

De waarde van diploma's in Nederland

De ESS-NL kwantificaties getoetst

Heike Schröder en Harry B.G. Ganzeboom

In een eerdere paper hebben we kwantificaties gepresenteerd van de 13 opleidingscategorieën die in ESS-NL onderscheiden worden. Deze kwantificaties werden bepaald als optimale schaling van de opleidingscategorieën. Uit de eerdere analyses van de ESS-data bleek dat de nieuwe schaal opleiding beter representeert dan bestaande kwantificaties.

In deze bijdrage presenteren we een validatie van de aldus bepaalde opleidingscores aan de hand van verse data, ontleend aan het ISSP-NL 2002-2008 (N = 4913). Onze algehele conclusie is dat de op basis van ESS-NL geconstrueerde opleidingskwantificatie een valide en sterke meting van het niveau van opleidingsniveau in Nederland oplevert.

Inleiding

Onderwijssystemen verschillen sterk tussen landen en nationale onderwijssystemen veranderen bovendien in de loop van tijd. Het Nederlandse onderwijssysteem is hiervan een goed voorbeeld. Het kent een grote hoeveelheid programma's en typen die bovendien door de tijd heen aan diverse hervormingen onderhevig zijn geweest. Deze grote hoeveelheid historische en huidige onderwijstypen maakt het meten van opleidingsniveau in vragenlijsten lastig. Een uitputtende opsomming van alle mogelijkheden ligt buiten het bereik van een gesloten vraagstelling. In plaats daarvan worden meestal onderwijskwalificaties die in verschillende historische perioden zijn behaald over één kam geschoren en wordt er volstaan met een enkele aanduiding en het opsommen van een beperkt aantal voorbeelden. Ook de diversiteit aan alternatieven binnen het gehele onderwijsstelsel wordt nogal eens onder het vloerkleed geveegd door een beperkte hoeveelheid alternatieven te presenteren. Op die manier worden verschillende soorten onderwijs a priori gelijkgesteld en de verantwoordelijkheid voor de juiste keuze binnen deze *mixed bag* wordt gelegd bij de respondenten. De hoeveelheid informatie die op deze manier verloren gaat kan aanzienlijk zijn en dit neemt toe naarmate er bij voorbaat minder categorieën onderscheiden worden.

In de European Social Survey is voor Nederland (ESS-NL) een nieuwe onderwijsvraagstelling gebruikt. Deze nieuwe vraagstelling is met 13 categorieën uitgebreider dan de meeste gangbare vraagstellingen. Wij hebben eerder (Ganzeboom & Schröder

2009) onderzocht hoe de verschillende opleidingscategorieën in ESS-NL zich ten opzichte van elkaar verhouden wat betreft inputs (ouderlijke opleidingsniveaus en beroepen) en outputs (verwerving van beroep en partner). Dit resulteerde in een optimale schaling van de onderscheiden opleidingscategorieën. Uit een vergelijking met andere opleidingsmetingen bleek dat deze gedetailleerde schaling het niveau van de opleiding merkbaar beter representeert dan zowel de internationaal geharmoniseerde opleidingsmaat in de ESS, ISCED (*International Standard Classification of Education*) als de eveneens in de ESS voorhanden alternatieve opleidingsmaat, de door de respondent gerapporteerde duur van de gehele onderwijsloopbaan.

In een volgende stap hebben we daarna de verschillende opleidingsmetingen gecombineerd in een multiple indicatoren model waardoor we de verschillende combinaties van metingen konden vergelijken met de afzonderlijke indicatoren. Uit deze vergelijkingen hebben wij geconcludeerd dat het gebruik van de metingen tegelijkertijd in een latent variabelenmodel de meetkwaliteit meer verbetert dan meer detail in een afzonderlijke indicator.

De door Ganzeboom & Schröder (2009) gegenereerde optimale schalingen van de Nederlandse opleidingskwalificaties waren verkregen in een analyse van de eerste drie rondes (R1-R3, 2002-2006) van de ESS. Inmiddels zijn ook de data van de vierde ronde (R4, 2008) beschikbaar. Wij hebben met gebruik hiervan de opleidingscategorieën in de ESS opnieuw geschaald, deze keer dus met de data van alle vier de rondes. De nieuwe schaal is bovendien van een naam voorzien: ISLED, de *International Standard Level of Education*. De Nederlandse variant duiden we aan als ISLED-NL.

We hebben in ons eerdere artikel aangetoond dat de meetkwaliteit van ISLED-NL superieur is aan die van zowel ISCED als de duurmaat. Men zou echter kritisch kunnen opmerken dat het nogal wies is dat met ESS-data ISLED beter werkt dan de alternatieven. Dat waren immers de brondata, en dus is de echte proef op de som nog niet genomen. Voor een meer serieuze validatie van de ISLED moeten we nagaan of deze schaal ook een betere meetkwaliteit vertoont als we met andere data werken. In dit hoofdstuk passen we de ISLED-NL toe op andere data, namelijk op gegevens uit 6 rondes ISSP-NL (International Social Survey Programme 2002-2008).

Theoretische achtergrond

Opleidingsniveau is een breed concept en voordat we kunnen bepalen hoe men dit het beste kan meten, moet wel helder zijn welke conceptuele betekenis van opleiding men voor ogen heeft (Braun & Müller, 1997). Vanuit een stratificatieperspectief zijn wij geïnteresseerd in de sociale betekenis van opleiding. Hierbij denken we niet alleen aan de waarde van kwalificaties op de arbeids- en huwelijksmarkt, maar ook aan de sociale factoren die bepalen welk opleidingsniveau mensen verwerven, in het bijzonder ouderlijke beroeps- en opleidingsniveaus.

We baseren ons hierbij op twee belangrijke theorieën over de rol van onderwijs. In de eerste plaats is dat de theorie zoals neergelegd in het statusverwervingsmodel van Blau en Duncan (1967). Volgens dit inmiddels klassieke model neemt opleiding een centrale plaats in het proces van sociale reproductie in: het opleidingsniveau van een

individu wordt bepaald door het opleidings- en beroepsniveau van de vader, en op haar beurt bepaalt de opleiding dan weer het (eerste en huidige) beroep van het individu zelf. Bij de ontwikkeling en validering van een maat voor opleidingsniveau maken we direct gebruik van dit model, maar breiden het uit door aan de inputkant opleiding en beroep van de moeder toe te voegen en aan de outputkant het opleidingsniveau van de partner, dat net als beroep sterk met opleidingsniveau correleert.

In de tweede plaats maken we gebruik van de theorie van positionele goederen (Hirsch 1976). Volgens dit idee produceert het onderwijssysteem een enkelvoudige hiërarchie van mensen die met hun kwalificaties op de arbeidsmarkt en huwelijksmarkt met elkaar wedijveren om aantrekkelijke banen en huwelijkspartners. Van belang in het positionele goederenperspectief is niet de absolute waarde van diploma's (d.w.z. de inhoud van het geleerde), maar hun relatieve waarde vergeleken met andere diploma's. Men kan deze theorie, al naar gelang de veronderstelde functie van diploma's, verschillend interpreteren. Wij sluiten ons aan bij Thurow (1975) die de functie van opleiding definieert in termen van sociale uitsluiting. Mensen met een hoger opleidingsniveau hebben een betere concurrentiepositie in de banenrij en bijgevolg toegang tot de eveneens hiërarchisch hoger geordende beroepen en hun beloningen. Het voordeel van sommige individuen gaat noodzakelijk ten koste van anderen, minder hoog opgeleiden die geen toegang hebben tot deze beroepen of daaruit worden verdrongen. Beroepen behoren eveneens tot de positionele goederen. In dit perspectief is er dus geen relatie tussen de inhoud van het onderwijs en de inhoud van de baan (laat staan de kwaliteiten van de partner), maar kunnen we opleidingsschalings op basis van criteriumvariabelen als een rangorde afleiden.

Implicaties voor het meten van opleidingsniveau

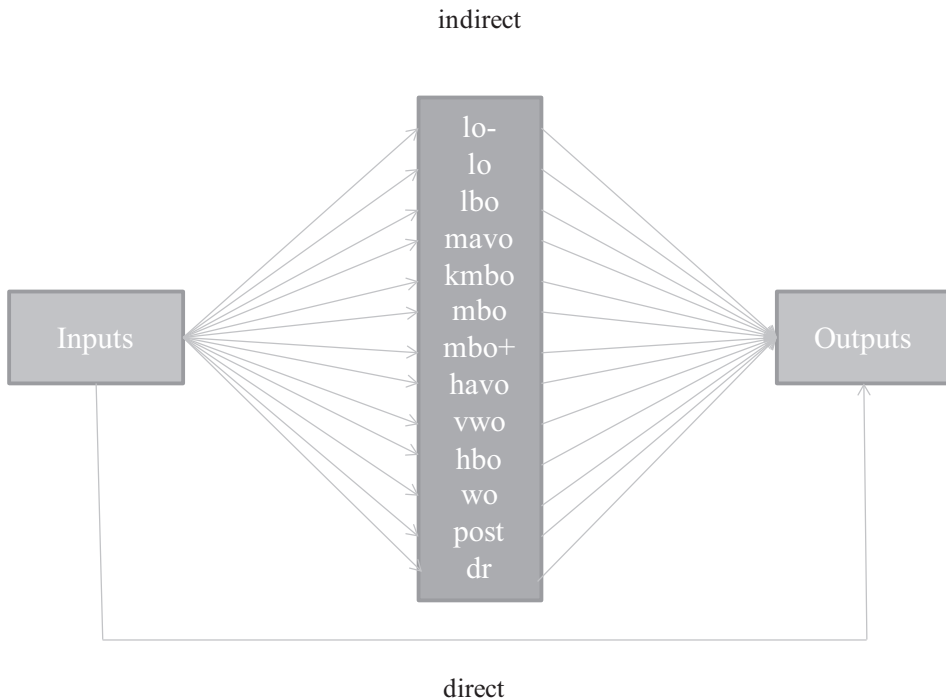
Door onze meetmethodes af te leiden uit sociale stratificatiemodellen verschuiven wij de focus van gangbare schalingsmethoden op basis van formele kenmerken van opleidingsprogramma's, zoals het aantal jaren dat nodig is om een opleiding af te ronden of de verkregen diploma's, naar een functioneel kenmerk, namelijk het bepalen van de plaats van mensen in de sociale hiërarchie. Anders uitgedrukt: wij vestigen de aandacht in plaats van op statische op proceskenmerken van opleidingssystemen. Concreet betekent een en ander dat wij nagaan hoe begeerd verschillende Nederlandse opleidingsniveaus zijn bij bepaalde sociale groepen, in termen van ouderlijke kenmerken als beroep en opleidingsniveau en wat ze uiteindelijk opleveren op de arbeids- en huwelijksmarkt, dat wil zeggen naar welke beroepen en partners zij leiden.

Een uit de positionele goederentheorie afkomstig grondbeginsel voor de hier voorgestelde meetprocedure is het idee dat elk onderwijsstelsel hoe dan ook altijd een enkelvoudige hiërarchie van hoger en lager opgeleiden produceert, die zich zowel uit in succes op arbeidsmarkt en huwelijksmarkt, als in de mate waarin groepen met een hogere maatschappelijke status erin slagen de aantrekkelijke positie voor hun kinderen te bestemmen en kinderen van lagere statusgroepen daarvan uit te sluiten. Van deze aanname maken we gebruik door na te gaan hoe deze hiërarchie zich uit in de samenhang met de gekozen criteriumvariabelen, te weten ouderlijke opleidings- en

beroepsniveaus, het beroep van de respondent zelf en het opleidingsniveau van de partner.

Methodologische achtergrond

Figuur 1 MIMIC-model



Onze meetprincipes zijn direct te vertalen in een MIMIC-model (zie figuur 1). In dit model intermedieert opleidingsniveau tussen meerdere inputvariabelen en meerdere outputvariabelen (MIMIC staat voor Multiple Indicators, Multiple Causes). De relevantie van dit MIMIC-model voor het bepalen van opleidingsniveau kan niet alleen substantieel, vanuit het statusverwervingsmodel en de positionele goederentheorie, worden beredeneerd, maar ook vanuit de klassieke meettheorie en wel als volgt. In een causale keten van drie variabelen veronderstellen we dat X een effect heeft op Y uitsluitend en alleen via Z. Of deze veronderstelling in de praktijk met empirische gegevens overeenstemt, hangt in sterke mate af van de meetkwaliteit van de tussenliggende variabele Z (Kelley 1973). Als X of Y niet perfect gemeten is, vinden we verzwakte coëfficiënten $b(YZ)$ en $b(ZX)$, maar blijft gelden $b(YX|Z) = 0$. Als daarentegen Z niet goed gemeten is, verandert dit en zal in het algemeen gelden: $b(YX|Z) > 0$. Om een causale keten adequaat te schatten is het daarom noodzakelijk om meetproblemen in Z te corrigeren. Deze redenering kunnen we toepassen op het probleem van optimale schaling van opleidingscategorieën. Een suboptimale schaling is er een

die de intermediërende rol van opleiding afzwakt en grotere directe effecten van inputs op outputs te zien geeft. Een optimale schaling daarentegen is er een die maximaal recht doet aan de mediërende rol van het opleidingsniveau en de indirecte effecten van inputs op outputs (via opleidingsniveau) zo groot mogelijk maakt, ofwel de directe effecten van inputs op outputs minimaliseert.

Nu we het theoretisch raamwerk voor het meten van opleidingsniveau hebben afgebakend, is de volgende stap de voor- en nadelen van verschillende meetprincipes tegen elkaar af te wegen. Als we opleidingen uit verschillende landen of perioden op vergelijkbare manier willen meten, hebben we de keuze uit drie verschillende werkwijzen die we achtereenvolgens kort de revue laten passeren: geharmoniseerde opleidingstypologieën via een grootste gemene deler, opleidingsduur en opleidingschalen (voor een uitvoerigere inleiding zie Ganzeboom & Schröder 2009).

De meest gebruikte aanpak is de strategie van de grootste gemene deler (GGD), waarbij men zoekt naar benoeming van de gemeenschappelijke elementen in de te harmoniseren indelingen. In de praktijk van cross-nationaal vergelijkend onderzoek, zo ook in de ESS en in de ISSP, wordt er voor een GGD strategie veelal gebruikgemaakt van de eerdergenoemde ISCED-classificatie. De moeizame uitkomsten van deze strategie zijn gemakkelijk te voorzien. Ten eerste zal een GGD-strategie altijd leiden tot een verlies aan informatie doordat een GGD per definitie minder details kent dan de te harmoniseren brongegevens. Ten tweede is het soms niet mogelijk om voor een aantal categorieën een gemeenschappelijke noemer te vinden. Deze problemen worden bovendien groter naarmate men meer gegevensbronnen harmoniseert: meer gegevens leidt tot minder details in de meting. Dat de problemen van GGD-classificaties in de praktijk grote schadelijke gevolgen hebben voor uitkomsten is herhaaldelijk aangetoond (Kerckhoff & Dylan 1999; Schneider 2008).

Een tweede mogelijke strategie om opleidingscategorieën vergelijkbaar te maken is gebaseerd op de tijdsduur, uitgedrukt in het aantal jaren dat individuele opleidingstrajecten in beslag nemen. Deze eveneens veelvuldig in survey-onderzoek toegepaste strategie (wederom ook in ISSP en ESS) is gebaseerd op de veronderstelling dat naarmate men langer in het onderwijssysteem verblijft, men ook steeds hogere kwalificaties verkrijgt. Ook duurmaten zijn echter al vaak problematisch gebleken (Hout & DiPrete 2006; Schneider 2007). Dat geldt vooral voor sterk gestratificeerde opleidingsystemen, waar eenzelfde lengte van programma's tot kwalificaties van verschillend niveau kan leiden. Dit is ook voor Nederland het geval, waar bijvoorbeeld vwo en mbo even lang duren.

Een derde, minder gebruikelijke manier om opleidingscategorieën op een noemer te brengen is via een gemeenschappelijke schaling. Bij het schalen van opleiding worden opleidingsscores dusdanig gegenereerd dat de correlatie tussen de geschaalde opleidingsvariabele en een gegeven criteriumvariabele gemaximeerd wordt (Braun & Müller 1997). Er zijn twee strategieën van opleidingsschalings in de literatuur te vinden. Eén werkwijze is die van Treiman en Terrell (1975) die in een vergelijking van de Amerikaanse en Britse onderwijssystemen zogenaamde *effect-proportional scaling* toepassen om vergelijkbare opleidingsscores te construeren. Zij gebruiken een output-

variabele, de beroepsstatus van de respondent, als criteriumvariabele. Smith en Garnier (1986) genereren juist een opleidingschaal door (in loglineaire modellen) een inputvariabele als criteriumvariabele in te zetten, het beroep van de vader. Hun aanpak zou als oorzaak-proportioneel (*cause-proportional*) schalen betiteld kunnen worden.

Wij sluiten in onze methode bij de schalingstrategie aan. We gaan een stap verder door de twee genoemde benaderingen te integreren door zowel in- en outputvariabelen als criteriumvariabelen voor onze opleidingschaal te benutten (*cause- and effect-proportional scaling*). En we wijken nog op een tweede manier af van onze voorgangers door zowel aan de input- als ook aan de outputkant niet één, maar meerdere criteriumvariabelen te gebruiken.

Het Nederlandse onderwijssysteem in survey-onderzoek

Nederland kent een zeer gedifferentieerd en sterk gestratificeerd onderwijssysteem dat bovendien in de loop der jaren ingrijpend en herhaaldelijk is hervormd (denk bijvoorbeeld aan de Mammoetwet van 1968 of de Wet Educatie en Beroepsonderwijs 1996). Zonder zelfs nog naar inhoudelijke variatie te verwijzen zijn er daardoor tientallen verschillende schooltypen en vervolgopleidingen te onderscheiden, verdeeld over (de verschillende niveaus van) basis-, voortgezet en hoger onderwijs.

Op grond van deze diversiteit van huidige en historische opleidingstypen is het meten van opleidingsniveau in vragenlijsten bepaald geen gemakkelijke taak. Aangezien het theoretisch moeilijk is om al die opleidingen a priori hiërarchisch te ordenen en praktisch onmogelijk om een uitputtende opsomming ervan in een vragenlijst op te nemen, zijn compromissen in survey-onderzoek onontkoombaar. Opvallend is dat er geen standaardvraagstelling naar het Nederlandse onderwijssysteem blijkt te bestaan. Wie vragenlijsten uit verschillende onderzoeken naast elkaar legt, komt tot de ontvutsende ontdekking dat opleidingsniveau in bijna elke vragenlijst anders is gemeten. De voorgelegde indelingen lijken uiteraard op elkaar, maar vragenlijsten verschillen wat betreft de precieze vraagstelling, het aantal onderscheiden categorieën, de (impliciet gepresenteerde) hiërarchische ordening en de concrete voorbeelden van de categorieën waaruit de respondent moet kiezen.

Een ander probleem is dat er binnen één vragenlijst vaak naar het opleidingsniveau van meerdere personen wordt gevraagd en dat ook hier geen sprake is van een standaard. Hierbij valt onmiddellijk op dat er vaak aanzienlijk meer zorg wordt besteed aan het meten van het opleidingsniveau van de respondent dan aan dat van andere personen, zoals partner of ouders van de respondent. Vaak is het opleidingsniveau van de respondent gedetailleerder en dubbel in kaart gebracht. Aangezien minder detail neerkomt op verlies van informatie en derhalve afgezwakte regressiecoëfficiënten is dit betreurenswaardig.

Het Europees Sociaal Survey (ESS)

In de ESS wordt opleidingsniveau voor de respondent op twee onafhankelijke manieren gemeten. Er wordt zowel gevraagd naar het aantal jaren onderwijs dat mensen hebben afgerond als naar hun hoogst behaalde kwalificatie. Terwijl de vraag naar de

opleidingsduur voor alle landen in de ESS hetzelfde is, zijn de vragen naar de hoogst behaalde opleiding specifiek naar land gesteld in de vorm van meer of minder gedetailleerde onderwijsclassificaties die worden voorgelegd aan de respondent. Deze land-specifieke classificaties lopen in de ESS sterk uiteen wat betreft het aantal onderscheiden categorieën, variërend van 5 categorieën voor Groot-Brittannië tot en met 19 voor Luxemburg. Met 13 categorieën hoort Nederland tot de landen met de meest gedetailleerde classificaties. Om deze maten vergelijkbaar te maken worden ze vervolgens geharmoniseerd in de 7 hoofdcategorieën van ISCED. Voor de meeste landen betekent dit dat het aantal onderscheiden categorieën sterk wordt gereduceerd, terwijl er belangrijke onderscheidingen wegvallen. In de praktijk is een aantal van de ISCED hoofdcategorieën nauwelijks gevuld.

Voor Nederland wordt er in de ESS een tamelijk gedetailleerde onderwijsclassificatie gebruikt die met 13 categorieën uitgebreider is dan welke andere bestaande indeling dan ook (zie Appendix). Ten opzichte van in ander Nederlands onderzoek gebruikte metingen zijn er met name meer niveaus onderscheiden in het middelbaar beroeps onderwijs (mbo) en is postacademisch onderwijs toegevoegd. Harmonisatie via ISCED betekent voor ESS-NL dat alle verschillen tussen lbo en mavo, tussen havo, mbo en vwo en tussen hbo en wo wegvallen. De oorspronkelijke variabele bevat dus meer informatie en heeft daardoor een hogere meetkwaliteit.

Voor de respondent wordt opleidingsniveau in de ESS dubbel gemeten. Er wordt namelijk ook gevraagd naar het aantal jaren dat men in opleiding heeft doorgebracht.

De meting van het opleidingsniveau van partners en ouders zit in de ESS anders in elkaar dan die van de respondent. Hier wordt er volstaan met één meting en daarenboven is helaas niet ESS-NL, maar alleen ISCED opgenomen in het databestand. De oorspronkelijke variabelen die zijn gebruikt om ISCED te coderen zijn niet beschikbaar, waardoor we de optimale schaling alleen voor de respondent kunnen toepassen.

Het International Social Survey Programme (ISSP)

Ook het ISSP kent een dubbele meting van opleidingsniveau voor de respondent. Net als in de ESS wordt er zowel naar het aantal jaren gevraagd als naar de hoogst behaalde kwalificatie. De Nederlandse ISSP-classificatie is met 9 categorieën minder gedetailleerd dan die van de ESS en de niveaus zijn bij de presentatie aan de respondent ook iets anders geordend. Gelukkig is het wel zo dat dezelfde kwalificaties bij elkaar zijn gegroepeerd. Op één uitzondering na: de programma's van de bbl (beroepsbegeleidende leerweg) en bol (beroepsopleidende leerweg). Die worden in ESS-NL gezien als zijnde van eenzelfde niveau van mbo, terwijl ze in ISSP-NL beide bij het kort middelbaar beroepsonderwijs (kmbo) geplaatst zijn. Dit is wel een probleem, maar afgezien daarvan zijn de twee classificaties compatibel en kunnen er gemakkelijk ESS-ISLED-scores aan de overeenkomstige ISSP-NL niveaus toegekend worden. Tabel 1 toont de twee classificaties naast elkaar.

Tabel 1 *Het Nederlandse opleidingssysteem in de ESS en in de ISSP*

ESS-NL		ISSP-NL	
1	lo-		
2	lo	1	Lager onderwijs
3	vmbo	2	lbo, hhs, lts, vmbo-b, lhno, vbo
4	mavo	3	mavo, ulo, mulo, vmbo-t
5	kmbo	6	kmbo, leerlingwezen, bbl-bol
6	mbo		
7	mbo+	7	mbo
8	havo	4	havo, mms
9	vwo	5	vwo
10	hbo	8	hbo
11	wo		
12	wo+	9	wo
13	dr		

Verder kunnen we bij de ISSP gebruikmaken van het feit dat er niet alleen voor de respondent maar ook voor partner en ouders gedetailleerde opleidingsvariabelen beschikbaar zijn. Afgezien daarvan verkeren we voor Nederland in de gelukkige omstandigheid dat ook het opleidingsniveau van de partner dubbel gemeten is, namelijk via een duurmaat. Dit stelt ons in staat om niet alleen voor random, maar ook voor systematische meetfout te corrigeren.

Onderzoeksontwerp en methode

Het genereren van ISLED

Ten zichte van onze eerdere analyses (Ganzeboom & Schröder 2009) zijn er wat betreft het maken van een ISLED-schaal twee dingen veranderd. Ten eerste zijn inmiddels de ESS-data van ronde 4 beschikbaar, waarvan we dankbaar gebruik hebben gemaakt. De nieuwe ISLED-schaal is dus met vier in plaats van drie rondes ESS (2002, 2004, 2006, 2008) gegenereerd. Ten tweede hebben we onze optimale schaal voorzien van een naam: ISLED (*International Standard Level of Education*) en van een metriek (meeteenheid).

Net zoals bij onze eerdere analyses brengen we de gedetailleerde opleidingsgegevens van de respondent in verband met input- en outputvariabelen. We passen ons eerder gebruikte algoritme toe.

- Kies criteriumvariabelen die een direct effect hebben op, respectievelijk direct beïnvloed worden door opleidingskwalificaties. Als inputvariabelen kiezen we ouderlijke beroepen en opleidingen en als outputvariabelen beroep van de respondent en opleiding van de partner. De beroepen zijn geschaald naar sociaal-economische status (ISEI), de opleidingen als een eenvoudige lineaire schaling van de ISCED categorieën.

- Construeer een additieve indexvariabele voor respectievelijk alle inputs en outputs, door de betreffende z-gestandaardiseerde indicatoren ongewogen te middelen (correlatie tussen inputs en outputs: 0,85).
 - Regresseer outputs op opleidingsdummies. Behoud verwachte waarden.
 - Regresseer inputs op opleidingsdummies. Behoud verwachte waarden.
 - Vorm optimale schaling uit een zodanige weging van de twee sets verwachte waarden dat:
 - de invloed van inputs op geschaalde opleidingsniveaus maximaal is,
 - de invloed van geschaalde opleidingsniveaus op outputs maximaal is,
 - de directe invloed van inputs op outputs minimaal is.
 - De gevonden optimale gewichten zijn 0,61 voor inputs en 0,39 voor outputs.
- Let wel: de gewichten zijn in alle landen hetzelfde, terwijl de inputs en de outputs (en de correlatiestructuur) wel verschillen.

Optimaal schalen op deze manier leidt wel tot de beste meting, maar nog niet tot een tussen landen vergelijkbare meting. Hiervoor hebben we een metriek nodig. De metriek hebben we als volgt verkregen:

- We definiëren een vergelijkende metriek door het gemiddelde en de spreiding tussen de optimale schaal (in een over het gehele databestand gestandaardiseerde Z-term) gelijk te maken aan gemiddelde en standaarddeviatie van de duurmaat.
- Vervolgens projecteren we terug in een 0-100 metriek door een logistische transformatie toe te passen: $ISLED = 100 * (\exp(Z) / (1 + \exp(Z)))$.

Gemiddelde en spreiding van de ISLED zijn hierdoor evenredig met die van de duurmaat, maar worden uitgedrukt in een metriek van 0 tot 100.

Het valideren van ISLED

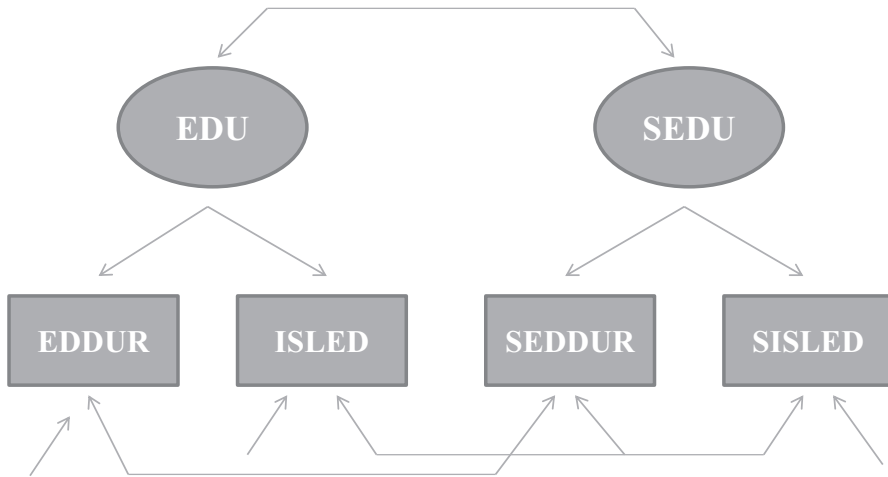
Eenmaal gemaakt is het zaak om ISLED aan nadere validaties te onderwerpen. In onze eerdere bijdrage hebben wij validatieanalyses met de ESS-data laten zien. Aangezien dit dezelfde data waren als die waarmee de ISLED is gegenereerd, onderwerpen we de ISLED deze keer aan een strengere toets door de ESS-ISLED toe te passen op verse data. We gebruiken hiervoor zes rondes van de ISSP-NL data (2002-2008).

Net als alle andere variabelen zijn ook sociale achtergrondvariabelen vatbaar voor meetfouten. De klassieke meettheorie biedt ons werktuigen om deze meetfouten te diagnosticeren en te corrigeren (*disattenuation*). We gebruiken hiervoor multiple indicatoren (duur en ISLED), die we beschouwen als congenerieke metingen van hetzelfde onderliggende concept. Een simultaan vergelijkingenmodel (enkele gemeenschappelijke factor) diagnosticeert en corrigeert random meetfouten.

Op grond van de aanwezigheid in de ISSP-data van dubbele opleidingsmetingen voor niet alleen respondent, maar ook partner, kunnen wij niet alleen random maar ook systematische meetfouten opsporen en corrigeren. Van systematische meetfouten is sprake indien respondenten telkens dezelfde fout maken bij het representeren van een opleidingsniveau. De oorzaak kan liggen in een verkeerde vraagstelling. Een belangrijk type systematische meetfout is gecorreleerde meetfout, i.e. fouten die elke

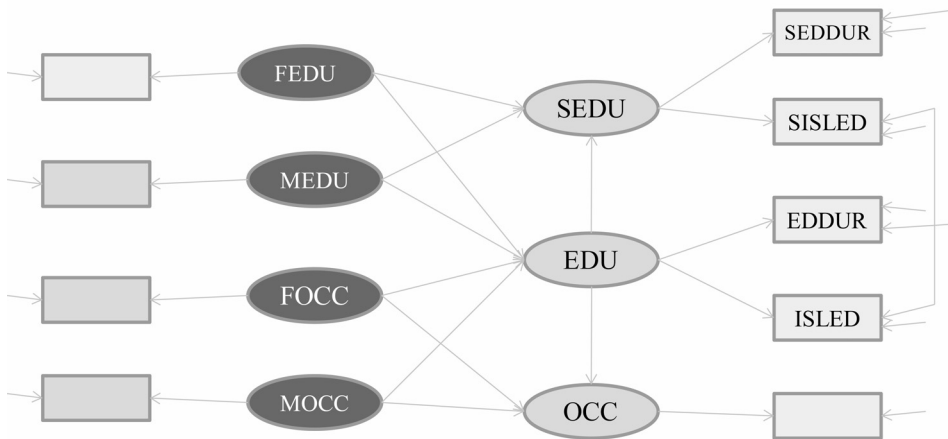
keer te zien zijn wanneer we opleidingsniveau met eenzelfde vraagstelling meten. Correctie voor systematische meetfouten is mogelijk in een MTMM-model (*Multiple Trait, Multiple Methods*), voor het eerst gepresenteerd door psychometrici in de jaren vijftig, maar tegenwoordig vooral toegepast in attitudeonderzoek om responsstijlen te modelleren. Random meetfouten kunnen we opsporen door metingen te herhalen (*multiple indicators*) en systematische meetfouten door de fout zelf te herhalen (*multiple traits*). MTMM combineert de twee methoden. Wij herhalen hetzelfde meetmodel voor de opleiding van verschillende personen, in dit geval de respondent en diens partner. Voor beiden hebben we zowel de duur en ISLED (als ook ISCED) ter beschikking.

Figuur 2 MTMM-model



Figuur 2 geeft ons basis-MTMM-model schematisch weer. De twee latente opleidingsvariabelen (ovalen) voor respondent (EDU) en partner (SEDU) worden telkens met twee indicatoren (rechthoeken) gemeten: een duurmaat en ISLED. De losstaande pijlen naar de gemeten variabelen vertegenwoordigen hun random meetfouten. De verbonden pijlen tussen de twee duurmaten (EDDUR en SEDDUR) en tussen de twee ISLEDs (ISLED en SISLED) representeren correlaties tussen hun residuen. Als zodanig is dit model niet geïdentificeerd. Dat wordt het wel als we het uitbreiden met andere criteriumvariabelen. Dit zijn beroep van de respondent (OCC), vaders en moeders beroepen (FOCC, MOCC) en hun opleidingen (FEDUC, MEDUC). Het volledige model is te zien in figuur 3. Alleen de gemeten opleidingsvariabelen hebben hier een naam, omdat zij dubbel gemeten zijn en belangrijk zijn voor het meetmodel.

Figuur 3 Het volledige SEM-model



Analyse

Het genereren van ISLED

Het eerste deel van de analyse is een replicatie van onze vorige analyses, alleen nu met vier in plaats van drie rondes ESS. Tabel 2 laat de oude en de nieuwe gestandaardiseerde schaalwaarden zien (kolommen 1 en 2). De nieuwe waarden lijken sterk op die van de eerder gerapporteerde analyse en geven eenzelfde volgorde weer van de Nederlandse opleidingsniveaus. Deze volgorde komt grotendeels overeen met de volgorde die in de ESS aan de respondenten is gepresenteerd (de volgorde van ISSP-NL wijkt hiervan af). De enige uitzondering is dat in de ESS-typologie havo volgt op mbo+ (vet gedrukt), terwijl deze volgorde in onze analyses omdraait en havo juist minder waard blijkt te zijn dan mbo+.

In de derde kolom van tabel 2 zien we de nieuwe schaalwaarden getransformeerd in de ISLED-metrick. Voorzien van een metrick zijn ze nu ook bruikbaar voor vergelijkend onderzoek en kunnen we er een interpretatie aan toekennen. Aangezien de ISLED-metrick gebaseerd is op de gehele ESS (35 Europese landen), kunnen we aan de waarden aflezen waar zich een bepaald opleidingsniveau bevindt in de Europese opleidingshiërarchie. De waarden geven globaal (namelijk onder veronderstelling van normaliteit) aan hoeveel procent van de Europese bevolking boven dan wel beneden een bepaald opleidingsniveau zit. Voor het laagste Nederlandse opleidingsniveau, niet voltooid basisonderwijs (lo-), betekent dit dat er 21% van de Europese bevolking een opleiding heeft die minder waard is op arbeids- en huwelijksmarkt en in termen van competitie tussen sociale groeperingen. Voor het hoogste opleidingsniveau, promotieonderzoek, betekent dit dat toch nog 7% van de Europese bevolking een opleiding heeft die meer waard is.

Tabel 2 *Schalings van ESS-NL-opleidingsniveaus met drie respectievelijk vier rondes ESS*

	Oud (3 rondes ESS)	Nieuw (4 rondes ESS)	ISLED
lo-	-2,221	-1,8075	21,1
lo	-1,627	-1,5265	25,8
vmbo	-1,228	-1,1861	31,8
mavo	-0,572	-0,4779	45,9
kmbo	-0,422	-0,4244	47,2
mbo	-0,196	-0,1411	53,0
mbo+	0,312	0,3828	63,9
havo	0,132	0,2459	61,2
vwo	0,668	0,6749	69,4
hbo	1,040	1,0105	75,0
wo	1,681	1,6693	84,1
wo+	2,160	2,0270	87,8
dr	2,416	2,7567	92,9

Het valideren van ISLED

Tabellen 3a en 3b laten meetmodellen zien. In een simultaan vergelijkingenmodel hebben wij de ISLED vergeleken met zowel ISCED als ook de duurmaat. In tabel 3a treffen we de meetcoëfficiënten voor de ESS-data (rondes 1 t/m 4) aan. We zien dat de duurmaat de slechtste meting is. Als we alleen deze maat gebruiken, verliezen we bijna 22% van de informatie (1-0,782). Het gebruik van ISCED vermindert dit informatieverlies behoorlijk. Nu verliezen we nog maar 10% (1-0,899). Vervangen we ISCED door ISLED, kunnen we het informatieverlies met nog eens 3% beperken. Maar 7% informatieverlies betekent dat ook ISLED geen perfecte meting is.

Om de kritiek te beantwoorden dat ISLED alleen met ESS-data zou werken omdat die daarmee is gemaakt, passen we ISLED-NL vervolgens toe op ISSP-data. Tabel 3b toont de uitkomsten van vier verschillende analyses. Deze analyses verschillen van elkaar op twee dimensies, namelijk hoe de ouderlijke opleidingen gemeten zijn (links: ISCED, rechts: ISLED) en of er wel of niet gecorrigeerd is voor systematische meetfouten (boven: niet, beneden: wel). Deze verfijningen waren met de ESS-data niet mogelijk, omdat hier de opleiding van de ouders alleen maar in ISCED-formaat beschikbaar is en omdat we daar niet over een duurmeting voor de partner beschikken. Het eerste dat in deze tabel opvalt, is dat de ISLED wederom in alle analyses verreweg het beste presteert. Opvallend is dat de coëfficiënten nauwelijks van elkaar verschillen (alleen in het derde decimaal). Wel doet deze keer de duurmaat het iets beter dan ISCED. Als we alleen met ISCED meten, verliezen we 20% informatie, met de duurmaat 18%. Bij gebruik van ISLED daalt dat verlies naar 8%, een behoorlijke winst. Meten we ook de ouderlijke opleidingen in ISLED, dan wordt het nog eens 0,2% minder.

Als we bovendien voor systematische meetfout corrigeren, behalen we extra winst van nog eens 1,7%. Deze winst is toe te schrijven aan het feit dat er wel systematische meetfouten in de duurmaat blijken te zitten, maar niet in ISLED. Maar let op: correctie

Tabel 3a Meetmodel ESS (3 opleidingsindicatoren)

Indicator	Meetcoëfficiënt
Duur	0,782
ISCED	0,899
ISLED	0,929

**Tabel 3b Meetmodellen ISSP (3 opleidingsindicatoren)
Ouderlijke opleidingen in ISCED resp. ISLED**

	Meetcoëfficiënt	
	ISCED	ISLED
Zonder correctie voor systematische meetfout		
Duur	0,819	0,817
ISCED	0,801	0,801
ISLED	0,921	0,923
Met correctie voor systematische meetfout		
Duur	0,803	0,801
ISCED	0,814	0,813
ISLED	0,938	0,940

voor een systematische meetfout in duur leidt tot een zwakkere meetcoëfficiënt voor die variabele, die nu toch, net als in de ESS, een slechtere meting blijkt dan ISCED.

Maar wat maakt het nou uit hoe goed we opleiding meten? Tabel 4 geeft daar uitsluitend over. Er zijn twee manieren om naar deze tabel te kijken. In de eerste plaats kunnen we de verklaarde variantie in de afhankelijke variabelen vergelijken. We stellen vast dat die voor alle drie de afhankelijke variabelen, opleiding respondent, opleiding partner en beroep respondent de verklaarde variantie voor ISLED aanzienlijk hoger is dan voor ISCED of duur en wel gemiddeld 6% voor opleiding respondent, 4% voor opleiding partner en maar liefst 12% voor beroep respondent. In de tweede plaats kunnen we de regressiecoëfficiënten vergelijken. We zien dat de directe effecten van inputs op outputs, bijvoorbeeld vaders en moeders opleiding (FEDUC/MEDUC) op opleiding partner gemiddeld kleiner worden naarmate we beter meten, terwijl de indirecte effecten, bijvoorbeeld van opleiding respondent (EDUC) op beroep of opleiding partner juist groter worden. Dit heeft ook inhoudelijk relevante gevolgen. Zo blijkt het effect van moeders beroep (MOCC) op de opleiding van de respondent niet meer significant te zijn als we opleidingsniveau adequaat meten.

Maar ook al levert ISLED veel winst op vergeleken bij ISCED en de duurmaat, we verliezen in het beste geval nog steeds 6% van de informatie. Door duur en geschaalde kwalificaties in één model te combineren, kunnen we ook dit verlies ondergaan en bereiken we onvertekende regressiecoëfficiënten en navenant hogere percentages verklaarde variantie in de afhankelijke variabelen. Een en ander is te zien in kolom 4 van tabel 4. Dit is het best mogelijke model, met ISLED-metingen voor de ouders en gecorrigeerd voor zowel random als systematische meetfout. De verklaarde

Tabel 4 *Modelparameters voor Nederland ISSP 2003-2008 (N = 4913)*
Enkelvoudige metingen van opleidingsniveau voor vader en moeder in
ISCED (model 1 en 2) resp. ISLED (model 3 en 4)

	1 ISCED	2 EDDUR	3 ISLED	4 EDDUR & ISLED
A. Structural models				
OPLEIDING RESPONDENT				
FEDUC	0,188	0,187	0,172	0,228
MEDUC	0,095	0,125	0,138	0,125
FOCC	0,087	0,117	0,150	0,117
MOCC	0,077	0,044#	0,074	0,093
R2	0,147	0,166	0,214	0,251
OPLEIDING PARTNER				
FEDUC	0,132	0,089	0,120	0,106
MEDUC	0,069	0,093	0,065	0,054
EDUC	0,467	0,509	0,514	0,591
R2	0,314	0,351	0,370	0,459
BEROEP RESPONDENT				
FOCC	0,196	0,182	0,134	0,078
MOCC	0,050	0,053	-0,008#	-0,017#
EDUC	0,547	0,540	0,686	0,762
R2	0,441	0,427	0,558	0,628
B. Fit statistics				
FIML Chi-square	78,165	76,294	91,303	114,169
RMSEA	0,0214	0,0211	0,0239	0,0206

Gestandaardiseerde coëfficiënten. # = niet significant.

variantie neemt toe voor alle drie de afhankelijke variabelen: met 3,5% voor opleiding respondent, bijna 9% voor opleiding partner en 7% voor het beroep van de respondent. Ook de regressiecoëfficiënten veranderen weer op dezelfde, voorspelbare manier: gemiddeld nemen de indirecte effecten toe en de directe af.

Als we nu de zwakste en de sterkste metingen naast elkaar leggen, dan komen we tot de opzienbarende conclusie dat we van opleiding respondent 10%, van opleiding partner 14,5% en van beroep respondent 18,7% meer variantie kunnen verklaren als we voor alle meetfouten corrigeren en alle opleidingsvariabelen optimaal meten. Tussen de regressiecoëfficiënten zien we overeenkomstige grote discrepanties.

Conclusie

De ISLED-NL blijkt heel goed bruikbaar te zijn ook voor ISSP-data. De ISLED-schaling verbetert ook hier merkbaar de meetkwaliteit van de opleidingsvariabele. De ISLED is gedetailleerder dan de ISCED en bevat minder random meetfout dan zowel de ISCED

en de duurmaat. Bovendien blijkt de ISLED in tegenstelling tot de duurmaat geen systematische meetfouten te bevatten.

Maar, net als bij de ESS-data, geldt ook voor de ISSP-data dat een combinatie van twee metingen maximale meetkwaliteit oplevert. ISLED presteert weliswaar aanzienlijk beter dan ISCED en duur, maar is niet perfect en we houden wel een verlies van 6% van de informatie over. Het feit dat de ISSP-data niet alleen voor de respondent maar ook voor de partner dubbele opleidingsmetingen bevatten, maakt het mogelijk om deze imperfectie volledig te compenseren. Door in een MTMM-model niet alleen voor random maar ook voor systematische meetfouten te corrigeren, bereiken wij een onvertekende meting die haar weerslag krijgt in hogere regressiecoëfficiënten en hogere percentages verklaarde variantie in de afhankelijke variabelen.

Literatuur

- Braun, M. en W. Müller (1997). Measurement of Education in Comparative Research, *Comparative Social Research*, 16: 163-201.
- Ganzeboom, H.B.G. en H. Schröder (2009). De waarde van diploma's: een kwantificatie van de ESS-NL-categorieën. In: H.B.G. Ganzeboom en M. Wittenberg (red.) *Nederland in vergelijkend perspectief, Proceedings Tweede Nederlandse Workshop European Social Survey*. Amsterdam: Aksant.
- Grusky, D.B. en S.E. Van Rompaey (1992). The Vertical Scaling of Occupations: Some Cautionary Comments and Reflections, *American Journal of Sociology*, 97 (6): 1712-1728.
- Hirsch, F. (1976). *Social Limits to Growth*. London: Routledge & Kegan Paul.
- Hout, M. en T.A. DiPrete (2006). What we Have Learned: RC28's Contributions to Knowledge about Social Stratification, *Research in Social Stratification and Mobility*, 24 (1): 1-20.
- ISCED, http://www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm
- Kelley, J. (1973). Causal Chain Models for the Socioeconomic Career, *American Sociological Review*, 38: 481-493.
- Kerckhoff, A.C. en M. Dylan (1999). Problems with International Measures of Education, *The Journal of Socio-Economics*, 28: 759-775.
- Schneider, S. (2007). *Measuring Educational Attainment in Cross-National Surveys: The Case of the European Social Survey*, paper presented at the EDUC workshop of the EQUAL-SOC network, Dijon, 22-24 Nov.
- Schneider, S. (2008). *Nominal Comparability is not Enough: Evaluating Cross-national Measures of Educational Attainment Using ISEI Scores*, Sociology Working Papers.
- Schröder, H. (2008). *Scaling the Grades*, unpublished manuscript, Vrije Universiteit Amsterdam.
- Smith, H.L. en M.A. Garnier (1987). Scaling via Models for the Analysis of Association: Social Background and Educational Careers in France, *Sociological Methodology*, 17: 205-245.
- Thurow, L.C. (1975). *Generating Inequality*. London: The Macmillan Press Ltd.
- Treiman, D.J. en K. Terrell (1975). The Process of Status Attainment in the United States and Great Britain, *American Journal of Sociology*, 81: 563-583.

Appendix Het Nederlandse onderwijssysteem in de ESS

ESS		ISCED
lo-	niet voltooid lager onderwijs	0
lo	lager onderwijs (lo), basisschool, lager speciaal onderwijs	1
vmbo	lager beroepsonderwijs (lbo), lagere technische school (lts), nijverheidsonderwijs, huishoudschool, voorbereidend middelbaar beroepsonderwijs (vmbo): basisberoepsgerichte of kaderberoepsgerichte leerweg	2
mavo	middelbaar algemeen voortgezet onderwijs (mavo), (meer) uitgebreid lager onderwijs (mulo, ulo), voorbereidend middelbaar beroepsonderwijs (vmbo): theoretische of gemengde leerweg	2
kmbo	kort middelbaar beroepsonderwijs (kmbo)	3
mbo	middelbaar beroepsonderwijs (mbo), beroepsopleidende leerweg (bol), beroepsbegeleidende leerweg (bbl)	3
mbo+	mbo-plus voor toegang tot het hbo, korte hbo-opleiding (korter dan 2 jaar)	4
havo	hoger algemeen voortgezet onderwijs (havo), vooropleiding hoger beroepsonderwijs (vhbo), middelbare meisjesschool (mms)	3
vwo	vorbereidend wetenschappelijk onderwijs (vwo), gymnasium, atheneum, hogere burgerschool (hbs)	3
hbo	hoger beroepsonderwijs (hbo), kweekschool, conservatorium, mo-akten, nieuwe stijl hogescholen	5
wo	wetenschappelijk onderwijs (wo), universiteit, technische/economische hogeschool oude stijl	5
wo+	postdoctorale opleiding (leraren- en beroepsopleidingen zoals medici, apotheker)	6
dr	aio/oio of andere promotieopleiding tot graad van doctor	6

NB: In ESS is de ISCED indeling gerepresenteerd door de variabele EDULVL/EDULVLA.