

Tijdreeksen werkzame beroepsbevolking naar beroepsprestige, sociaal- economische status en sociale klasse

Paul den Boer

Publicatiedatum CBS-website: 2 december 2010



Verklaring van tekens

.	= gegevens ontbreken
*	= voorlopig cijfer
**	= nader voorlopig cijfer
x	= geheim
–	= nihil
–	= (indien voorkomend tussen twee getallen) tot en met
0 (0,0)	= het getal is kleiner dan de helft van de gekozen eenheid
niets (blank)	= een cijfer kan op logische gronden niet voorkomen
2008–2009	= 2008 tot en met 2009
2008/2009	= het gemiddelde over de jaren 2008 tot en met 2009
2008/'09	= oogstjaar, boekjaar, schooljaar enz., beginnend in 2008 en eindigend in 2009
2006/'07–2008/'09	= oogstjaar, boekjaar enz., 2006/'07 tot en met 2008/'09

In geval van afronding kan het voorkomen dat het weergegeven totaal niet overeenstemt met de som van de getallen.

Colofon

Uitgever

Centraal Bureau voor de Statistiek
Henri Faasdreef 312
2492 JP Den Haag

Prepress

Centraal Bureau voor de Statistiek - Grafimedia

Omslag

TelDesign, Rotterdam

Inlichtingen

Tel. (088) 570 70 70
Fax (070) 337 59 94
Via contactformulier: www.cbs.nl/infoservice

Bestellingen

E-mail: verkoop@cbs.nl
Fax (045) 570 62 68

Internet

www.cbs.nl

Samenvatting

Deze nota beschrijft de methode die is gebruikt voor het maken van de tijdreeksen werkzame beroepsbevolking naar drie maten van sociaal-economische status. Deze methode is gebaseerd op het artikel van Bart Bakker et al. (1997) dat beschrijft hoe maten voor beroepsprestige (U&S-92), sociaal-economische status (ISEI-92) en sociale klasse (EGP-92) kunnen worden afgeleid uit de Standaard Beroepen Classificatie 1992 (SBC-92). Het ISMF-project (International Stratification and Mobility File) waaraan Harry Ganzeboom en Donald Treiman (1993) hebben gewerkt, levert tools om beroep en opleidingsniveau te hercoderen naar internationaal en historisch vergelijkbare maten. Deze tools zijn gebruikt om de U&S-92, ISEI-92 en EGP-92 af te leiden uit oudere beroepenclassificaties, waaronder Beroepen Classificatie 1984 (BC-84) en International Standard Classification of Occupations 1968 (ISCO-68). Uit de Arbeidskrachtentelling (AKT) 1975–1985 en de Enquête Beroepsbevolking (EBB) 1987–2007 zijn tijdreeksen 1975–2007 afgeleid van U&S-92, ISEI-92 en EGP-92.

Beroepsprestige (U&S-92) en sociaal-economische status (ISEI-92)

De U&S-92 en ISEI-92 zijn afgeleid uit het beroep. In de tijdreeksen 1975–2007 van U&S-92 en ISEI-92 zitten elk drie trendbreuken, zodat telkens vier deelreeksen kunnen worden onderscheiden (1975–1977, 1981–1985, 1987–2003, 2004–2007). De trendbreuken zijn veroorzaakt door gebruik van verschillende bronnen (AKT, EBB) en door gebruik te maken van verschillende afleidingsschema's (Bakker et al. 1997; Ganzeboom & Treiman 1993). De trendbreuken zijn gerepareerd met de methode van stapsgewijze regressie. Hierbij zijn zogenaamde dummyvariabelen (waarde 1 of 0) als exogene variabelen ingezet, om per deelreeks te bepalen welk deel van de trendbreuk toegeschreven kan worden aan het methode-effect. Verder zijn in het regressiemodel nog een aantal exogene variabelen meegenomen die volgens de geraadpleegde literatuur (De Graaf & Luijkx 1995; Van der Lippe & Van Doorne-Huiskens 1995; Ganzeboom & Luijkx 1995) samenhangen met beroepsprestige en sociale status. Deze variabelen zijn het aandeel van personen met hoger opleidingsniveau (hbo, wo) in de totale werkzame beroepsbevolking, het aandeel vrouwen in de totale werkzame beroepsbevolking, en het aandeel van de dienstverlening (inclusief overheid) in de totale economie. De regressie geeft een goede modelfit, zodat het goed mogelijk is de trendbreuken te repareren en consistente tijdreeksen 1975–2007 te maken van U&S-92 en ISEI-92.

Sociale klasse (EGP-92)

1975–1985

De EGP-92 is afgeleid uit het beroep, de positie in de werkring, en het aantal personen waarover leiding wordt gegeven. In 1981–1985 kan de EGP-92 op twee manieren worden afgeleid. De eerste manier is door een koppeling (via het persoons-identificatienummer) tussen het opgehoogde AKT-bestand (variabelen beroep, positie in de werkring) en de op het Data Service Center gearhiveerde originele AKT-steekproef (variabele aantal personen waarover leiding wordt gegeven). De tweede manier is EGP-92 direct af te leiden uit deze originele AKT-steekproef (alle benodigde variabelen). Gekozen is voor de tweede manier, omdat EGP-92 dan vaker is af te leiden, en omdat bij de eerste manier een trendbreuk zou ontstaan tussen 1981 en 1983, wat vreemd is gezien het feit dat beide jaren uit dezelfde bron en via hetzelfde afleidingsschema zijn afgeleid.

1987–2007

In het bestand EBB 1987–1995 ontbreekt het aantal personen waarover leiding wordt gegeven. Maar in de jaarbestanden EBB 1992–1995 zitten de variabele die aangeeft of iemand leiding geeft en het aantal personen waarover leiding wordt gegeven. Daarom is het bestand EBB 1987–1995 gekoppeld (via adresnummer, huishoudnummer en

persoonsnummer) aan de jaarbestanden EBB 1992–1995. Ook in de periode 1996–2007 missen soms de variabele die aangeeft of iemand leidinggeeft en het aantal personen waarover leiding wordt gegeven.

Deze variabelen zijn hercodeerd als dichotome variabelen (twee waarden: leiding = ja of nee, Nleiding = geeft leiding aan 1–9 personen of aan minstens 10 personen). Met behulp van de imputatiemethode van logistische regressie zijn deze variabelen bijgeschat. Omdat in de jaren 1987–1991 het aantal personen waarover leiding wordt gegeven nooit bekend is, is de oplossing van het logistische regressiemodel in de periode 1992–1995 ook ingezet in de jaren 1987–1991. Dit is gedaan onder de aanname dat de periode 1992–1995 representatief is voor de jaren 1987–1991. Exogene variabelen in de logistische regressie zijn geslacht, leeftijdsgroep, beroepsniveau en manager (geeft aan of iemand manager is of niet). De logistische regressie geeft een goede modelfit. Ook uit een plausibiliteitscontrole, uit de methodenreeks (Israëls, Pannekoek & Schulte Nordholt 2007), blijkt dat de imputatiemethode betrouwbaar is. Verder is aangenomen dat zelfstandigen met personeel in dienst ook leidinggeven. Daarna is EGP-92 afgeleid.

Trendbreuken

In de tijdreeksen 1975–2007 van EGP-92 zitten trendbreuken, veroorzaakt door het gebruik van verschillende bronnen en afleidingsschema's. Deze trendbreuken zijn gerepareerd met behulp van stapsgewijze regressie. In het regressiemodel zijn weer dummyvariabelen meegenomen, en verder dezelfde andere exogene variabelen als gebruikt bij de reparatie van trendbreuken in de tijdreeksen U&S-92 en ISEI-92. De regressie geeft een goede modelfit. Tot slot zijn de zeven reeksen (zeven sociale klassen) consistent gemaakt met behulp van een Lagrange-functie, waarbij ook rekening is gehouden met de kwaliteit van de schatting van de gerepareerde reeksen (Kardal & Van der Stegen 2009). Zo tellen de zeven reeksen met het aandeel in de werkzame beroepsbevolking naar sociale klasse weer op tot 100 procent.

Inhoudsopgave

1. Inleiding	6
2. Benodigde variabelen in de AKT en EBB	7
2.1 AKT	7
2.1.1 1975–1979	7
2.1.2 1981–1985	7
2.2 EBB	7
2.2.1 1987–1995	8
2.2.2 1996–2007	8
2.3 Non-respons van benodigde variabelen	8
3. Afleidingsschema's	10
3.1 1975–1985	10
3.2 1987–2007	10
3.2.1 Afleidingsschema	10
3.2.2 Aanpassingen in setup	11
4. Beroepsprestige (U&S-92) en sociaal-economische status (ISEI-92)	13
4.1 Afleiden van U&S-92 en ISEI-92	13
4.1.1 1975–1985	13
4.1.2 1987–2007	13
4.2 Trendbreuken	13
4.3 Reparatie van trendbreuken	14
4.3.1 Dummyvariabelen	14
4.3.2 Andere exogene variabelen uit de literatuur	15
4.3.3 Regressiemodel	15
4.3.4 Methode-effect	15
4.3.5 Gerepareerde tijdreeksen U&S-92 en ISEI-92	16
5. Sociale klasse (EGP-92)	17
5.1 1975–1979	17
5.2 1981–1985	17
5.2.1 Manier 1: EGP-92 afleiden uit koppeling AKT en DSC AKT	17
5.2.2 Manier 2: EGP-92 afleiden uit DSC AKT	18
5.2.3 Vergelijking tussen manier 1 en 2	18
5.3 1987–2007	19
5.3.1 EGP-92 afleiden uit beschikbare data	19
5.3.2 Koppelen van bestand EBB 1987–1995 en jaarbestanden EBB 1992–1995	19
5.3.3 Analyse van leidinggeven naar achtergrondvariabelen	20
5.3.4 Imputeren van leidinggeven (logistische regressie)	21
5.3.5 Plausibiliteitscontrole	22
5.3.6 EGP-92 afleiden uit beschikbare en geïmputeerde data	23
5.4 Trendbreuken	23
5.5 Reparatie van trendbreuken	23
5.5.1 Regressiemodel	23
5.5.2 Consistentie van de gerepareerde reeksen	24
6. Resultaten van beroepsprestige (U&S-92), sociaal-economische status (ISEI-92) en sociale klasse (EGP-92)	26
7. Referenties	27

Inleiding

De wens bestaat om een tijdreeks te maken van de werkzame beroepsbevolking naar sociaal-economische status. De werkzame beroepsbevolking zal worden afgeleid naar drie maten voor beroepsprestige, sociaal-economische status en sociale klasse:

- Ultee & Sixma-schaal voor beroepsprestige (Sixma & Ultee 1983), afgekort als U&S-92
- International Socio-Economic Index (Ganzeboom, De Graaf & Treiman 1992), afgekort als ISEI-92
- Indeling naar sociale klasse (Erikson, Goldthorpe & Portocarero 1979), afgekort als EGP-92.

Het artikel van Bart Bakker et al. (1997) beschrijft hoe deze maten voor beroepsprestige (U&S-92), sociaal-economische status (ISEI-92) en sociale klasse (EGP-92) kunnen worden afgeleid uit de Standaard Beroepen Classificatie 1992 (SBC-92). Verder levert het ISMF project (International Stratification and Mobility File), waaraan Harry Ganzeboom en Donald Treiman (1993) hebben gewerkt, tools om beroepsprestige, sociaal-economische status, en sociale klasse af te leiden uit oudere beroepenclassificaties, waaronder de Beroepen Classificatie 1984 (BC-84) en International Standard Classification of Occupations 1968 (ISCO-68).

In paragraaf 2 wordt bekeken welke benodigde variabelen voor het afleiden van U&S-92, ISEI-92 en EGP-92 beschikbaar zijn in de AKT en de EBB. In paragraaf 3 staan de gebruikte tools en setups met afleidingsschema's, en enkele aanpassingen in deze setups. Paragraaf 4 beschrijft de berekening van U&S-92 en ISEI-92. In paragraaf 5 wordt de werkzame beroepsbevolking ingedeeld naar EGP-92.

De resultaten van voor trendbreuken gerepareerde tijdreeksen van U&S-92, ISEI-92 en EGP-92 staan in paragraaf 6. Tot slot volgt de referentielijst (paragraaf 7).

2. Benodigde variabelen in de AKT en EBB

Onderzocht is welke van de benodigde variabelen beschikbaar zijn in de AKT (paragraaf 2.1) en EBB (paragraaf 2.2). Door Piet Prak zijn de AKT van 1973, 1975, 1977, 1979, 1981, 1983 en 1985 gearchiveerd in SPSS bestanden op het Data Service Center (DSC). Deze AKT-bestanden (akt73_def.sav t/m akt85_def.sav) zijn niet gecorrigeerd naar definitie CCS 1991. Voor een duidelijk onderscheid zullen deze originele AKT verder genoemd worden DSC AKT. Paragraaf 2.3 geeft een non-respons analyse van de benodigde variabelen.

2.1 AKT

2.1.1 1975–1979

In de DSC AKT 1975, 1977 en 1979 staan de volgende benodigde variabelen. Bestanden akt75_def.sav en akt77_def.sav

- Br (beroep, waarschijnlijk volgens Beroepen Classificatie 1984)
- WrkPos2 (positie in de werkkring)
- LeidPers (Geeft u leiding aan personeel?)
- AprsLeid (Aan hoeveel medewerkers geeft u leiding?)

Bestand akt79_def.sav

- WrkPos2 (positie in de werkkring)

In de DSC AKT 1979 ontbreken dus het beroep en het aantal personen waarover leiding wordt gegeven.

2.1.2 1981–1985

In de DSC AKT 1981, 1983 en 1985 staan de volgende benodigde variabelen. Bestanden akt81_def.sav, akt83_def.sav en akt85_def.sav

- Br (beroep, waarschijnlijk volgens Beroepen Classificatie 1984)
- WrkPos(4) (positie in werkkring)
- AprsLeid (Aan hoeveel medewerkers geeft u leiding?)

In de AKT 1981, 1983 en 1985 staan de volgende benodigde variabelen. Bestanden akt81reeksNb.sav, akt83reeksNb.sav en akt85reeksNb.sav

- br9 (beroep, waarschijnlijk volgens Beroepen Classificatie 1984)
- pos19 (positie in de werkkring)
- poswrk (positie in de werkkring)

In de AKT 1981, 1983 en 1985 is niet gedocumenteerd waar de waarden van de variabelen pos19 en poswrk voor staan. Door een vergelijking met de variabele WrkPos (positie in werkkring 1) uit de DSC AKT 1981, 1983 en 1985 kan de variabele pos19 toch gecodeerd kan worden (1=zelfstandig, 2=meewerkend, 3=loon-dienst). Het aantal personen waarover leiding wordt gegeven staat niet in de AKT 1981, 1983 en 1985. Deze variabele (AprsLeid) staat wel in de DSC AKT 1981, 1983 en 1985.

2.2 EBB

In de EBB 1987–2007 staan de volgende benodigde variabelen.

2.2.1 1987–1995

Bestand tijdreeks_1987–1995.sav

- brp84 (Beroepenindeling 1984)
- brp92 (StandaardBeroepenClassificatie (SBC) 1992)
- brpgr92 (SBC-92 beroepsgroep)
- brpk192 (SBC-92 beroepsklasse)
- brpni92 (SBC-92 beroepsniveau)
- posit (positie in de 1^e werkkring)

Het aantal personen waarover leiding wordt gegeven ontbreekt in EBB 1987–1995.

Jaarbestanden totjr_92.sav, totjr_93.sav, totjr_94.sav en totjr_95.sav

- beroep84 (Beroep 84)
- beroep92 (Beroep 92)
- poswrk1 (positie in de werkkring (eerste werkkring))
- leidw1 (QT_LeidingGeven)
- nleidw1 (QT_LeidingGevenAantal)

2.2.2 1996–2007

Bestanden reeks1996–2000.sav en reeks2001–2007.sav

- Beroep84 (Beroepentyping van de eerste werkkring (1984))
- Beroep92 (Beroepentyping van de eerste werkkring (1992))
- ABER92GR (Beroepsgroep (Beroepenclassificatie 1992))
- ABER92KL (Beroepsklasse (Beroepenclassificatie 1992))
- ABER92NI (Beroepsniveau (Beroepenclassificatie 1992))
- PosWrk1 (positie in de werkkring (eerste werkkring))
- LeidW1 (geeft persoon in de eerste werkkring leiding aan medewerkers)
- NLeidW1 (aantal personen waarover leiding wordt gegeven in de eerste werkkring)

2.3 Non-respons van benodigde variabelen

Omdat de werkloze beroepsbevolking en de niet-beroepsbevolking geen beroep uitoefenen, is alleen van de werkzame beroepsbevolking een non-respons analyse gedaan voor de benodigde variabelen. Frequentietellingen geven per databron het percentage van het aantal cases dat de benodigde variabelen missend of onbekend zijn (tabel 1).

Tabel 1
Non-respons van benodigde variabelen (% van aantal cases)

Databron	Beroep	Positie in de werkkring	Leidinggeven	Aantal personen waarover leiding wordt gegeven
DSC AKT 1975	1,2	0,0	0,3	0,8
DSC AKT 1977	1,5	0,0	0,3	0,6
DSC AKT 1979	100,0	0,0	100,0	100,0
DSC AKT 1981	2,0	0,0	100,0	1,2
DSC AKT 1983	0,0	0,0	100,0	1,6
DSC AKT 1985	0,0	0,0	100,0	1,7
AKT 1981	0,0	0,0	100,0	100,0
AKT 1983	0,0	0,0	100,0	100,0
AKT 1985	0,0	0,0	100,0	100,0
EBB 1987–1995	2,6	0,0	100,0	100,0
EBB 1992	51,6	38,1	54,1	87,7
EBB 1993	8,5	0,0	16,6	78,5
EBB 1994	4,9	0,0	17,3	79,7
EBB 1995	0,8	0,0	17,4	80,2
EBB 1996–2000	1,3	0,1	21,0	80,5
EBB 2001–2007	1,1	0,0	12,6	77,0

In de DSC AKT 1979 ontbreken beroep, leidinggeven en aantal personen waarover leiding wordt gegeven, zodat U&S-92, ISEI-92 en EGP-92 niet afgeleid kunnen worden. In de AKT 1981, 1983 en 1985 ontbreken leidinggeven en aantal personen waarover leiding wordt gegeven. Maar in de DSC AKT 1981, 1983 en 1985 is het aantal personen waarover leiding wordt gegeven wel beschikbaar. In de EBB 1987–1995 ontbreken de variabele leidinggeven en aantal personen waarover leiding wordt gegeven. Maar in de jaarbestanden EBB 1992–1995 zijn deze variabelen vaak wel beschikbaar. Ook in de EBB 1996–2000 en 2001–2007 zijn leidinggeven en aantal personen waarover leiding wordt gegeven vaak missend of onbekend.

3. Afleidingsschema's

Voor het afleiden van beroepsprestige (U&S-92), sociaal-economische status (ISEI-92) en sociale klasse (EGP-92) is gebruik gemaakt van afleidingsschema's, die staan beschreven in SPSS-setups. In de jaren 1975–1985 (paragraaf 3.1) zijn andere afleidingsschema's ingezet dan in de jaren 1987–2007 (paragraaf 3.2).

3.1 1975–1985

Het ISMF project (International Stratification and Mobility File) levert tools om beroep en opleidingsniveau te hercoderen naar internationaal en historisch vergelijkbare maten (Ganzeboom & Treiman, 1993). Deze tools maken het mogelijk om de maten U&S-92, ISEI-92 en EGP-92 af te leiden uit oudere beroepenclassificaties, waaronder de Beroepen Classificatie 1984 (BC-84) en de International Standard Classification of Occupations 1968 (ISCO-68). De Beroepen Classificatie 1984 is de voorloper van de Standaard Beroepen Classificatie 1992, en een enigszins aangepaste versie van de classificatie die bij de Volkstelling 1971 werd gebruikt. De syntax met labels van BC-84, de syntax om U&S-92 af te leiden uit BC-84, en de syntax voor het omzetten van BC-84 naar ISCO-68 staan op <http://users.fmg.uva.nl/vandewerfhorst/occrecode.htm>. Om in de periode 1975–1985 de U&S-92, ISEI-92 en EGP-92 af te leiden uit BC-84 of ISCO-68 is gebruik gemaakt van de volgende tools:

- Cbs84vallab.inc (labels van BC-84)
- Cbsusinc (afleiding van U&S-92 uit BC-84)
- Cbsisco.inc (omzetten van BC-84 naar ISCO-68)
- iscolab.sps (labels van ISCO-68)
- iscoisei.sps (afleiding van ISEI-92 uit ISCO-68)
- iscoegp (afleiding van EGP-92 uit ISCO-68)

3.2 1987–2007

3.2.1 Afleidingsschema

De U&S-92, ISEI-92 en EGP-92 zijn afgeleid uit de Standaard Beroepen Classificatie 1992 door gebruik te maken van de ontwikkelde afleidingsschema's (Bakker et al. 1997). De afleidingsschema's staan in de SPSS-setups Alfstra.SPS en Alfstrah.SPS (beschikbaar op de cd-rom 'Standaard beroepenclassificatie 1992 verbeterde editie 2001'). De Standaard Beroepen Classificatie 1992 bestaat uit:

- 1211 beroepen (5-cijferige SBC-92)
- 121 beroepsgroepen (3-cijferige SBC-92)
- 43 beroepsklassen (2-cijferige SBC-92)
- 5 beroepsniveaus (1-cijferige SBC-92)

De beroepen kunnen niet geaggregeerd worden naar beroepsgroepen, beroepsklassen en beroepsniveaus door de codes in te korten tot codes van 3, 2 of 1 cijfer. Het bestand Beroepsgr SBC naar klasse en niveau.txt¹⁾ beschrijft de samenhang tussen beroepsgroepen, beroepsklassen en beroepsniveaus. Het bestand SBCLABEL.TXT²⁾ bevat de betekenis van de codes van de SBC-92 (variabele sbc92).

¹⁾ Beroepsbevolking\SES\SBC\Schakelschema's\BC84_bergrSBC92.

²⁾ Beroepsbevolking\SES\AFEGP92.

De U&S-92 en ISEI-92 worden direct van de 5-cijferige SBC-92 afgeleid. EGP-92 wordt afgeleid van de 5-cijferige SBC-92, en in 20 procent van de SBC-92 codes óók van twee hulpvariabelen:

- Positie in de werkkring (pos: 1=loondienst of meewerkend gezinslid, 2=zelfstandig).
- Het aantal personen waarover leiding wordt gegeven (lep: 0=geeft geen leiding, 5=geeft leiding aan 1-9 personen, 11=geeft leiding aan 10 personen of meer).

Het is belangrijk eerst een waardecontrole (komen waarden van variabelen sbc92, pos en lep overeen met waarden die deze variabelen mogen aannemen) te doen vóór de afleiding, anders loopt het afleidingsschema stuk. De SPSS-setups voor de afleiding van U&S-92, ISEI-92 en EGP-92 geven de volgende output:

- U&S-92 (uens: 15-86, 99)
- ISEI-92 (isei92: 16-87, 99)
- EGP-92 (egp92: 01–11, 99. Naar positie in de werkkring en het aantal personen waarover leiding wordt gegeven)
- Soort arbeid (srt: 1=hoofdarbeid, 2=handarbeid en niet agrarisch, 3=handarbeid, agrarisch)
- Niveau van de benodigde bekwaamheden (niv: 0=onbekend, 1=elementair, 2=lager, 3=middelbaar, 4=hoger, 5=wetenschappelijk)
- Hoofdarbeidersberoepen waarin leiding wordt gegeven aan handarbeiders (lhand: 1=leidinggevend aan handarbeiders, 0=overig)
- Waardecontrole (controle: 0=alleen toegestane waarden in variabelen sbc92, pos, lep, srt, niv).

3.2.2 Aanpassingen in setups

Eerst zijn de variabelen brp92 (beroep SBC-92) uit EBB 1987–1995 en Beroep92 (beroep SBC-92) uit EBB 1996–2007 hercodeerd als variabele sbc92 (beroep SBC-92). De variabelen posit (positie in de eerste werkkring) uit EBB 1987–1995 en PosWrk1 (positie in de eerste werkkring) uit EBB 1996–2007 zijn hercodeerd als variabele pos (positie in de werkkring), waarbij de waarde 99 is toegekend bij missend of onbekend. Het aantal personen waarover leiding wordt gegeven staat niet in de EBB 1987–1995. Dus is in de periode 1987–1995 de variabele lep (het aantal personen waarover leiding wordt gegeven) aangemaakt met in alle cases de waarde 99 (missend of onbekend).

In de periode 1996–2007 is de variabele lep gebaseerd op de variabelen LeidW1 (geeft persoon in de eerste werkkring leiding aan medewerkers) en NLeidW1 (aantal personen waarover leiding wordt gegeven in de eerste werkkring). Wanneer de variabele lep niet kan worden afgeleid uit de variabelen LeidW1 en NLeidW1 (doordat deze missend of onbekend zijn) krijgt de variabele lep de waarde 99 (missend of onbekend).

Vervolgens bleken in de SPSS-setup voor het afleiden van U&S-92, ISEI-92 en EGP-92 (Alfstra.SPS) enkele aanpassingen nodig. Bij sbc92=99999 (onbekend) en bij missende waarden van sbc92 krijgt de variabele srt (soort arbeid) de waarde 1 en krijgen de variabelen uens92 (U&S-92) en isei92 (ISEI-92) de waarde 99 (onbekend), zoals dat ook bij sbc92=0 (onbekend) in de SPSS-setup stond. Als sbc92=99999 dan krijgt ook de variabele egp92 (EGP-92) de waarde 99. Anders zou in de SPSS-setup de hulpvariabele niv (niveau van de benodigde bekwaamheden), berekend als eerste digit van sbc92 dus niv=9, de waarde 5 (wetenschappelijk) worden toegekend, wat uiteindelijk zou resulteren in EGP-92 klasse 1 (hogere leidinggevendenden, academici). Als de variabele sbc92 missend is, hoeft de SPSS-setup niet te worden aangepast, want dan wordt berekend dat niv=99 en wordt met behulp van de variabelen pos, srt en lep de variabele egp92 bepaald. Om praktische redenen worden bij uens92=99 en isei92=99 de waarden gezien als missende waarden (missing values), zoals ook bij egp=99 in de SPSS-setup stond.

Tot slot blijkt in de EBB bestanden de variabele voor het beroep SBC-92 (brp92, Beroep92) soms een waarde te hebben die niet voorkomt in de Standaard Beroepenclassificatie 1992. Het gaat om de waarden 67200, 86300, 86400, 86500, 86600, 87200. Daarvoor zijn de waarden van de variabelen srt, uens92 en isei92 genomen die gekoppeld zijn aan de variabele sbc92 waarvan de eerste vier digits overeenkomen, maar waarvan de laatste digit niet gelijk aan 0 is. Bijvoorbeeld: voor sbc92=67200 is gekeken naar sbc92 met waar-

den 67201, 67202 en 67203. Voor de SBC-92 beroepen 67201, 67202 en 67203 geldt: srt=1, uens92=63, isei92=55. De SBC-92 beroepen die genomen worden hebben steeds eenzelfde waarde voor de variabelen srt, uens92 en isei92 (tabel 2).

Tabel 2
Variabele brp92/Beroep92 (EBB bestand) met niet bestaande waarde gekoppeld aan soort arbeid

brp92, Beroep92 (EBB bestand)	sbc92	srt	uens92	isei92
<i>schaalscore</i>				
67200	67201	1	63	55
	67202	1	63	55
	67203	1	63	55
86300	86301	1	76	69
	86302	1	76	69
	86303	1	76	69
	86304	1	76	69
	86305	1	76	69
	86306	1	76	69
	86307	1	76	69
	86308	1	76	69
	86309	1	76	69
86400	86401	1	76	65
	86402	1	76	65
	86403	1	76	65
86500	86501	1	76	68
	86502	1	76	68
	86503	1	76	68
	86504	1	76	68
	86505	1	76	68
86600	86601	1	76	69
	86602	1	76	69
	86603	1	76	69
	86604	1	76	69
	86605	1	76	69
	86606	1	76	69
87200	87201	1	76	65
	87202	1	76	65
	87203	1	76	65
	87204	1	76	65
	87205	1	76	65
	87206	1	76	65
	87207	1	76	65
	87208	1	76	65

4. Beroepsprestige (U&S-92) en sociaal-economische status (ISEI-92)

Met behulp van de in paragraaf 3 genoemde tools en setups zijn het beroepsprestige (U&S-92) en de sociaal-economische status (ISEI-92) afgeleid (paragraaf 4.1). Na afleiding blijken in de tijdreeksen trendbreuken te zitten. Paragraaf 4.2 geeft een analyse van deze trendbreuken. In paragraaf 4.3 staat beschreven hoe de trendbreuken worden gerepareerd.

4.1 Afleiden van U&S-92 en ISEI-92

4.1.1 1975–1985

De U&S-92 is afgeleid uit de Beroepen Classificatie 1984 (BC-84). De ISEI-92 is afgeleid uit de International Standard Classification of Occupations 1968 (ISCO-68). Daarvoor was het eerst nodig BC-84 om te zetten naar ISCO-68. In de periode 1975–1985 kunnen U&S-92 en ISEI-92 bijna altijd (tenminste 97,0 procent van de cases) afgeleid worden. De reden dat U&S-92 niet altijd kan worden afgeleid is dat het beroep soms onbekend is. Dit geldt ook voor ISEI-92, bovendien kan beroep BC-84 soms niet omgezet worden naar beroep ISCO-68. In 1979 kunnen U&S-92 en ISEI-92 nooit afgeleid worden, want in de DSC AKT 1979 ontbreekt het beroep.

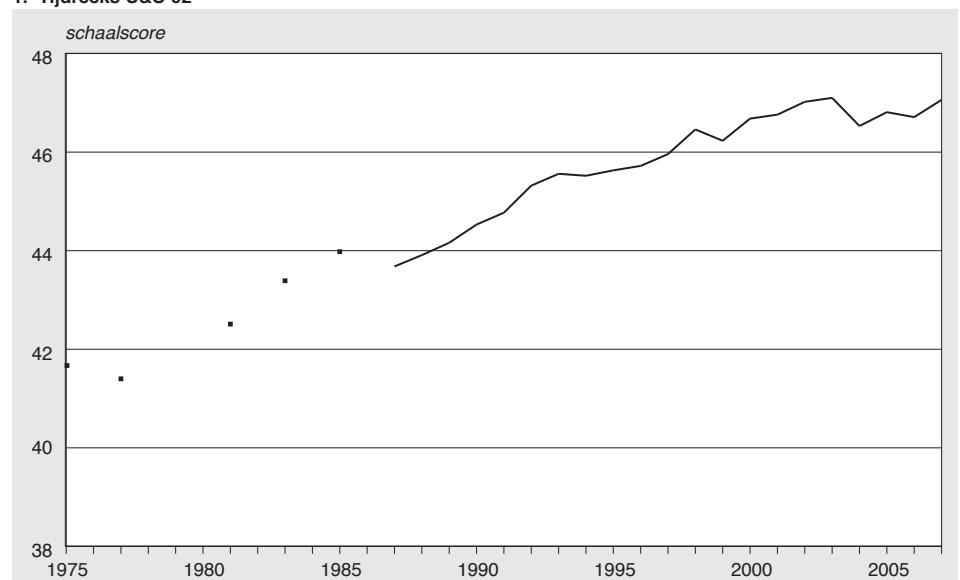
4.1.2 1987–2007

De U&S-92 en ISEI-92 zijn afgeleid uit de Standaard Beroepen Classificatie 1992 (SBC-92). In de periode 1987–2007 kunnen U&S-92 en ISEI-92 bijna altijd (tenminste 97,4 procent van de cases) afgeleid worden. De reden dat U&S-92 en ISEI-92 niet altijd kunnen worden afgeleid is dat het beroep SBC-92 soms missend of onbekend is.

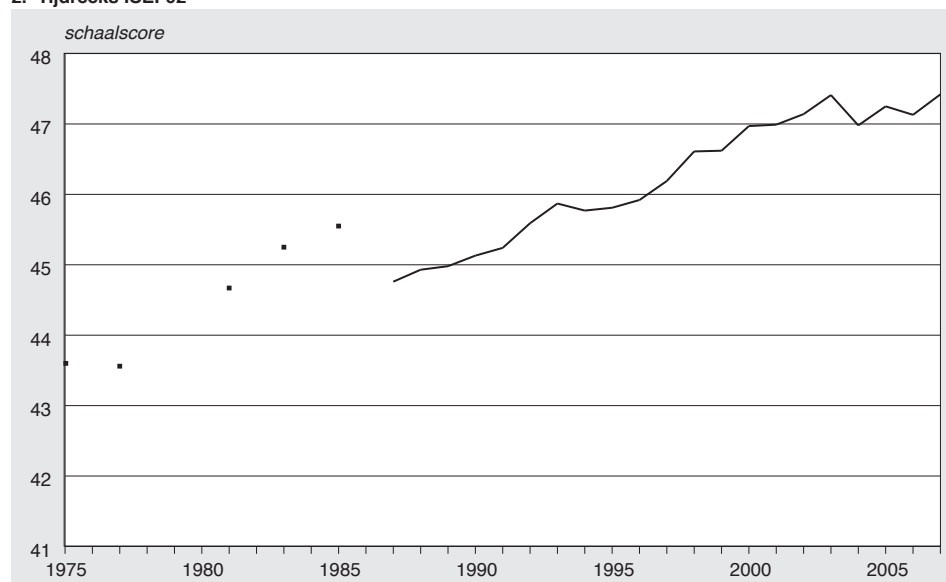
4.2 Trendbreuken

Per jaar is een (met de werkzame beroepsbevolking) gewogen gemiddelde van U&S-92 en ISEI-92 berekend. De resultaten van de tijdreeksen 1975–2007 staan in figuur 1 (U&S-92) en 2 (ISEI-92).

1. Tijdreeks U&S-92



2. Tijdreeks ISEI-92



Beide grafieken laten in de periode 1975–2007 een stijgende trend zien. Dit betekent dat het beroepsprestige en de sociaal-economische status van de werkzame beroepsbevolking zijn gestegen. In deze reeksen zitten drie trendbreuken, namelijk tussen 1977 en 1981, tussen 1985 en 1987, en tussen 2003 en 2004.

1. De trendbreuk tussen 1977 en 1981 valt te verklaren doordat 1975 en 1977 zijn gebaseerd op de DSC AKT (steekproef, niet gecorrigeerd naar definitie CCS 1991), terwijl 1981, 1983 en 1985 zijn gebaseerd op de AKT (opgehoogde aantallen, wel gecorrigeerd naar definitie CCS 1991).
2. De trendbreuk tussen 1985 en 1987 komt door de overgang van de AKT (1981, 1983, 1985) naar de EBB (1987–2007).
3. De trendbreuk tussen 2003 en 2004 wordt veroorzaakt doordat in de EBB sinds 2004 het beroep anders wordt getypeerd, namelijk door andere CBS werknemers dan de enquêteurs/enquêtrices (blijkt uit een gesprek met Bart Bakker). Deze trendbreuken moeten gerepareerd worden.

4.3 Reparatie van trendbreuken

Het repareren van deze trendbreuken vindt plaats in een aantal stappen. Eerst worden zogenaamde dummyvariabelen toegekend aan de afzonderlijke deelreeksen. Daarna wordt vanuit de literatuur bekeken welke andere exogene variabelen van betekenis zijn op de ontwikkeling van beroepsprestige en sociaal-economische status. Vervolgens worden de dummyvariabelen en de gekozen andere exogene variabelen meegenomen in een regressiemodel. Uit de regressie-analyse valt af te leiden welk deel van de trendbreuk veroorzaakt wordt door een methode-effect (gebruik van andere bronnen en methoden). Tot slot worden de tijdreeksen gecorrigeerd voor dit methode-effect en daarmee gerepareerd.

4.3.1 Dummyvariabelen

Zowel bij U&S-92 als ISEI-92 bestaat de tijdreeks uit vier deelreeksen:

1. Jaren 1975 en 1977
2. Jaren 1981, 1983 en 1985
3. Jaren 1987–2003
4. Jaren 2004–2007.

Het repareren van de trendbreuken vindt zo plaats dat de gehele tijdreeks consistent wordt gemaakt met de vierde deelreeks (jaren 2004–2007). Om te bepalen welk deel van de trendbreuk kan worden toegeschreven aan het methode-effect, zijn bij de eerste drie deelreeksen dummyvariabelen (waarde 1 of 0) ingezet:

Variabele	Waarde	Jaren
dummy1	1	1975, 1977
	0	overige jaren
dummy2	1	1981, 1983, 1985
	0	overige jaren
dummy3	1	1987–2003
		overige jaren

4.3.2 Andere exogene variabelen uit de literatuur

Met behulp van een regressieanalyse wordt het verband geschat tussen de afhankelijke variabelen uens92 (U&S-92) en isei92 (ISEI-92), en een aantal exogene variabelen. Uit de bestudeerde literatuur blijkt dat er al veel onderzoek is gedaan naar de variabelen die samenhangen met beroepsprestige en sociale status. Beroepsstatus wordt doorgegeven van generatie op generatie, zowel direct van vader op zoon als indirect via het onderwijs. De periode 1930–1980 laat de trend zien dat het bereikte beroepsniveau steeds sterker verbonden is aan het onderwijs dat men gevolgd heeft, en minder afhangt van sociale herkomst (De Graaf & Luijkx 1995). Vrouwen hebben in opleidingsniveau een sterke inhaalslag gemaakt, en in 1991 ligt het gemiddelde opleidingsniveau van jonge vrouwen even hoog als dat van mannen. De arbeidsmarktparticipatie van vrouwen is de afgelopen decennia enorm gestegen. Maar de traditionele scheiding tussen mannen- en vrouwenberoepen is blijven bestaan. Dit kwam voornamelijk doordat de werkgelegenheid groeide in die sectoren (vooral de dienstensector) waarin altijd al relatief veel vrouwen werkzaam waren. Vrouwenberoepen blijken gemiddeld niet een veel lager maatschappelijk aanzien te hebben dan mannenberoepen (Van der Lippe & Van Doorne-Huiskens 1995). In de loop van de tijd heeft de beroepsstructuur grote veranderingen ondergaan: midden jaren negentig zijn er meer hogere beroepen en veel minder lagere beroepen dan begin jaren vijftig (Ganzeboom & Luijk 1995). Op basis van deze literatuur is gekozen om naast de dummyvariabelen ook de volgende exogene variabelen mee te nemen in de regressie:

- aandeel van personen met hoger opleidingsniveau (hbo, wo) in de totale werkzame beroepsbevolking (opl_hoog_tot)
- aandeel vrouwen in de totale werkzame beroepsbevolking (wzbb_vrouw_tot)
- aandeel van de dienstverlening (inclusief overheid) in de totale economie (wp_dv_tot)

4.3.3 Regressiemodel

Met behulp van stapsgewijze regressie wordt bepaald welke variabelen uiteindelijk meegenomen worden in het model. De regressie geeft een erg goede modelfit, zowel bij uens92 (aangepaste $R^2=0,987$) als ook bij isei92 (aangepaste $R^2=0,969$). Gedetailleerde resultaten van de stapsgewijze regressie staan in een interne CBS nota (Den Boer 2010). De oplossing van het regressiemodel:

$$\text{uens92} = 30,441 + 2,025 \cdot \text{dummy1} + 1,998 \cdot \text{dummy2} + 1,136 \cdot \text{dummy3} + 12,972 \cdot \text{opl_hoog_tot} + 29,158 \cdot \text{wzbb_vrouw_tot}$$

$$\text{isei92} = 31,482 + 2,907 \cdot \text{dummy1} + 2,245 \cdot \text{dummy2} + 0,939 \cdot \text{dummy3} + 18,544 \cdot \text{opl_hoog_tot} + 18,399 \cdot \text{wp_dv_tot}$$

4.3.4 Methode-effect

De oplossing van het model is gebruikt om U&S-92 en ISEI-92 in de ontbrekende jaren (1976, 1978, 1979, 1980, 1982, 1984, 1986) te schatten. De parameters van de dummyvariabelen geven in de jaren met een trendbreuk aan welk deel van de trendbreuk toegeschreven wordt aan het methode-effect.

Methode-effect in tijdreeks U&S-92:

- 2,025 (jaren 1975 en 1977)
- 1,998 (jaren 1981, 1983 en 1985)
- 1,136 (jaren 1987–2003)

Methode-effect in tijdreeks ISEI-92:

- 2,907 (jaren 1975 en 1977)
- 2,245 (jaren 1981, 1983 en 1985)
- 0,939 (jaren 1987-2003)

4.3.5 Gerepareerde tijdreeksen U&S-92 en ISEI-92

De trendbreuken in de tijdreeksen van U&S-92 en ISEI-92 zijn gerepareerd door te corrigeren voor het methode-effect.

$$R_uens92 = uens92 - (2,025 * dummy1 + 1,998 * dummy2 + 1,136 * dummy3)$$

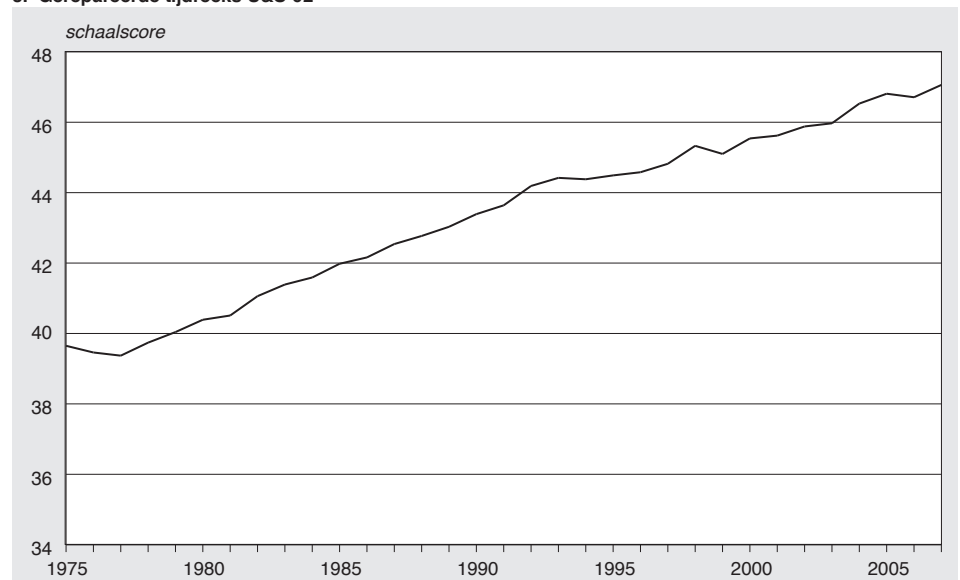
$$R_isei92 = uens92 - (2,907 * dummy1 + 2,245 * dummy2 + 0,939 * dummy3)$$

R_uens92 U&S-92 gerepareerd

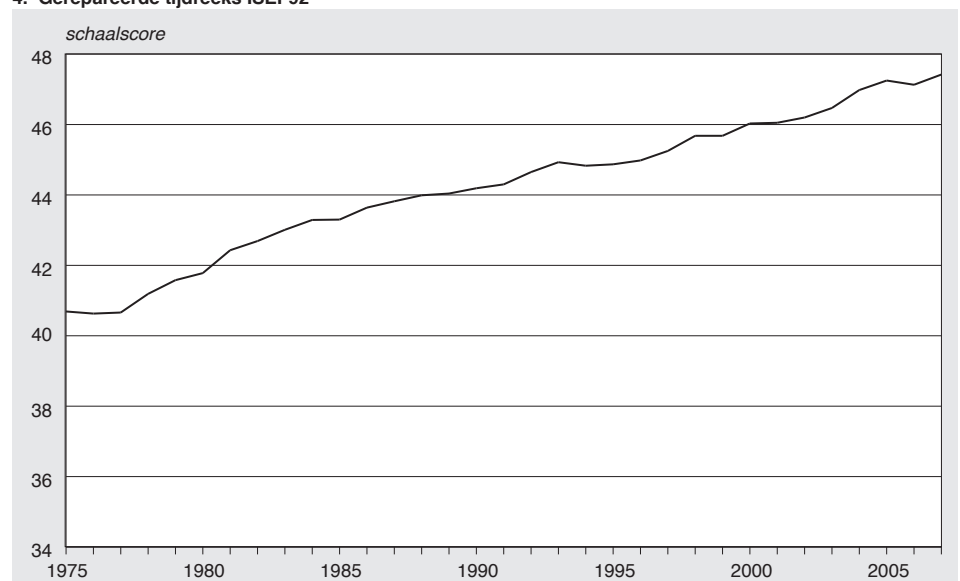
R_isei92 ISEI-92 gerepareerd

De gerepareerde tijdreeksen staan in figuur 3 (U&S-92) en 4 (ISEI-92).

3. Gerepareerde tijdreeks U&S-92



4. Gerepareerde tijdreeks ISEI-92



5. Sociale klasse (EGP-92)

De sociale klasse (EGP-92) is afgeleid met behulp van de in paragraaf 3 genoemde tools en setups. Bij het afleiden van de EGP-92 worden verschillende perioden onderscheiden. De periode 1975–1979 staat beschreven in paragraaf 5.1. In de periode 1981–1985 kan de EGP-92 op twee manieren worden afgeleid. Paragraaf 5.2 beschrijft beide manieren en vergelijkt ze met elkaar, om uiteindelijk een manier te kiezen. De afleiding van de EGP-92 in de jaren 1987–1995 staat in paragraaf 5.3. In de periode 1996–2007 kost het afleiden van de EGP-92 relatief meer werk, omdat het aantal personen waarover leiding wordt gegeven eerst wordt bijgeschat (paragraaf 5.4). Na afleiding blijken in de tijdreeks in verschillende sociale klassen trendbreuken te zitten. In paragraaf 5.5 staat een analyse van deze trendbreuken. Paragraaf 5.6 beschrijft hoe de trendbreuken worden gerepareerd.

5.1 1975–1979

In de jaren 1975 en 1977 is de sociale klasse (EGP-92) afgeleid uit de International Standard Classification of Occupations 1968 (ISCO-68), de positie in de werkring, en het aantal personen waarover leiding wordt gegeven. Daarvoor was het eerst nodig de Beroepen Classificatie 1984 (BC-84) om te zetten naar ISCO-68. Het aantal personen waarover leiding wordt gegeven is afgeleid uit de variabelen `LeidPers` (Geeft u leiding aan personeel?) en `AprsLeid` (Aan hoeveel medewerkers geeft u leiding?).

In de jaren 1975 en 1977 kan EGP-92 bijna altijd (respectievelijk 98,4 procent en 98,2 procent van de cases) worden afgeleid. EGP-92 kan soms niet worden afgeleid omdat het beroep dan onbekend is en dus niet omgezet kan worden naar beroep ISCO-68, of omdat het aantal personen waarover leiding wordt gegeven dan onbekend is. In 1979 kan de EGP-92 helemaal niet worden afgeleid omdat dan het beroep en het aantal personen waarover leiding wordt gegeven niet beschikbaar is.

5.2 1981–1985

De EGP-92 wordt afgeleid uit het beroep, de positie in de werkring, en het aantal personen waarover leiding wordt gegeven. Uit een inventarisatie van de benodigde variabelen (paragraaf 2) bleek al dat het aantal personen waarover leiding wordt gegeven niet beschikbaar is in de AKT, maar wel in de DSC AKT. In de jaren 1981–1985 kan de EGP-92 op twee manieren worden afgeleid.

De eerste manier is om de AKT 1981–1985 (beroep BC-84, positie in de werkring) te koppelen aan de DSC AKT 1981–1985 (aantal personen waarover leiding wordt gegeven) en daarna de EGP-92 af te leiden uit het gekoppelde bestand. De tweede manier is om EGP-92 direct af te leiden uit de DSC AKT 1981–1985 (beroep BC-84, positie in de werkring, aantal personen waarover leiding wordt gegeven). Nadat beide manieren met elkaar zijn vergeleken, is gekozen op welke manier EGP-92 wordt afgeleid.

5.2.1 Manier 1: EGP-92 afleiden uit koppeling AKT en DSC AKT

De AKT 1981–1985 en de DSC AKT 1981–1985 zijn aan elkaar gekoppeld via de variabele `persid` (persoons-identificatienummer), die zowel in de AKT als ook in de DSC AKT voorkomt. Voordat koppeling mogelijk is, moeten eerst een aantal bewerkingen op deze variabele plaatsvinden. Eerst is de variabele omgezet naar een string variabele en eventueel vooraan opgevuld met nullen. Daarna zijn in de AKT positie 1–8 en positie 11–12 van variabele `persid` samengevoegd tot de nieuwe variabele `str_persid_new` (tien posities). In de DSC AKT zijn positie 1-3 en positie 6–11 van variabele `persid`, na toevoeging van een voorloopnul, samengevoegd tot de nieuwe variabele `str_persid_new` (tien posities). De noodzakelijke bewerkingen op de variabele `persid` hebben tot gevolg dat een waarde van de nieuwe variabele `str_persid_new` meerdere keren kan voorkomen. Om de AKT via de

variabele str_persid_new aan de DSC AKT te kunnen koppelen is het nodig dat er alleen unieke cases van deze variabele voorkomen. Tabel 3 laat zien hoeveel unieke cases van variabele str_persid_new overblijven in de AKT en de DSC AKT 1981–1985.

Tabel 3
Unieke cases str_persid_new in AKT

	AKT	DSC AKT
	% van aantal cases	
1981	96,4	89,3
1983	91,2	75,5
1985	91,4	75,6

Vooraf in de DSC AKT blijven minder unieke cases van str_persid_new over, zodat relatief veel cases in de DSC AKT met informatie over het aantal personen waarover leiding wordt gegeven niet gekoppeld kunnen worden aan de AKT.

Na koppeling is de EGP-92 afgeleid. Hiervoor moest beroep BC-84 weer eerst omgezet worden naar beroep ISCO-68. In 1981, 1983 en 1985 kan de EGP-92 in respectievelijk 92,2 procent, 80,1 procent en 80,1 procent van de cases worden afgeleid. De EGP-92 kan niet altijd worden afgeleid omdat beroep BC-84 soms onbekend is en dus niet omgezet kan worden naar beroep ISCO-68, maar voornamelijk omdat het aantal personen waarover leiding wordt gegeven vaak ontbreekt. Deze laatste oorzaak komt doordat de AKT niet altijd aan de DSC AKT gekoppeld kan worden.

5.2.2 Manier 2: EGP-92 afleiden uit DSC AKT

In plaats van de EGP-92 af te leiden uit de AKT gekoppeld aan de DSC AKT, is het ook mogelijk de EGP-92 direct af te leiden uit de DSC AKT, net zoals in 1975 en 1977 gedaan is. De EGP-92 is afgeleid uit het beroep BC-84, de positie in de werkring, en het aantal personen waarover leiding wordt gegeven, alle uit de DSC AKT 1981–1985. Beroep BC-84 moest weer eerst omgezet worden naar beroep ISCO-68. In 1981, 1983 en 1985 kan de EGP-92 in respectievelijk 97,7 procent, 97,5 procent en 97,7 procent van de cases worden afgeleid. Dit is aanzienlijk vaker dan bij de afleiding uit de AKT gekoppeld aan de DSC AKT. De reden dat EGP-92 niet altijd kan worden afgeleid is meestal dat beroep BC-84 dan onbekend is en dus niet omgezet kan worden naar beroep ISCO-68, en soms dat het aantal personen waarover leiding wordt gegeven dan ontbreekt.

5.2.3 Vergelijking tussen manier 1 en 2

De resultaten van de EGP-92 afgeleid op manier 1 (afleiden uit koppeling tussen AKT en DSC AKT) en manier 2 (direct afleiden uit DSC AKT) zijn met elkaar vergeleken. Voor de

Tabel 4
EGP-92 afgeleid op manier 1 (uit koppeling tussen AKT en DSC AKT) en manier 2 (direct uit DSC AKT)

Sociale klasse (EGP-92)	Manier 1			Manier 2		
	1981	1983	1985	1981	1983	1985
	%					
I Hogere leidinggevendenden, academici	8	8	8	8	8	9
II Supervisoren hoofdarbeid, hooggeschoolde hoofdarbeid	18	21	22	18	20	20
IIIa Geschoolde hoofdarbeid	20	22	23	18	18	18
IIIb Half- en ongeschoolde hoofdarbeid	9	10	10	8	8	8
Iva Kleine zelfstandigen met minder dan 10 man personeel	1	0	0	1	1	1
Ivb Kleine zelfstandigen zonder personeel	3	1	1	4	4	4
Ivc Zelfstandige en leidinggevende agrariërs	2	2	3	2	2	2
V Supervisoren handarbeid, hooggeschoolde handarbeid	15	7	7	17	16	15
VI Geschoolde handarbeid	18	20	20	18	18	18
VIIa Halfgeschoolde en ongeschoolde handarbeid	3	3	3	2	3	3
VIIb Agrarische arbeid	3	4	3	3	3	3

volledigheid is ook de EGP-92 in 1975 en 1977 (beide afgeleid uit DSC AKT) in de tabellen geplaatst. Tabel 4 toont de resultaten van beide manieren.

Bij het afleiden van de EGP-92 op manier 1 lijkt er in sociale klasse V (Supervisors handarbeid, hooggeschoolde handarbeid) een trendbreuk te zitten tussen 1981 en 1983. Dit is vreemd omdat 1981 en 1983 uit dezelfde soort bron en op dezelfde manier zijn afgeleid. Bij het afleiden van de EGP-92 op manier 2 is deze trendbreuk niet aanwezig. Gekozen is om de EGP-92 in 1981, 1983 en 1985 af te leiden op manier 2 (direct afleiden uit DSC AKT), omdat de EGP-92 dan vaker kan worden bepaald. Voordeel van deze keuze is bovendien dat er geen trendbreuk zit in de tijdreeks sociale klasse V tussen 1981 en 1983.

5.3 1987–2007

In de periode 1987–2007 is de EGP-92 afgeleid uit het beroep (SBC-92), de positie in de werkring, en het aantal personen waarover leiding wordt gegeven. Waar het aantal personen waarover leiding wordt gegeven mist, is deze variabele geschat met behulp van de imputatiemethode van logistische regressie. Daarna is de EGP-92 afgeleid uit de beschikbare en geïmputeerde data. Tot slot is een plausibiliteitscontrole gedaan om – naast de resultaten van de logistische regressie – iets over de betrouwbaarheid van de imputatiemethode te zeggen.

5.3.1 EGP-92 afleiden uit beschikbare data

In de periode 1987–1995 kan de EGP-92 in eerste instantie slechts in 11,3 procent van de cases worden afgeleid uit de beschikbare data in de EBB 1987–1995. In de overige cases zou EGP-92 alleen bepaald kunnen worden met behulp van de twee hulpvariabelen positie in de werkring en het aantal personen waarover leiding wordt gegeven. De laatste hulpvariabele ontbreekt in de EBB 1987–1995.

In de perioden 1996–2000 en 2001–2007 kan de EGP-92 in respectievelijk 81,2 procent en 88,5 procent van de cases worden afgeleid uit de beschikbare data. De reden dat EGP-92 niet altijd kan worden bepaald, is dat het beroep, de positie in de werkring, en het aantal personen waarover leiding wordt gegeven niet altijd bekend zijn. In respectievelijk 78,5 en 86,8 procent van de cases zijn deze variabelen wel bekend. Deze percentages liggen iets lager dan de eerder genoemde percentages over het kunnen afleiden van EGP-92, omdat de variabelen beroep, positie in de werkring, en het aantal personen waarover leiding wordt gegeven niet altijd allemaal nodig zijn om EGP-92 te bepalen. In respectievelijk 21,4 procent en 13,2 procent van de cases is alléén het aantal personen waarover leiding wordt gegeven onbekend. Dus kan geconcludeerd worden dat het ontbreken van alléén het aantal personen waarover leiding wordt gegeven de voornaamste reden is dat EGP-92 niet altijd kan worden bepaald.

5.3.2 Koppelen van bestand EBB 1987–1995 en jaarbestanden EBB 1992–1995

In de jaarbestanden EBB 1992–1995 zitten de variabele die aangeeft of iemand leiding geeft (leidw1) en het aantal personen waarover leiding wordt gegeven (nleidw1). Daarom is het bestand EBB 1987–1995 gekoppeld aan de jaarbestanden EBB 1992–1995. De koppeling is gemaakt via de variabelen adresid (adresnummer), huinhnr (huishoudnummer) en persnr (persoonsnummer) in het bestand EBB 1987–1995 en de variabelen adresnr (adresnummer), aandnr (huishoudnummer) en h_op (persoonsnummer) in de jaarbestanden EBB 1992–1995.

Voordat koppeling mogelijk is, moeten eerst een aantal bewerkingen op deze variabelen plaatsvinden. Eerst zijn de variabelen adresid, huinhnr en persnr omgezet naar string-variabelen. Vervolgens zijn deze stringvariabelen samengevoegd tot de nieuwe variabele str_id_new (string, 9 posities). Ook de variabelen adresnr, aandnr en h_op zijn samengevoegd tot de nieuwe variabele str_id_new (string, 9 posities). Om het bestand EBB

1987–1995 via de variabele *str_id_new* aan de jaarbestanden EBB 1992–1995 te kunnen koppelen is het nodig dat er alleen unieke cases van deze variabele voorkomen. Tabel 5 laat zien hoeveel unieke cases van variabele *str_id_new* overblijven in het bestand EBB 1987–1995 en de jaarbestanden EBB 1992–1995.

Tabel 5
Unieke cases *str_persid_new* in bestand EBB 1987–1995 en jaarbestanden EBB

	Bestand EBB 1987–1995	Jaarbestanden EBB 1992–1995
	<i>% van aantal cases</i>	
1992–1995	99,8	86,4

Na koppeling van het bestand EBB 1987–1995 aan de jaarbestanden EBB 1992–1995 kan de EGP-92 in de periode 1987–1995 in 42,9 procent van het aantal cases worden afgeleid worden. In de jaren 1992–1995 ligt dit percentage op 78,5 procent.

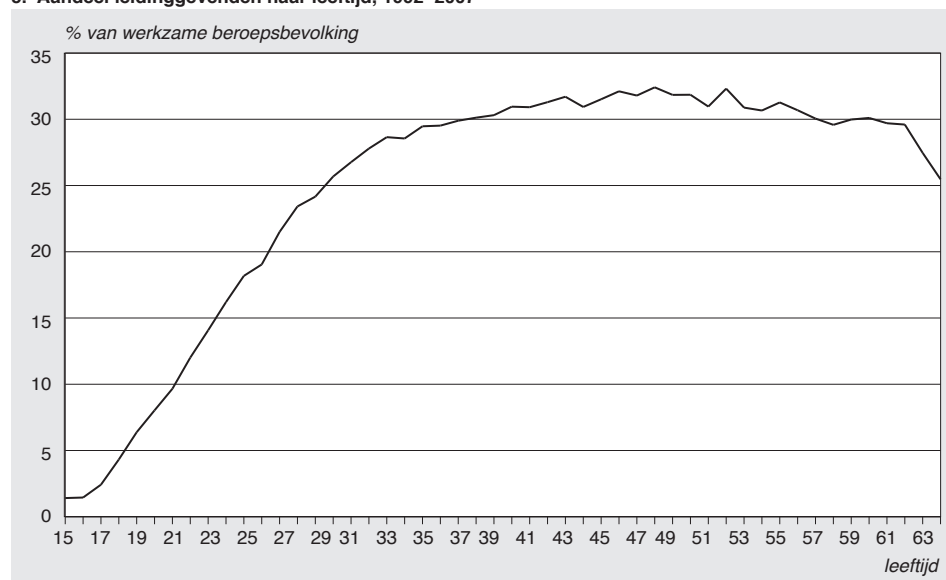
5.3.3 Analyse van leidinggeven naar achtergrondvariabelen

Of men al dan niet leiding geeft hangt vermoedelijk samen met allerlei achtergrondvariabelen, zoals geslacht en leeftijd. De volgende analyse van deze achtergrondvariabelen heeft betrekking op de werkzame beroepsbevolking waarvan bekend is of leiding wordt gegeven, en geeft altijd een gemiddelde over de periode 1992–2007. De gegevens van 1992–1995 komen uit de koppeling tussen het bestand EBB 1987–1995 en de jaarbestanden EBB 1992–1995. Tabel 6 toont het aandeel leidinggevendenden naar geslacht.

Tabel 6
Aandeel leidinggevendenden naar geslacht, 1992–2007

	Geeft niet leiding	Geeft wel leiding
	<i>% van werkzame beroepsbevolking</i>	
Mannen	67	33
Vrouwen	82	18
Totaal	74	26

5. Aandeel leidinggevendenden naar leeftijd, 1992–2007



Uit de tabel blijkt dat 26 procent van de werkzame beroepsbevolking leidinggeeft. Er zijn wel duidelijke verschillen tussen mannen en vrouwen; 33 procent van de mannen geeft leiding, bij de vrouwen ligt het aandeel leidinggeevenden met 18 procent aanzienlijk lager. Figuur 5 geeft het aandeel leidinggeevenden naar leeftijd.

De grafiek laat zien dat het aandeel leidinggeevenden sterk toeneemt naarmate men ouder wordt. Vanaf een leeftijd van ongeveer 30 jaar stijgt dit aandeel minder hard en tot een leeftijd van ongeveer 62 jaar blijft het percentage leidinggeevenden veel stabiel. Opvallend is dat personen van 63 en 64 jaar weer minder leidinggeven. Op basis van deze analyse zijn daarom de volgende leeftijdsgroepen gemaakt: 15–17 jaar, 18–19 jaar, 20–24 jaar, 25–29 jaar, 30–62 jaar, 63–64 jaar.

Een andere belangrijke variabele die samenhangt met leiding geven is beroepsklasse (SBC-92), en wel specifiek de waarden 78 managers (hbo werk- en denkniveau) e.d. en 98 managers (wetenschappelijk werk- en denkniveau). De grootte van beroepsklasse 78 is gemiddeld 0,6 procent en die van beroepsklasse 98 gemiddeld 1,4 procent van de werkzame beroepsbevolking. Tabel 7 toont het aandeel leidinggeevenden onder managers.

Tabel 7
Aandeel leidinggeevenden bij managers (in procenten van de werkzame beroepsbevolking), 1992–2007

Beroepsklasse (SBC-92)	Geeft niet leiding	geeft wel leiding
	%	
78 managers (hbo werk- en denkniveau)	7	93
98 managers (wetenschappelijk werk- en denkniveau)	5	95
totaal (78 en 98)	6	94

Uit de tabel blijkt dat 6 procent van de managers niet leiding geeft. Managers hoeven dus kennelijk niet per se leiding te geven. Het aandeel leidinggeevenden onder managers is (zoals toch verwacht mocht worden) veel hoger dan de 26 procent (tabel 6) van de totale werkzame beroepsbevolking. Voor de logistische regressie is de dummyvariabele manager aangemaakt:

- 1 manager (beroepsklasse 78 en 98)
- 0 geen manager (overige beroepsklassen)

5.3.4 Imputeren van leidinggeven (logistische regressie)

Het aantal personen waarover leiding wordt gegeven is afgeleid uit de variabelen LeidW1 (geeft persoon in de eerste werkring leiding aan medewerkers) en NLeidW1 (aantal personen waarover leiding wordt gegeven in de eerste werkring). Omdat de variabelen LeidW1 en NLeidW1 vaak missend of onbekend zijn, wordt het aantal personen waarover leiding wordt gegeven in die gevallen geschat. Deze schatting is gedaan met de imputatiemethode van logistische regressie.

Eerst zijn de variabelen LeidW1 en NLeidW1 hercodeerd als een dichotome variabelen (twee waarden). De variabele LeidW1 is hercodeerd als dichotome variabele leiding:

- 1 (geeft wel leiding)
- 0 (geeft niet leiding)

De variabele NLeidW1 is hercodeerd als dichotome variabele Nleiding:

- 1 (geeft leiding aan 10 personen of meer)
- 0 (geeft leiding aan 1-9 personen)

Vervolgens vinden twee logistische regressies plaats, de eerste om de variabele leiding te schatten, de tweede om de variabele Nleiding te schatten.

1. Logistische regressie van variabele leiding

De eerste logistische regressie heeft betrekking op de werkzame beroepsbevolking waarvan de variabele LeidW1 bekend is. De gegevens van 1992–1995 komen uit de koppeling

tussen het bestand EBB 1987–1995 en de jaarbestanden EBB 1992–1995. De variabele leiding is geschat met exogene variabelen geslacht, leeftijdsgroep (LFTG: 15–17, 18–19, 20–24, 25–29, 30–62, 63–64 jaar), beroepsniveau SBC-92 (brpni92/ABER92NI), manager (dummyvariabele). De regressie geeft een goede modelfit. Chi^2 (df=11) = 29980 in 1992–1995, Chi^2 (df=11) = 37320 in 1996–2000 en Chi^2 (df=11) = 44400 in 2001–2007. Alle regressiecoëfficiënten zijn significant. Gedetailleerde resultaten van de logistische regressie van variabele leiding staan in een interne CBS nota (Den Boer 2010).

2. Logistische regressie van variabele Nleiding

De tweede logistische regressie is gedaan voor de werkzame beroepsbevolking waarvan de variabele NLeidW1 bekend is. De gegevens van 1992–1995 komen weer uit de koppeling tussen het bestand EBB 1987–1995 en de jaarbestanden EBB 1992–1995. De variabele Nleiding is geschat met exogene variabelen geslacht, leeftijdsgroep (LFTG2: 15–24, 25–34, 35–44, 45–54, 55–64 jaar), beroepsniveau SBC-92 (brpni92/ABER92NI), manager (dummyvariabele). Ook deze regressie heeft een goede modelfit. Chi^2 (df=10) = 6978 in 1992–1995, Chi^2 (df=10) = 9209 in 1996–2000 en Chi^2 (df=10) = 13201 in 2001–2007. Alle regressiecoëfficiënten zijn weer significant. De logistische regressie van variabele Nleiding staat gedetailleerder beschreven in een interne CBS nota (Den Boer 2010). Het is overigens normaal dat zelfs bij goede logistische regressiemodellen de R^2 -waarden relatief laag zijn in vergelijking met de R^2 -waarden bij goede lineaire regressiemodellen (Hosmer & Lemeshow 2000, p. 167). Ze zijn vooral gebruikt als hulpmiddel bij het evalueren van verschillende modellen.

De kans P (leiding=1) is toegekend aan de werkzame beroepsbevolking waarvan de variabele LeidW1 missend of onbekend is. Hierbij is de kans P (leiding=1) in de periode 1992–1995 toegekend aan de periode 1992–1995 en ook aan de periode 1987–1991. Dit is gedaan onder de aanname dat de jaren 1992–1995 representatief zijn voor de jaren 1987–1991. Op dezelfde wijze is de kans P (Nleiding=1) toegekend.

Tot slot is in de periode 1996–2007 voor een specifieke groep personen, namelijk de zelfstandigen werkzaam in eigen bedrijf met personeel in dienst, aangenomen dat zij leiding geven. Het aantal personen waarover zelfstandigen (PosWrk1=21) met personeel (Person1=1) leidinggeven wordt gehaald uit de bedrijfsomvang van het eigen bedrijf (BdrOmvz1). In 1987–1995 is de groep zelfstandigen werkzaam in eigen bedrijf veel kleiner (15 procent), en bovendien is in 1987–1995 geen informatie over personeel in dienst. Daarom is in 1987–1995 de imputatiemethode van logistische regressie aangehouden.

5.3.5 Plausibiliteitscontrole

Om een indruk te krijgen van de kwaliteit van de imputatiemethode van logistische regressie, heeft een plausibiliteitscontrole plaatsgevonden. Deze plausibiliteitscontrole staat genoemd in de methodenreeks (Israëls, Pannekoek & Schulte Nordholt 2007). Bij de werkzame beroepsbevolking waarvoor de variabele LeidW1 bekend is, is de imputatiemethode van logistische regressie ingezet om de variabele leiding te schatten. Vervolgens is de bekende variabele LeidW1 vergeleken met de geschatte variabele leiding. Dezelfde analyse is gedaan voor de variabelen NLeidW1 en Nleiding. De vergelijking is gemaakt naar geslacht.

Tabel 8
Plausibiliteitscontrole: leidinggeven in database en door imputatie, 1992–2007

Periode	Geslacht	Geeft leiding		Geeft leiding aan 10 of meer personen	
		Database	Imputatie	Database	Imputatie
<i>% van werkzame beroepsbevolking</i>					
1992–1995	mannen	31,0	30,8	37,3	37,2
	vrouwen	15,1	14,9	25,4	25,8
1996–2000	mannen	31,8	31,5	38,4	38,3
	vrouwen	15,8	15,5	27,2	27,4
2001–2007	mannen	34,2	34,4	39,4	39,4
	vrouwen	17,9	18,3	29,3	29,3

Tabel 8 laat zien dat de bekende waarden in de database (EBB bestanden) en de geschatte waarden uit de imputatiemethode van logistische regressie goed overeenkomen.

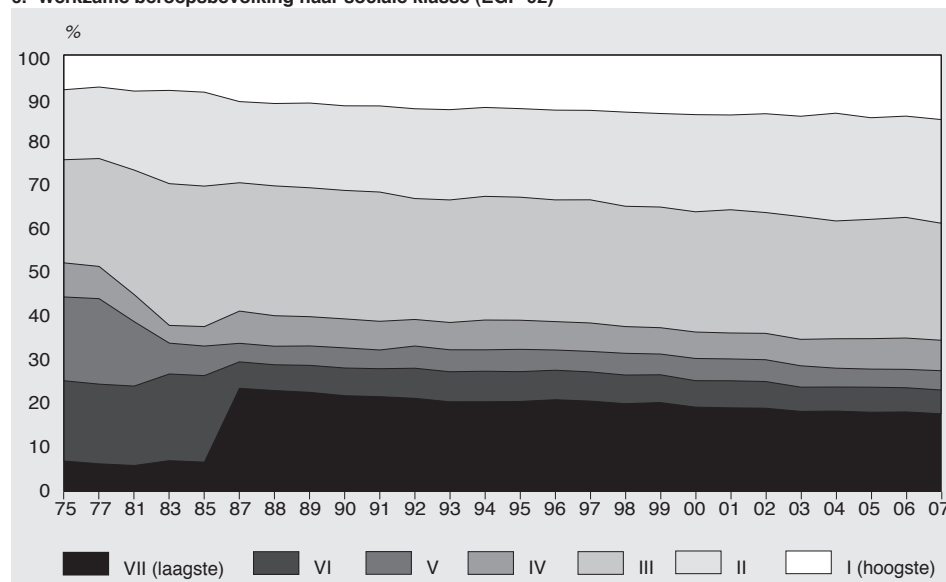
5.3.6 EGP-92 afleiden uit beschikbare en geïmputeerde data

De percentages dat in de perioden 1996–2000 en 2001–2007 alléén het aantal personen waarover leiding wordt gegeven onbekend is (21,4 procent en 13,2 procent) wordt door gebruik van de geschatte variabelen leiding en Nleiding gereduceerd naar respectievelijk 0,3 procent en 0,1 procent. In de perioden 1987–1995, 1996–2000 en 2001–2007 kan de EGP-92 nu bijna altijd (respectievelijk 97,4 procent, 99,6 procent en 99,9 procent van de cases) bepaald worden.

5.4 Trendbreuken

De werkzame beroepsbevolking is ingedeeld naar sociale klasse (EGP-92). In figuur 6 staan de resultaten van 1975–2007, geaggregeerd naar zeven sociale klassen. In de tijdreeksen naar sociale klasse zitten trendbreuken.

6. Werkzame beroepsbevolking naar sociale klasse (EGP-92)



5.5 Reparatie van trendbreuken

5.5.1 Regressiemodel

Het repareren van de reeksen vindt weer in een aantal stappen plaats, en wel zo dat de gehele tijdreeks consistent wordt gemaakt met de deelreeks 2004–2007. Eerst zijn dummyvariabelen aangemaakt.

Variabele	Waarde	Jaren
dummy1	1	1975, 1977
	0	overige jaren
dummy2	1	1981, 1983, 1985
	0	overige jaren
dummy3	1	1987–2003
	0	overige jaren

Daarna zijn de dummyvariabelen en een aantal andere exogene variabelen meegenomen in een regressiemodel. Doelvariabele is het aandeel in de werkzame beroepsbevolking naar zeven sociale klassen. Er worden dus zeven reeksen gerepareerd. Naast de dummyvariabelen zijn dezelfde exogene variabelen gebruikt als bij de reparatie van trendbreuken in de tijdreeksen U&S-92 en ISEI-92. Uit de literatuur bleek namelijk dat deze variabelen samenhangen met beroepsprestige en sociaal-economische status, en waarschijnlijk ook met sociale klasse:

- aandeel van personen met hoger opleidingsniveau (hbo, wo) in de totale werkzame beroepsbevolking (opl_hoog_tot)
- aandeel vrouwen in de totale werkzame beroepsbevolking (wzbb_vrouw_tot)
- aandeel van de dienstverlening (inclusief overheid) in de totale economie (wp_dv_tot).

Met behulp van stapsgewijze regressie is bepaald welke exogene variabelen uiteindelijk worden meegenomen. De oplossing van het regressiemodel (Den Boer 2010) is ingezet om het aandeel in de werkzame beroepsbevolking naar sociale klasse te schatten in de onbekende jaren (1976, 1978, 1979, 1980, 1982, 1984, 1986).

Vervolgens zijn de tijdreeksen gecorrigeerd voor het methode-effect, dat is gekwantificeerd door de parameter van de dummyvariabele.

5.5.2 Consistentie van de gerepareerde reeksen

Na reparatie van de trendbreuken blijkt dat de zeven reeksen met het aandeel in de werkzame beroepsbevolking niet meer optellen tot 100 procent. Met behulp van een Lagrange-functie, waarbij ook rekening is gehouden met de kwaliteit van de schatting van de gerepareerde reeksen, kan het (zeer geringe) verschil verdeeld worden over de zeven reeksen (Kardal & Van der Stegen 2009). De oplossing in formulevorm:

Het oplossen van het stelsel:

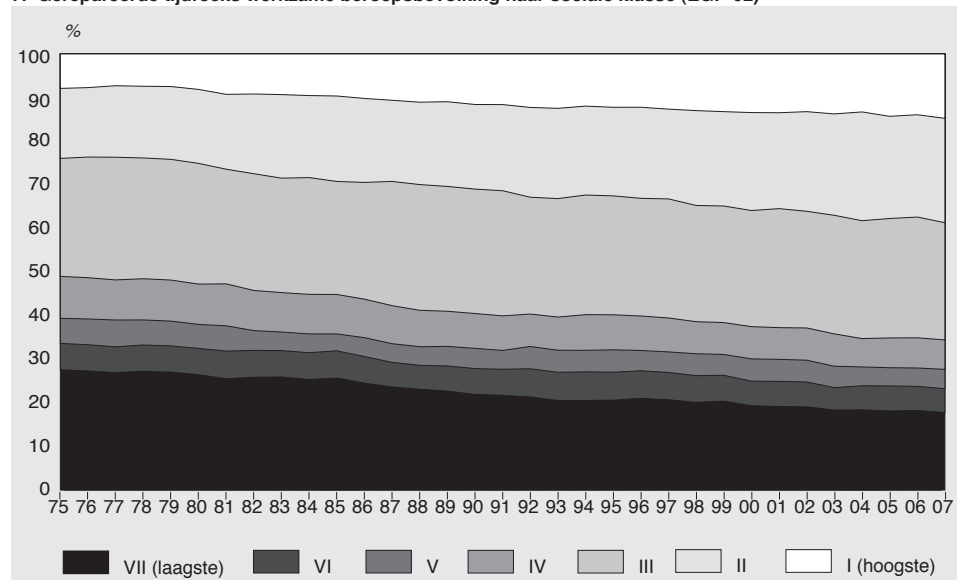
$$\min_{y_t^*} \left\{ \begin{array}{l} \left\| \frac{y_t^* - y_t}{v_t} \right\|_2 \\ \Delta_t y_t = 0 \end{array} \right.$$

waarbij v_t een component van de variantiematrix V_t .
geeft als oplossing:

$$y_t^* = y_t - V_t \Delta' (\Delta V_t \Delta')^{-1} \Delta y_t$$

De restrictiematrix Δ_t beschrijft de lineaire afhankelijkheden van de verschillende reeksen op tijdstip t . De componenten van de verschillende reeksen op tijdstip t staan in de vector y_t . Het

7. Gerepareerde tijdreeks werkzame beroepsbevolking naar sociale klasse (EGP-92)



oplossen van $\Delta_t y_t = 0$ (de reeksen moeten optellen tot 100 procent) vindt plaats door rekening te houden met de kwaliteit van de schatting van de gerepareerde reeksen, gekwantificeerd door de standaardfout van de dummyvariabelen, die staat in de diagonaalmatrix V_t . De resultaten van de gerepareerde en consistent gemaakte reeksen staan in figuur 7.

Resultaten van beroepsprestige (U&S-92), sociaal-economische status (ISEI-92) en sociale klasse (EGP-92)

De resultaten van de tijdreeksen die gerepareerd zijn voor trendbreuken staan in tabel 9 (U&S-92 en ISEI-92) en tabel 10 (EGP-92). De cijfers in tabel 9 en 10 wijken iets af van eerder in deze publicatie beschreven. De cijfers van 2004–2007 zijn gewijzigd door een herziening van de EBB cijfers. Daarnaast zijn inmiddels cijfers van de jaren 2008 en 2009 toegevoegd.

Tabel 9
Gerepareerde tijdreeksen beroepsprestige (U&S-92) en sociaal-economische status (ISEI-92)

Jaar	U&S-92	ISEI-92
1975	39,7	40,7
1976	39,5	40,6
1977	39,4	40,7
1978	39,7	41,2
1979	40,0	41,6
1980	40,4	41,8
1981	40,5	42,4
1982	41,1	42,7
1983	41,4	43,0
1984	41,6	43,3
1985	42,0	43,3
1986	42,2	43,6
1987	42,5	43,8
1988	42,8	44,0
1989	43,0	44,0
1990	43,4	44,2
1991	43,6	44,3
1992	44,2	44,7
1993	44,4	44,9
1994	44,4	44,8
1995	44,5	44,9
1996	44,6	45,0
1997	44,8	45,3
1998	45,3	45,7
1999	45,1	45,7
2000	45,5	46,0
2001	45,6	46,1
2002	45,9	46,2
2003	46,0	46,5
2004	46,2	46,6
2005	46,4	46,9
2006	46,4	46,8
2007	46,7	47,1
2008	46,9	47,2
2009	46,9	47,2

Tabel 10
Gerepareerde tijdreeks werkzame beroepsbevolking naar sociale klasse (EGP-92)

jaar	I	II	III	IV	V	VI	VII
1975	7,9	16,1	27,0	9,6	5,8	6,0	27,6
1976	7,7	15,9	27,7	9,5	5,9	6,0	27,3
1977	7,3	16,4	28,1	9,2	6,1	5,9	26,9
1978	7,4	16,4	27,7	9,5	5,7	6,0	27,2
1979	7,5	16,7	27,7	9,4	5,7	6,0	27,0
1980	8,1	17,0	27,7	9,3	5,5	6,0	26,5
1981	9,3	17,2	26,3	9,6	5,8	6,4	25,5
1982	9,2	18,3	26,8	9,2	4,5	6,1	25,9
1983	9,3	19,2	26,2	9,0	4,3	6,0	25,9
1984	9,5	18,8	26,8	9,1	4,3	6,1	25,3
1985	9,6	19,6	25,9	9,0	3,9	6,2	25,7
1986	10,2	19,3	26,8	8,8	4,3	6,1	24,5
1987	10,6	18,6	26,7	8,7	4,3	5,6	23,6
1988	11,1	18,9	28,8	8,4	4,3	5,5	23,1
1989	11,0	19,4	28,6	8,1	4,5	5,8	22,7
1990	11,6	19,4	28,5	8,0	4,6	5,9	21,9
1991	11,6	19,8	28,7	7,9	4,3	6,0	21,7
1992	12,3	20,6	26,8	7,4	5,1	6,5	21,3
1993	12,5	20,7	27,2	7,6	5,1	6,5	20,5
1994	12,0	20,4	27,4	8,2	4,9	6,6	20,5
1995	12,2	20,3	27,3	8,0	5,1	6,4	20,6
1996	12,2	20,9	27,0	7,9	4,7	6,3	21,0
1997	12,7	20,6	27,3	7,8	4,7	6,3	20,7
1998	13,0	21,8	26,7	7,4	5,0	6,1	20,1
1999	13,2	21,6	26,7	7,3	4,8	5,9	20,4
2000	13,4	22,5	26,6	7,4	5,1	5,6	19,3
2001	13,5	22,0	27,3	7,3	5,1	5,8	19,2
2002	13,2	22,9	26,8	7,4	5,0	5,7	19,1
2003	13,8	23,3	27,2	7,4	4,9	5,1	18,4
2004	12,9	24,3	26,7	7,6	4,3	5,5	18,7
2005	13,8	22,8	27,1	7,8	4,2	5,8	18,6
2006	13,5	22,9	27,4	7,7	4,2	5,6	18,6
2007	14,3	23,5	26,6	7,7	4,4	5,5	18,0
2008	14,8	23,2	26,7	7,6	4,6	5,2	17,9
2009	14,7	23,7	26,8	7,9	4,6	5,1	17,4

Referenties

- Bakker, B., I. Sieben, P. Nieuwbeerta & H. Ganzeboom (1997), Maten voor prestige, sociaal-economische status en sociale klasse voor de Standaard beroepenclassificatie 1992. In: *Sociale Wetenschappen*, 40, nr. 1, p. 1–22.
- Boer, P.M. den (2010), *Tijdreeksen werkzame beroepsbevolking naar beroepsprestige (U&S-92), sociaal-economische status (ISEI-92) en sociale klasse (EGP-92)*. Interne nota, Centraal Bureau voor de Statistiek.
- Ganzeboom, H.B.G. & R. Luijkx (1995), Intergenerationele beroepsmobiliteit in Nederland: patronen en historische veranderingen. In: Dronkers, J. & W.C. Ultee (red.), *Verschuivende ongelijkheid in Nederland. Sociale gelaagdheid en mobiliteit*, p. 14–30, Assen, Van Gorcum.
- Graaf, P.M. de & R. Luijkx (1995), Paden naar succes: geboorte of diploma's? In: Dronkers J. & W.C. Ultee (red.), *Verschuivende ongelijkheid in Nederland. Sociale gelaagdheid en mobiliteit*, p. 31–45, Assen, Van Gorcum.
- Israëls, A.Z., J. Pannekoek & E. Schulte Nordholt (2007), *Methodenreeks, Thema: Controle en correctie/imputatie, Deelthema: Imputatie*. Interne nota, Voorburg, Centraal Bureau voor de Statistiek.
- Kardal, M. & R.H.M. van der Stegen (2009), *Methodebeschrijving breukreparatie toegepast op PS tijdreeksen handel en diensten*. Interne nota, Centraal Bureau voor de Statistiek.
- Lippe, T. van der & J. van Doorne-Huiskens (1995), Veranderingen in stratificatie tussen mannen en vrouwen? In: Dronkers, J. & W.C. Ultee (red.), *Verschuivende ongelijkheid in Nederland. Sociale gelaagdheid en mobiliteit*, p. 104–124, Assen, Van Gorcum.
- Erikson, R., J.H. Goldthorpe & L. Portocarero (1979), *International Class Mobility in Three Western European Countries*. In: *British Journal of Sociology* 30, p. 415–441.
- Ganzeboom, H.B.G., P.M. De Graaf & D.J. Treiman (1992), A standard international socio-economic index of occupational status. In: *Social Science Research*, 21, p. 1–56.
- Ganzeboom, Harry B.G. & Treiman, Donald J. (1993), *International Stratification and Mobility File: Conversion Tools*, <http://www.fsw.vu.nl/~h.ganzeboom/ismf>
- Hosmer, David W. & Stanley Lemeshow (2000), *Applied Logistic Regression*, second edition, New York, John Wiley & Sons, Inc.
- Sixma, H & W.C. Ultee (1983), Een beroepsprestigeschaal voor Nederland in de jaren tachtig. In: *Mens en Maatschappij* 4, p. 360–382.