kandidaat voor een dergelijke identificerende hulpvariabele. Met inzet van een of meerdere hulpvariabelen zijn alle zeven effecten [a] tot en met [g] afzonderlijk geïdentificeerd.

Tabel 3 geeft de parameterschattingen voor modellen berekend over alle 33 landen en vijf ronden van dataverzameling: de totale ESS-dataset. De effecten zijn berekend via *MLMV estimation* in Stata12 (een efficiënte manier om met missende waarden om te gaan), maar zouden er hetzelfde uitzien in LISREL 8.8 (FIML) of Mplus. In Model 1 schatten we het meetmodel met slechts één identificerende hulpvariabele (beroep respondent). In Model 2 voegen we nog vier hulpvariabelen toe (opleidingen van respondent, vader en moeder, en het huishoudinkomen). Model 1 laat zien dat de meetcoëfficiënten voornamelijk uiteenlopen tussen de toonkaartmeting (fcrude, mcrude) en de gecodeerde meting (fisco88, misco88). Voor vader en moeder is het beeld juist

Tabel 3 Meetkwaliteit geschat in twee modellen die verschillen naar hulpvariabelen. Volledig gestandaardiseerde parameters (met standaardfout)

	Model 1 R1-R5	Model 1a R1-R3	Model 1b R4-R5	Model 2 R1-R5	Model 2a R1-R3	Model 2b 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 commando's:

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)

lsei is de schaling van gedetailleerde beroepen (isco-88), osei van de toonkaartberoepen (crude) (zie Tabel 1)

min of meer hetzelfde. De gecodeerde gegevens leveren een betere meting op dan de voorgedecodeerde vraag: het verschil (0,85 versus 0,77) heeft als interpretatie dat elke correlatie geschat met de gecodeerde gegevens ruim 10% hoger uitvalt dan wanneer je de toonkaartmeting gebruikt. Het verschil is statistisch significant en inhoudelijk relevant. Tegelijk kun je je erover verbazen dat de toonkaartmeting nog zo goed uitpakt. Verder impliceren dezelfde cijfers dat de gecodeerde gegevens weliswaar 10% betere informatie leveren dan de toonkaartmeting, maar tegelijkertijd toch nog 15% afzwakking van ware effecten opleveren. Hier komt de betekenis van een tweede meting naar voren: ook al is een meting goed, ze is nooit perfect. De enige manier om daarachter te komen en de meetfouten uiteindelijk te corrigeren is het opnemen van een tweede meting, ook al is dat een slechte(re).

Er is nog iets bijzonders aan de toonkaartmeting, en dat is de omvangrijke gecorrigeerde residu-term. Dit betekent dat antwoorden op de toonkaarten voor vader en moeder vaker op dezelfde of soortgelijke manier gekozen zijn dan op basis van de gecodeerde open antwoorden en de samenhangen met de hulpvariabelen verwacht wordt. Een interpretatie van deze correlatie als een systematische meetfout ligt voor

de hand: niet wetende hoe op de mysterieuze toonkaart-categorieën te reageren, hebben de respondenten uiteindelijk voor vader en moeder maar hetzelfde uitgekozen. Een optimistische interpretatie is echter ook denkbaar: de toonkaart brengt iets over de ware samenhang in beroepsbeoefening van vaders en moeders in beeld dat verloren gaat in de beroepenstrings. De interpretatie als meetfout lijkt ons echter aanmerkelijk.

Model 1a en Model 1b schatten hetzelfde model, maar nu voor R1-R3 en R4-R5 apart. Deze modellen laten zien welk verschil de invoer van de nieuwe toonkaartmeting in R4 heeft gemaakt. Dat is aanzienlijk: vanaf R4 is de meetkwaliteit van de toonkaartmeting minstens zo goed als die van de gecodeerde beroepen (circa 0,85). Ook daarover kun je je verbazen: hoe kan het zo zijn dat een eenvoudige toonkaartmeting minstens zo goed is als onze moeizaam gecodeerde gedetailleerde beroepen? Overwegingen hierover worden behandeld door Ganzeboom (2005) en De Vries & Ganzeboom (2008), waar dezelfde bevinding wordt gedaan ten aanzien van andere databronnen, namelijk respectievelijk de ISSP 1987, waarin de tweede ESS toonkaart voor het eerst is toegepast, en de ISSP-NL (meerdere ronden sinds 1996) waarin deze toonkaart ook gebruikt is. De overwegingen komen erop neer dat het geven, verwerken en coderen van open antwoorden nu eenmaal op meerdere plaatsen en sequentieel gevoelig is voor random meetfouten: kleine foutjes blijven doorwerken in het proces en na verschillende stapjes kan dit behoorlijk cumuleren. De toonkaartmeting is minder gevoelig voor random verstoringen omdat de weg tussen het beroep van de respondent en de datamatrix uit veel minder stappen bestaat. De keerzijde hiervan kan zijn dat toonkaartmetingen gevoeliger zijn voor systematische meetfouten. Dit blijkt ook in R4-R5 nog steeds zo te zijn. Wel is in deze ronden de gecorreleerde residu-term dramatisch in omvang afgenomen en benadert deze inmiddels die van de gecodeerde open beroepenvraag.

Er is nog iets opmerkenswaardig in de vergelijking tussen Model 1a en 1b: de standaardfouten van de coëfficiënten zijn in 1b afgenomen, terwijl het aantal observaties juist kleiner is. Hoewel het invoeren van een betere tweede meting nauwelijks verschil maakt in de puntschatting van de meetkwaliteit van de eerste meting, vergroot het wel de precisie waarmee deze meetkwaliteit geschat wordt! Statistische significantie is in deze meetmodellen met zoveel eenheden natuurlijk geen interessante vraag, maar dat zal veranderen op het moment dat we de schattingen binnen kleinere steekproeven moeten doen.

Model 2 is uitgebreider: naast beroep respondent zijn nu nog vier andere hulpvariabelen toegevoegd: opleidingen van respondent, vader en moeder, en het huishoudinkomen van de respondent. Alle hulpvariabelen zijn enkelvoudig gemeten en hun onderlinge correlaties zijn perfect gefit – hun enige werking is dat ze ook meehelpen de meetcoëfficiënten van de twee metingen van vaders en moeders beroep te identificeren. Het meest spectaculaire effect van het gebruik van meer hulpvariabelen is dat de standaardfouten ongeveer halveren en dat de effecten dus met veel meer precisie gemeten zijn: meer metingen hebben hier dezelfde uitwerking als meer eenheden! Model 2 is daarom informatiever, met name wanneer we het in een kleinere dataset schatten. In Model 2a en Model 2b worden de exercities herhaald voor de twee vari-

anten van de toonkaart. De conclusies blijven hetzelfde: de tweede toonkaart is aanzienlijk beter dan de eerste, en de inzet van meer hulpvariabelen helpt in het efficiënter schatten van de meetcoëfficiënten.

We hebben het model ook per land per ronde geschat. De geschatte parameters (waarvan de in Tabel 3 aangegeven waarden de gemiddelden zijn) zijn vervolgens in een meta-analyse in verband gebracht met de volgende voorspellers:

- land en ronde;
- vader versus moeder;
- codeur: professionele versus naïeve codeurs.

Bij de resultaten in Tabel 4 is de eerste en belangrijkste conclusie dat er een indrukwekkende gelijkenis optreedt tussen de verschillende ronden en landen: eigenlijk is het nergens fout gegaan. Ook zijn de verschillen tussen de vader- en moeder-coderingen niet groot. De aard van de codeur heeft weinig invloed en dat laat zien dat het heel goed mogelijk is om met min of meer ongetrainde codeurs aan de slag te gaan. Daarin zit een selectie-effect: de slechtere codeurs hebben snel opgegeven of zijn vervangen door andere.

Tabel 4 geeft ook een inzicht in hoe de meetkwaliteit van de gesloten vraagstelling per land en ronde verschilt. Opnieuw valt op dat het hiermee eigenlijk redelijk meevalt, met twee uitzonderingen: Turkije en Nederland. Het gebruikte model verhuut overigens dat de meting in R1-R3 in Nederland nog veel beroerder was. De overtuigende presentatie waarmee Ganzeboom in 2007 de ESS ertoe bracht van toonkaart te wisselen, berustte achterafgezien op een toevalstreffer.

Twee codeurs

In een beperkt aantal gevallen zijn dezelfde gegevens door meerdere codeurs geclassificeerd. Omdat het al moeilijk genoeg was om een enkele codeur voor elk taalgebied te vinden, is dit in de praktijk beperkt gebleven, namelijk tot het Grieks (Cyprus en Griekenland), Italiaans, Sloveens, Lets en Zweeds. Soms is het databestand helemaal dubbel gecodeerd, soms alleen maar random delen met enige overlap. Voor de schatting van de parameters van het meetmodel maakt dat niet zo vreselijk veel uit. We kunnen nu hetzelfde model schatten voor gegevens waarin de beide codeurs de invulling van de indicatoren vormen. Tabel 5 geeft de schatting van deze parameters. Er doet zich maar één geval voor waarin de ene codeur het significant beter doet dan de andere: Letland. In de meeste gevallen liggen de codeurs dicht tegen elkaar aan in kwaliteit. De geschatte meetcoëfficiënten zijn rond de 0,93. Merk op dat dit getal uitsluitend een schatting is van de *door codeurs* aangebrachte meetfouten – omdat de toonkaartmeting in deze modellen niet meedoet, is dit niet gelijk aan de totale meetfout die in Tabel 2 werd gerapporteerd. De 15% totale meetfout is voor ongeveer de helft het gevolg van variantie tussen de codeurs, de rest is elders in het dataverzamelings- en dataverwerkingsproces ontstaan.

Tabel 4 *Determinanten van meetkwaliteit van gecodeerde beroepen en toonkaart-beroepen*

		GECODEERDE BEROEPEN				TOONKAARTBEROEPEN			
		VADER		MOEDER		VADER		MOEDER	
		[b]	se	[d]	se	[c]	se	[e]	se
AT	Oostenrijk	,791	,054	,801	,060	,708	,111	,687	,110
BE	België	,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	Tsjechië	,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-Brittannië	,730	,050	,706	,055	,725	,102	,756	,101
GR	Griekenland	,871	,035	,851	,039	,757	,073	,720	,072
HR	Kroatië	,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	Israël	,837	,038	,772	,042	,734	,078	,692	,077
IS	IJsland	,815	,053	,802	,059	,812	,110	,880	,109
IT	Italië	,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	Roemenië	,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	Slovenië	,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	Oekraïne	,816	,050	,898	,056	,802	,104	,775	,103
ESS-ronde									
	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
Kwaliteit codeur									
	Pro	,010	,024	,000	,027	,016	,049	,010	,049

We voegen nog twee observaties toe over de situaties waarin we met twee codeurs gewerkt hebben:

- Onze modelschattingen maken duidelijk dat er substantiële codeur-onbetrouwbaarheid is. Nog niet zo duidelijk is wat je daarmee nu precies moet doen. Als de ene codeur beter is dan de andere, is de keuze nog voor de hand liggend. Maar onze proefnemingen laten zien dat in de meeste gevallen de twee codeurs allebei

een stukje ware variantie pakken. In die gevallen zou het voor de hand liggen om een derde codeur in dienst te nemen, die per record een keuze maakt tussen de beide andere. Dit is in één geval – Italië – ook zo gedaan (de derde codeur was een professionele socioloog). De uiteindelijke keuze bleek wel beter dan de beide alternatieven, maar uiteindelijk ook niet perfect. Dat suggereert weer dat het beter is om een latente-variabelen-model te gebruiken, waarin beide coderingen in combinatie worden gebruikt.

- Echter, als we gecodeerde beroepen met een toonkaartmeting combineren in een model met latente variabelen, is het eigenlijk niet meer zo interessant een zo goed mogelijke codering van de open gegevens te verkrijgen. Verbetering van de codering heeft alleen maar een fractionele verbetering van de geschatte parameters tot gevolg, en deze gevolgen beperken zich tot de geschatte standaardfouten. In plaats van twee codeurs is het veel beter om twee vragen te stellen, een open en een gesloten, een design dat door een wonderlijk toeval in de ESS generaliseerd is.

Tabel 5 *Meetekwaliteit van dubbele coderingen, gepoolde schatting op basis van gelijkstelling vader en moeder*

	correlatie [a]	[b]=[d] codeur A	[c]=[e] codeur B
CY	0,876	0,939	0,901
GR	0,867	0,924	0,923
IT	0,880	0,958	0,918
LV	0,818	0,937	0,849
SE	0,950	0,960	0,980
SI	0,915	0,964	0,951

ISCO-08 en ISEI-08: naar een nieuwe classificatie en een nieuwe schaling

Zoals opgemerkt is de codering van de ESS-gegevens naar de nieuwe ISCO-08 classificatie beperkt tot landen die in R5 ook strings hebben gedeponereerd voor de beroepen van de respondent en diens partner. Dit bleek uiteindelijk voor vijftien landen het geval. We rapporteren hier over de acht landen waarvoor de ISCO-08 codering tot op heden is afgerond. De beperking tot R5 gegevens is ingegeven door de gedachte dat de nieuwe classificatie vooral van betekenis zou moeten zijn voor actuele beroepsgegevens: de gegevens van vader en moeder hebben betrekking op het tijdstip dat de respondent vijftien jaar oud was, ongeveer dertig jaar geleden. Onder vaders en moeders treffen we bijvoorbeeld weinig tot geen webdesigners aan en je kunt je afvragen of de nieuwe classificatie hier verbetering kan brengen. De codeerprocedure voor deze fase verliep als volgt:

- Er is nu een codeerfile geproduceerd met beroepen van vier personen: vader, moeder, respondent, partner. Ieder record is ook hier voorzien van een indicatie van zelfstandigheid (0/1) en leidinggevendheid (aantal ondergeschikten). Voor

respondenten en partner zijn vaak meer strings (tot vier) beschikbaar. Deze zijn allemaal aan de codeerfile toegevoegd.

- Vervolgens werd een automatische conversieslag gemaakt. De automatische conversieslag geeft de meest aannemelijke ISCO-08 bestemming bij elke ISCO-88 categorie. Bovendien werd een indicator (*howmany*) gecreëerd die aangeeft hoeveel andere bestemmingen ook mogelijk waren geweest. Dit maakt de codeur erop alert eens goed in de manual te kijken. De automatische conversieslag berust op informatie in de ISCO-08 manual.
- Daarna was de codeur aan de beurt: deze bekeek alle records vanuit de gekozen ISCO-08 codering en besloot of herziening een goed idee zou zijn. De codeurs voor deze fase van het project werden opnieuw opgeroepen voor een trainingssessie van circa twee uur. Hierin deden we uit de doeken hoe ISCO-08 zich verhoudt tot ISCO-88. We vroegen de codeurs alle automatisch gegenereerde ISCO-08 codes te bekijken en eventueel te herzien, ook als er een een-op-een relatie met de aanvankelijke ISCO-88 codering was (*howmany=0*).

In het navolgende maken we een vergelijking tussen de aanvankelijke ISCO-88 gegevens en nieuw toegevoegde ISCO-08 gegevens. Bij deze vergelijking zijn diverse overgangen in het spel:

- Allereerst kunnen we een vergelijking maken tussen de ISCO-88 gegevens en de automatisch geconverteerde ISCO-08 gegevens. Deze vergelijking heeft betrekking op een situatie die zich in de toekomst veelvuldig zal voordoen in onderzoek: ISCO-88 wordt automatisch geconverteerd naar ISCO-08 zonder dat de oorspronkelijke strings beschikbaar zijn.
- Vervolgens kunnen we een vergelijking maken tussen ISCO-88 en ISCO-08 zoals herzien door onze codeurs. Equivalent is een vergelijking te maken tussen de automatisch gegenereerde ISCO-08 coderingen en door codeurs herziene coderingen. Deze vergelijkingen leveren een indicatie of een dergelijke herziening inderdaad wat toevoegt. Hierbij moeten we ons wel realiseren dat het verschil niet alleen kan worden teruggevoerd op verschillen in kwaliteit tussen de ISCO-88 en ISCO-08 classificaties, maar ook op voortschrijdend inzicht van onze codeurs die in feite de aanvankelijk in ISCO-88 gegeven codering anders interpreteren in het licht van de ISCO-08 classificatie.
- Vergelijking vindt opnieuw plaats met behulp van de ISEI, maar ook hier zijn twee alternatieven mogelijk: we kunnen ISEI-88 vergelijken met ISEI-08, maar ook voor zowel ISCO-88 als ISCO-08 de ISEI-08 schaling gebruiken. Deze vergelijking geeft een inzicht in hoeverre veranderingen in meetkwaliteit terug zijn te voeren op verandering in kwaliteit van de beide ISEI-schalen.

Tabel 6 Meetkwaliteit van coderingen in ISCO-88 en ISCO-08, gepoolde schatting op basis van gelijkstelling meetcoëfficiënt vader, moeder, respondent, partner. Ongestandaardiseerde coëfficiënten

	(a) Referentie ISCO-88 ISEI-88	(b) Nieuwe schaling ISCO-88 ISEI-08	(c) Automatische conversie ISCO-08 ISEI-08	(d) Met correctie codeur 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	0,999
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

Tabel 6 geeft per land een schatting van het meetmodel weer, waarbij we de categorisering van beroepen in ISCO-88 en schaling in ISEI-88 als uitgangspunt nemen (kolom a). Door gelijkstelling van het meetmodel van vader, moeder, respondent en partner is het resultaat van de verschillende modellen eenvoudig uit te drukken in verhouding tot dit referentie-effect. In kolom (b) zien we allereerst wat er gebeurt als we de beroepenclassificatie hetzelfde laten, maar uitdrukken in de nieuwe schaling ISEI-08. We merken op dat dit in alle landen de meting verbetert, gemiddeld met zo'n 2%. Doen we in kolom (c) eerst een automatische conversie naar ISCO-08 en drukken we deze uit in ISEI-08, dan is er in doorsnee nog een kleinere verbetering, circa 1%. Deze verbetering is het gevolg van verbetering van de ISCO-classificatie zelf. In kolom (d) voegen we toe wat er gebeurt als we de door de codeurs herziene gegevens als meting nemen. De veranderingen die door onze codeurs op basis van de strings zijn aangebracht, pakken wisselend uit en brengen gemiddeld nog niet eens 1% verbetering. Alles bijeengenomen is de nieuwe meting zeker eerder een verbetering dan een verslechtering, maar de nieuwe ISEI-schaling draagt hier nog het meeste aan bij, niet de nieuwe classificatie.

Samenvatting en conclusies

In het gerapporteerde ESS Developmental Project *Improving the Measurement of Social Background in the European Social Survey* zijn waar het de beroepen aangaat tot op heden de volgende werkzaamheden verricht:

- Alle strings van ouderlijke beroepen (vader en moeder) in ESS R1-R5 zijn gecodeerd in ISCO-88. Deze resultaten staan inmiddels vrij ter beschikking aan de internationale onderzoekswereld. In totaal zijn 330.000 beroepen in 25 talen van een code voorzien. Dat heeft circa dertig codeurs ongeveer 1.800 arbeidsuren gekost. Het is de vraag of de besparing in werk die de ESS zich had toegedacht wel echt een verstandige zet is geweest. De internationale onderzoekswereld mag in

ieder geval het NWO dankbaar zijn dat het de hiervoor benodigde fondsen ter beschikking heeft gesteld. Dat is een zeer nuttige investering geweest.

- Uit R5 zijn tot op heden voor acht landen alle beroepsgegevens (niet alleen van ouders, maar ook van respondent en partner) omgezet in de nieuwe classificatie ISCO-08. Deze omzetting, die in feite een dubbele codering is, biedt ook gelegenheid de kwaliteit van de oorspronkelijke codering van de beroepen van respondent en partner te controleren – althans waar het deze acht landen in R5 aangaat.
- Beroepencodes worden niet als zodanig in statistisch onderzoek gebruikt. Daarvoor moeten we afgeleide schalen hebben, zoals de *International Socio-Economic Index of occupational status*. Tijdens het project is een nieuwe versie van deze schaal ontwikkeld om te gebruiken in combinatie met de nieuwe beroepenclassificatie: ISEI-08, die geacht kan worden verwisselbaar te zijn met ISEI-68 en ISEI-88, die voor eerdere versies van de ISCO zijn ontwikkeld. De nieuwe ISEI-08 is in feite niet ontwikkeld met behulp van ISCO-08 gegevens, maar via substituut ISCO-08 gegevens. Dit geeft de nieuwe schaal een voorlopige status, maar levert ook de mogelijkheid op om haar zowel toe te passen op ISCO-88 en ISCO-08 gegevens. Dit stelt ons in staat om uit onze modellen af te lezen in welke mate veranderingen in meetkwaliteit zijn terug te voeren op verandering van classificatie (ISCO-88 naar ISCO-08), dan wel op veranderingen in afgeleide statusschaal (ISEI-88 naar ISEI-08).
- Voor de evaluatie van de kwaliteit van de beroepencoderingen maken we gebruik van een factor-analytisch meetmodel dat ons (a) in staat stelt een zeer krachtige gepoolde schatting van de meetkwaliteit van elke meting te bepalen, en (b) de mogelijkheid biedt om gecorreleerde systematische meetfouten te schatten. Door in het model voor het beroep van de ouders ook gebruik te maken van een (optimaal geschaalde) meting op basis van de toonkaart kunnen we ook opsporen welke fouten er mogelijk gemaakt zijn bij de initiële codering.

De resultaten van deze werkzaamheden kunnen als volgt worden samengevat:

- Een eerste conclusie uit de analyses is dat de meetkwaliteit van de door ons geco-deerde ouderlijke beroepsgegevens in de ESS zeer redelijk is. We concluderen dit uit de meetindicator voor deze open beroepen, waarin deze in een model zijn samengebracht met de gesloten vraagstelling. De meetkwaliteit is rondom de 0,85, met slechts weinig fluctuaties tussen landen en rondes. De gesloten vraagstelling heeft een meetkwaliteit van rond de 0,75, aanzienlijk minder dus. De formulering van de gesloten vraagstelling is wel veranderd, tussen R3 en R4, en deze verandering is inderdaad een merkbare verbetering geweest van de meetkwaliteit van deze vraagstelling.
- Meetkwaliteit wordt niet alleen beïnvloed door de kwaliteit van de coderingen: ook het antwoordgedrag van de respondenten en alles wat er gebeurt voordat zo'n antwoord een string in het ESS data-archief wordt (interviewer, data-invoer, databewerking) is daarop van invloed. In een beperkt aantal gevallen hebben we gebruik kunnen maken van een dubbele codering voor een nationaal databestand, waardoor we codeervariantie kunnen scheiden van responsvariantie. Deze

voorbeelden wijzen uit dat onafhankelijk werkende codeurs gemiddeld genomen 0,86 intercodeurs-correlatie bereiken. In vrijwel alle gevallen ontlieden de beide codeurs elkaar niet veel in kwaliteit: ze maakten allebei fouten, de een niet veel meer dan de ander. Vanuit het gezichtspunt van een meetmodel kan men daaruit concluderen dat elke codering sowieso een vertekening van circa 7% teweegbrengt. Het lijkt niet zo veel, maar de effecten daarvan zijn bij het bestuderen van intergenerationale correlaties toch tamelijk dramatisch. Dat komt omdat de meetfout dan twee keer optreedt.

- Voor de acht landen waarbij R5 zowel in ISCO-88 en ISCO-08 gecodeerd is, kunnen we concluderen dat de overgang naar ISCO-08 en de bijbehorende ISEI-08 een duidelijke verbetering inhoudt. Gemiddeld nemen de meetcoëfficiënten met circa 3% toe. Met enige voorzichtigheid kan deze 3% gesplitst worden in een gedeelte dat te danken is aan een verandering van classificatie (ongeveer 1%) en een gedeelte dat voortkomt uit verandering naar de nieuwe ISEI-08 schaal (ongeveer 2%). Deze schattingen verschillen tussen de landen, maar geven toch aanleiding om, algemeen gesproken, de toepassing van de nieuwe classificatie en de daarbij behorende ISEI-schaling met vertrouwen tegemoet te zien.

Discussie

Het voorgenoemde productieplan van het ESS-DEVO-project is wat beroepen aangaat voor een groot deel gerealiseerd en de producten staan inmiddels ter beschikking van de internationale onderzoekswereld. Er zal nog verder gewerkt worden aan de ISCO-08 codering voor enige andere landen waarvoor dit zinvol en mogelijk is: dat zijn landen waarin men in R5 ook de strings van respondent en partner heeft gedeponed. Onze resultaten geven echter geen reden om de dubbele codering aanzienlijk uit te breiden. De door ons uitgewerkte automatische conversie van ISCO-88 naar ISCO-08 geeft ook bevredigende resultaten, zodat ook de aanpassing van bestaande gecodeerde beroepsgegevens naar de nieuwe standaard, indien de oorspronkelijke strings niet meer ter beschikking staan, met vertrouwen kan worden uitgevoerd.

Doelstelling van het ESS-DEVO-project was ook om een discrete beroepsklasse-indeling te ontwerpen voor de nieuwe ISCO-08 classificatie. Het werk aan deze zogenaamde International Socio-Economic Classes [ISEC-08] is inmiddels onderweg (Ganzeboom & Treiman, 2013).

Noten

1. Het artikel is het eerste verslag van het ESS Developmental Project #3 (NWO, 471-09-005), 'Improving the Measurement of Social Background in the European Social Survey'. Graag danken we allen die aan het project meewerkten. Assistenten in het project waren achtereenvolgens Natasha Nikitina, Tim Mickler en David Nikoloski. Medewerking werd verleend door ruim 30 beroepencodeurs (zie Appendix A). Het NSD data-archief in Bergen (NO) maakte de beroepenstrings toegankelijk. De resultaten van het project en de gebruikte *tools* zijn te vinden op home.fsw.vu.nl/hbg.ganzeboom/ess-devo.

2. Het ESS-DEVO-project omvat ook een component over opleidingen in de ESS. Hierbinnen is door Schröder & Ganzeboom (2013) de International Standard Level of Education [ISLED] ontwikkeld. Nader verslag over deze component is gedaan door Schröder & Ganzeboom (2010, 2011, 2013) en zal worden opgenomen in de dissertatie van Schröder en de daaruit voortvloeiende tijdschriftartikelen.

Literatuur

- Erikson, R., Goldthorpe, J.H. & Portocarero, L. (1979). 'Intergenerational class mobility in three Western European societies: England, France and Sweden.' *British Journal of Sociology* 30, 415-441.
- Ganzeboom, H.B.G. (2010). *Tools for deriving occupational status measures from ISCO-08 with interpretative notes to ISCO-08*. <http://home.fsw.vu.nl/hbg.ganzeboom/isco08/index.htm>.
- Ganzeboom, H.B.G., De Graaf, P. & Treiman, D.J. (1992). 'A standard international socio-economic index of occupational status.' *Social Science Research* 21, 1-56.
- Ganzeboom, H.B.G. & Treiman, D.J. (1996). 'Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations.' *Social Science Research* 25, 201-239.
- Ganzeboom, H.B.G. & Treiman, D.J. (2013). 'Occupational Stratification Measures for the New International Standard Classification of Occupation 2008 [ISCO-08]; with a Discussion of the New Classification.' Working Paper.
- Ganzeboom, H.B.G. & Schröder, H. (2009). 'De waarde van diploma's: een kwantificatie van de ESS-NL-categorieën.' in: Ganzeboom, H.B.G., Wittenberg, M. (Red.), 2009. *Nederland in Vergelijkend Perspectief. Proceedings Tweede Nederlandse Workshop European Social Survey*. Amsterdam: Aksant [DANS Symposium Publicaties #3], pp. 89-108.
- Ganzeboom, H.B.G. (2005). 'On the Cost of Being Crude: A Comparison of Detailed and Coarse Occupational Coding.' in: Hoffmeyer-Zlotnik, J.H.P., *Methodological Aspects of Cross-National Research, Mannheim: ZUMA-Nachrichten* [Special Issue #11], 2005, pp. 241-258.
- Ganzeboom, H.B.G. & Treiman, D.J. (2003). 'Three Internationally Standardised Measures for Comparative Research on Occupational Status.' in: Hoffmeyer-Zlotnik, J.H.P., Wolf, C. (Eds.). *Advances in Cross-National Comparison. A European Working Book for Demographic and Socio-Economic Variables*. New York: Kluwer Academic Press, pp. 159-193.
- Ganzeboom, H.B.G. (2010). *Do's and Don'ts of Occupation Coding*. <http://home.fsw.vu.nl/HBG.Ganzeboom/Pdf/2010-DO-and-DONTS-Occupation-coding.pdf>.
- ILO [International Labour Organisation] (1990). *International Standard Classification of Occupations ISCO-88*. Geneva: ILO.
- ILO [International Labour Organisation] (2012). *International Standard Classification of Occupations ISCO-08*. Geneva: ILO.
- Rose, D. & Harrison, E. (Eds.) (2010). *Social Class in Europe. An introduction to the European Socio-Economic Classification*. Routledge, Oxford.
- Schröder, H. & Ganzeboom, H.B.G. (2012). 'De waarde van diploma's in Nederland: de ESS-NL kwantificaties getoetst'. in: Aarts, K., Wittenberg, M. (Red.), *Nederland in de jaren nul. Proceedings Derde Nederlandse Workshop European Social Survey*. Amsterdam: Palas Publications, pp. 63-78.
- Saris, W.E. & Gallhofer, I.N. (2007). *Design, Evaluation, and Analysis of Questionnaires for Survey Research*. New York: Wiley.

- Schröder, H. & Ganzeboom, H.B.G. (2013). 'Measuring and Modeling Level of Education in European Societies'. *European Sociological Review* (forthcoming).
- Treiman, D.J. (1977). *Occupational Prestige in a Comparative Perspective*. New York: Wiley.
- Vries, J. de & Ganzeboom, H.B.G. (2008). 'Hoe meet ik beroep? Open en gesloten vragen naar beroep toegepast in een statusverwervingsmodel'. *Mens & Maatschappij* (83,1), pp. 71-96. + 'Rectificatie'. *Mens & Maatschappij* (83,2), pp. 190-191.

Appendix A Codeurs die meewerkten aan het codeerproject

Bekir, Semiha	BG	naïef
Bigaj, Maria	PL	naïef
Cappa, Ingrid	IT	naïef
Celisceva, Karina	LV	naïef
Cohen, Yasmin	IL	naïef
Delic, Jasna	HR, SI	naïef
Denaro, Daniela	IT	naïef
Domanski, Henryk	PL	pro
Duta, Adriana	RO	naïef
Dzeguze, Dace	LV	naïef
Erola, Jani	FI	pro
Ferre, Pernille	NO	naïef
Ganzeboom, Harry	NL, BE	pro
Güveli, Ayse	TR	pro
Ignasz, Szofia	HU	pro
Izlay, Zoltan	HU	naïef
Jönsas, Katia	FI	naïef
Joye, Dominique	CH	pro
Kalogreades, Lawrence	GR, CY	naïef
Karpejute, Evelina	LT	naïef
Kekic, Sabina	HR, SI	naïef
Kirköen, Benedicte	NO	naïef
Kmetova, Anna	SK	naïef
Kramberger, Anton	SI	pro
Marques, Laura	ES	naïef
Menegatou, Elini	GR, CY	naïef
Meraviglia, Cinzia	IT	pro
Mickler, Tim	DE, DK, GB	pro
Nettelbladt, Sonja	SE	naïef
Nikitina, Natasha	UA, RU	naïef
Nikoloski, David	BE	pro
Pauw, Lisanne	FR	naïef
Schröder, Heike	AT, DE	pro
Sigurardottir, Runa	IS	naïef
Taht, Kadri	ES	pro
Wettflö, Anton	SE	naïef
Zuccotti, Carolina	ES, PT	pro

Appendix B Aantal in ISCO-88 gecodeerde strings, per land

	Coding-status			Totaal
	Geen beroep	Nog niet af	Gecodeerd	
AT	2.489	0	11.721	14.210
BE	2.282	0	11.370	13.652
BG	452	0	10.253	10.705
CY	756	0	4.398	5.154
CZ	2.086	0	13.916	16.002
DE	4.366	0	19.448	23.814
DK	1.095	0	12.078	13.173
EE	1.243	0	9.758	11.001
ES	1.715	0	11.215	12.930
FI	285	0	15.939	16.224
FR	1.119	0	12.431	13.550
GB	2.077	0	14.711	16.788
GR	37	1	14.597	14.635
HR	13	0	2.927	2.940
HU	1.999	0	10.862	12.861
CH	5.721	0	11.895	17.616
IE	3.317	0	11.109	14.426
IL	2.218	0	9.224	11.442
IS	17	0	858	875
IT	275	0	2.835	3.110
LT	2.087	0	5.271	7.358
LU	214	0	3.783	3.997
LV	286	0	5.677	5.963
NL	1.944	0	11.312	13.256
NO	208	0	12.738	12.946
PL	284	0	13.801	14.085
PT	3.951	0	13.496	17.447
RO	3.917	0	4.653	8.570
RU	1.985	0	11.259	13.244
SE	3.861	0	13.588	17.449
SI	9	5	6.387	6.401
SK	69	712	9.431	10.212
TR	48	0	3.248	3.296
UA	1.217	588	11.009	12.814
Totaal	53.642	1.306	337.198	392.146