

## Bijlage G: Het berekenen van een index zonder gewichten

In de hoofdtekst is een alternatieve leefsituatie-index gepresenteerd waarbij alle indicatoren simpelweg bij elkaar opgeteld worden. In deze bijlage beschrijven we enkele verdere aanpassingen op die methodiek en bekijken welke gevolgen dat heeft voor de berekening van de leefsituatie-scores.

De enige aanpassing die we bij het simpelweg optellen van de indicatoren toepassen, is dat er zodanig gehercodeerd wordt dat de antwoordmogelijkheid waarvan wordt aangenomen dat die een positieve bijdrage levert aan de leefsituatie-index de hoogste score krijgt (hoe groter de kamer, hoe hoger de categoriekwantificatie; hetzelfde geldt bv voor het aantal sporten, geringe mate van eenzaamheid en afwezigheid van belemmeringen). Het resultaat is een index waarvan de minimale score die behaald kan worden 19 is (op alle vragen het minst positieve antwoord) en de maximale score 93 (op alle vragen het meest positieve antwoord). In de praktijk loopt de score van minimaal 22 tot maximaal 77. Deze index hangt sterk samen met de index die met behulp van *Overals* geconstrueerd wordt: de correlatie-coëfficiënt is 0,91.

Er zijn echter grote verschillen in de maximale scores die per indicator en vervolgens ook per domein kunnen worden behaald: voor vrijetijdsbesteding 21 en voor vakantie 5. Hoe hoger de maximale score per domein, hoe groter de kans dat er grotere variatie is, waardoor het belang van dat domein toeneemt. Let wel: de maximale totaal-score is 93, dus de score van vrijetijdsbesteding bepaalt al voor een groot deel de uitkomst.

### *Een leefsituatie-index zonder gewichten*

We kunnen de berekeningen nog wat verder verfijnen om het ‘nadeel’ van de grotere verschillen in maximale scores (en dus grote verschillen in bijdragen aan het eindresultaat) te verkleinen. Ten eerste kan een oververtegenwoordiging van indicatoren met veel antwoordmogelijkheden (en daardoor van clusters met veel indicatoren) voorkomen worden door eerst de indicatoren te delen door het aantal antwoordmogelijkheden. Bijvoorbeeld: de indicator ‘diversiteit vrijetijdsactiviteiten’ kent een minimale score van 1 en een maximale van 8. Door van deze indicator de score te delen door 8, en door bij alle overige indicatoren ook te delen door het maximum, is de maximale score voor elke indicator gelijk (namelijk 1). Een andere manier om dit te bereiken is door eerst de indicatoren per domein op te tellen en vervolgens de score per domein te delen door het maximum van dat domein.

Ten tweede kan er voor worden gezorgd dat één missende waarde op een indicator niet automatisch betekent dat geen index-score wordt berekend. We kunnen bijvoorbeeld besluiten om alleen het domein waarbinnen die indicator valt als missing aan te merken, maar voor de andere domeinen wel een score te berekenen.

Tot slot kunnen we de nominale indicatoren zodanig hercoderen dat de werkelijk beste optie de hoogste score krijgt. Omdat we niet weten hoe deze volgorde te bepalen is, maken we gebruik van de resultaten uit de *Overals*-analyse. Dat is natuurlijk niet helemaal eerlijk, want we wilden met deze test nu juist kijken of we *zonder* die analyse de leefsituatie kunnen bepalen door indicatoren op te tellen. De reden om het toch te doen is dat zo een beeld ontstaat of dat überhaupt gaat lukken. We kunnen dan bijvoorbeeld besluiten om eens in de zoveel tijd de nominale indicatoren op deze wijze (dus met behulp van *Overals*) te ijken om vervolgens de index ongewogen opgeteld te presenteren.<sup>1</sup>

Deze drie aanpassingen leiden tot de volgende vijf indexen (waarbij voor elke index dus geldt dat woningtype en mobiliteit gehercodeerd zijn volgens de zojuist beschreven stap):

1. *geengew* een gewone optelsom van alle indicatoren, waarbij gehercodeerd is tot een hoge score voor positieve antwoorden (heeft iemand op een indicator een missende waarde, dan krijgt hij geen leefsituatie-score);

---

<sup>1</sup> De hercodering naar de *Overals*-uitkomsten leidt tot een volgorde van minst positief tot meest positief: ouderwoning; overige woning; vrijstaande woning; woning plus winkel; flat of appartement; eengezinswoning; studentenwoning. Deze vreemd aandoende volgorde ontstaat door de combinatie van woninggerelateerde indicatoren: zie verder bijlage B.

2. *geengew2* nu worden de indicatoren gedeeld door de maximale score en vervolgens worden de indicatoren opgeteld (heeft iemand op een indicator een missende waarde, dan krijgt hij geen leefsituatie-score);
3. *geengew3* lijkt op de vorige, maar nu worden eerst per domein de indicatoren opgeteld, vervolgens worden de domeinscores gedeeld door de maximumscore per domein (heeft iemand op een indicator een missende waarde, dan krijgt hij geen leefsituatie-score);
4. *geengew4* lijkt op 2 (dus eerst indicatoren delen door maximum; per domein optellen), maar nu zorgt een missende waarde alleen voor een missende score op het domeinen waarbinnen de betreffende indicator valt; voor de overige domeinen krijgt hij wel een score (iedereen krijgt een leefsituatie-score);
5. *geengew5* lijkt op 3 (dus per domein indicatoren optellen, dan delen door maximum per domein), maar nu zorgt een missende waarde alleen voor een missende score op het domein waarbinnen de betreffende indicator valt (iedereen krijgt een leefsituatie-score).

Uit de resultaten blijkt dat de tweede optie om een index zonder gewichten te berekenen de meeste samenhang heeft met de leefsituatie-index gebaseerd op Overals, maar alle vijf de alternatieve methodes hebben een grote samenhang met de oorspronkelijke index (de correlatie-coëfficiënten zijn hoger dan 0.90 - zie tabel G.1 en bijlage B voor uitleg over de werking van Overals). Ook als we de resultaten uitsplitsen naar enkele achtergrondkenmerken zijn de verschillen gering (zie verderop in de bijlage).

Tabel G.1 samenhang tussen de leefsituatie index op basis van overals, enkele indexen zonder gewichten en afzonderlijke domeinscores

	leefsituatie-index	geengew5	geengew	geengew2	geengew3	geengew4
leefsituatie-index	1,00					
geengew5	0,94	1,00				
geengew	0,92	0,97	1,00			
geengew2	0,96	0,99	0,97	1,00		
geengew3	0,94	1,00	0,97	0,99	1,00	
geengew4	0,94	0,97	0,97	1,00	0,99	1,00
<i>domeinen</i>						
woning	0,67	0,55	0,58	0,65	0,55	0,65
vrijtijdsbesteding	0,72	0,66	0,72	0,67	0,66	0,66
sociale participatie	0,47	0,49	0,56	0,49	0,50	0,47
sport	0,57	0,61	0,62	0,57	0,60	0,58
vakantie	0,66	0,66	0,54	0,63	0,65	0,63
consumptiegoederen	0,68	0,61	0,52	0,62	0,60	0,61
mobiliteit	0,58	0,46	0,42	0,50	0,46	0,48
gezondheid	0,48	0,55	0,44	0,50	0,54	0,50

NB: de clusters zijn hier geoptimaliseerd voor samenhang met de leefsituatie-index (dwz geherkodeerd naar overals-kategoriekwantificaties)

Bron: SCP (CV'04).

De veranderingen zijn iets groter als we kijken naar de samenhang tussen de indexen en de afzonderlijke clusters (tabel G.1). De samenhang van de index met de domeinen wonen, mobiliteit en bezit consumptiegoederen wordt kleiner als geen gewichten worden gebruikt, die met gezondheid groter. Dit komt waarschijnlijk doordat Overals uitgaat van de samenhang tussen indicatoren en domeinen: gezondheid hangt niet erg samen met de andere domeinen, terwijl wonen, mobiliteit en het bezit van consumptiegoederen sterk met elkaar samenhangen (het zijn de drie meest materiële clusters). Door deze samenhang niet meer mee te nemen krijgen de sterk samenhangende cluster een minder grote rol en voorheen minder sterke clusters een grotere rol.

De resultaten voor 'geengew 5' uit bovenstaande analyse, zijn eerder ook gedaan voor 1993 en 1997. Het gaat dus om een index waarbij de indicatoren bij elkaar worden opgeteld, maar waarbij per cluster wel is gedeeld door het maximaal aantal te behalen punten. Overigens is getracht de

index voor 1993 zoveel mogelijk te laten lijken op de index die we sinds 1997 gebruiken (in 1993 was nog geen sprake van domeinen en werd Overals niet gebruikt). De correlatie tussen beide methoden is in 1993 ook erg hoog 0,98. Verder valt op dat de verschillen in correlaties met de clusters niet bijster groot zijn (zie tabel G.2).

Tabel G.2 correlatie tussen leefsituatie-index 1993 oorspronkelijk (wel Overals gebruikt) en opgeteld zonder gewichten en de 8 clusters

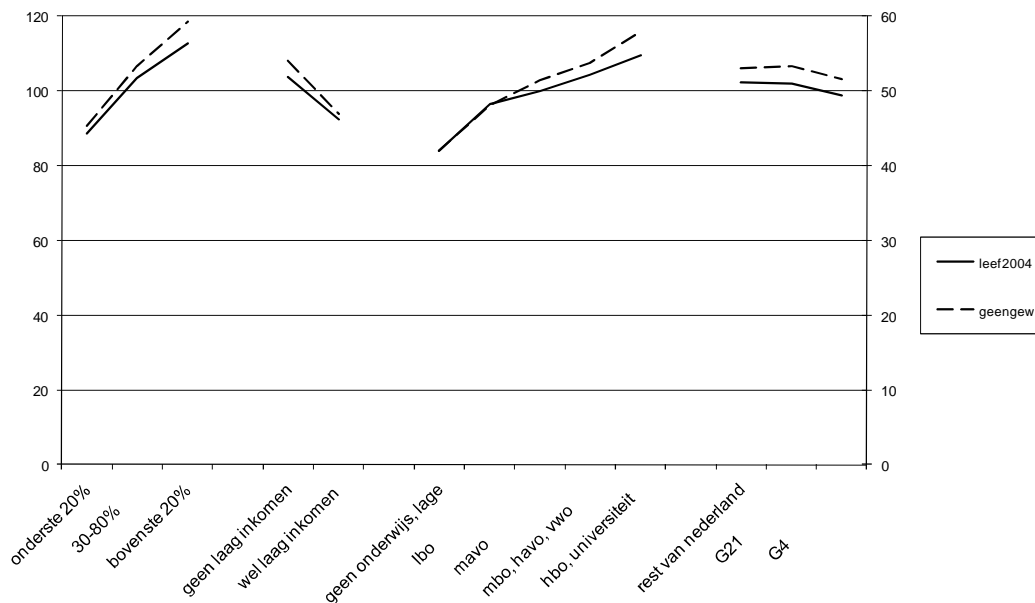
	woning	gezondheid	autobezit	duurzame consumptiegoederen	vakantie	participatie	sport	sociaal-culturele vrijetijdsbesteding
<b>1993</b>								
leefsituatie-index	0,50	0,33	0,62	0,64	0,58	0,48	0,56	0,42
index zonder gewichten	0,48	0,39	0,60	0,60	0,53	0,55	0,60	0,46
<b>1997</b>								
leefsituatie-index	0,44	0,38	0,51	0,65	0,64	0,71	0,56	0,58
index zonder gewichten	0,45	0,47	0,49	0,63	0,66	0,71	0,63	0,59

Bron: CBS DLO 1993; POLS 1997.

In 1997 is de samenhang 0,96; gezondheid en sportbeoefening spelen een belangrijker rol dan in de Overals-oplossing.

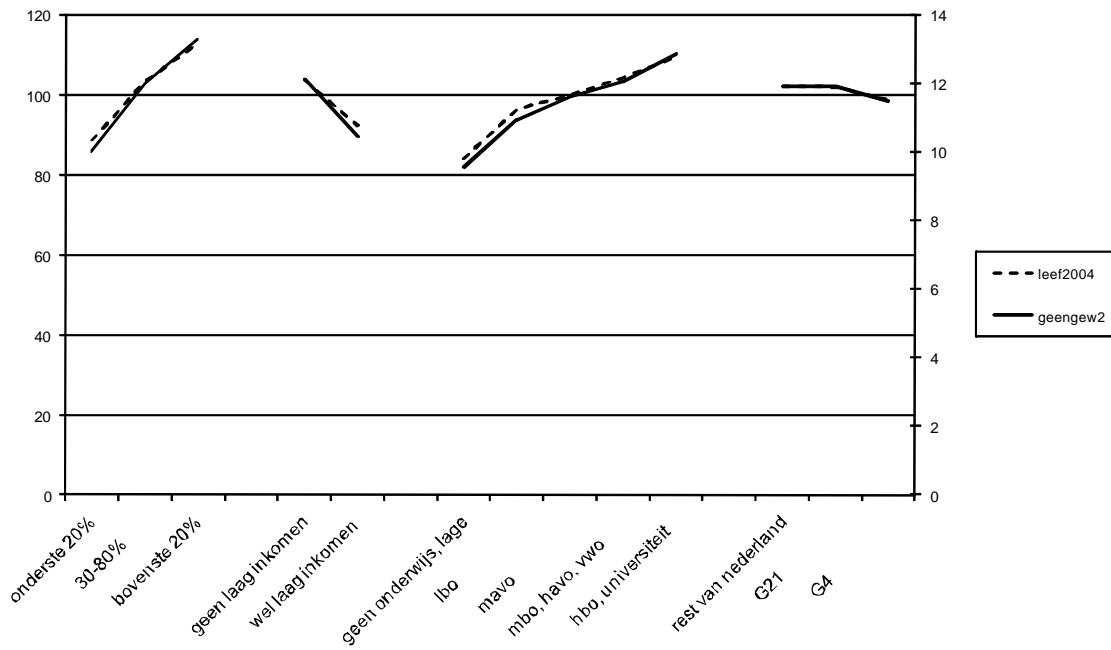
Ook uitgesplitst naar enkele achtergrondkenmerken zijn de verschillen zeer klein. Zie onderstaande figuren voor een vergelijking met Geengew, geengew2 en geengew5.

Figuur G.1: Leefsituatie-score en een indexscore gebaseerd op 'geengew' voor enkele sociale groepen (2004)



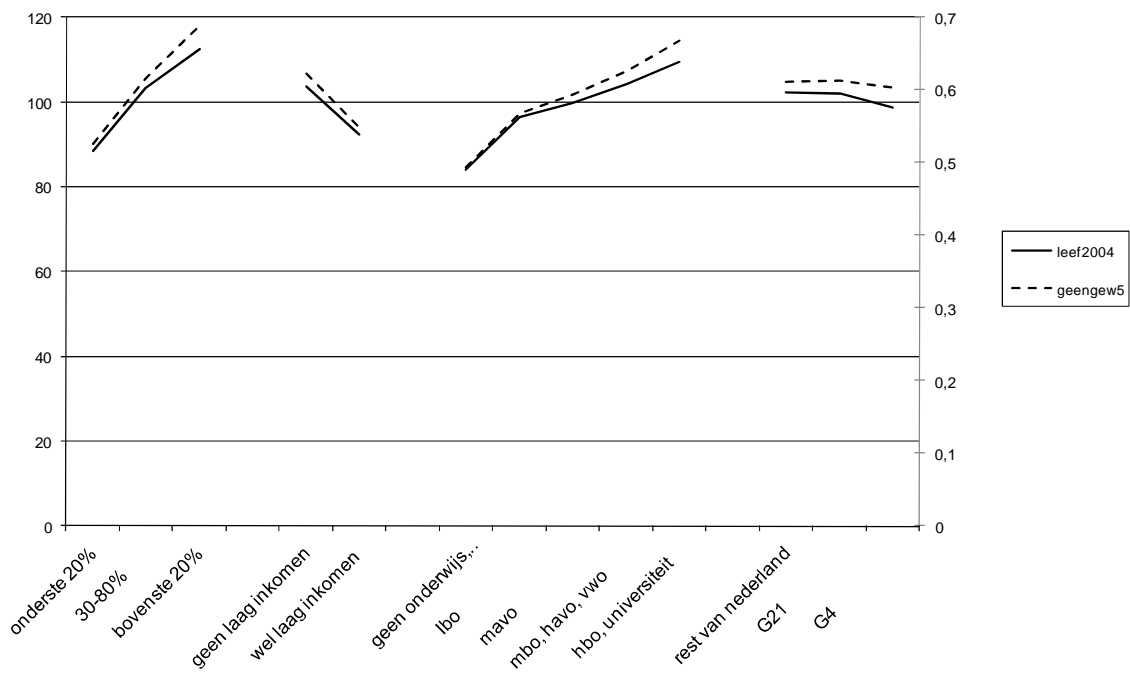
Bron: SCP (CV'04).

Figuur G.2: Leefsituatie-score en een indexscore gebaseerd op 'geengew2' voor enkele sociale groepen (2004)



Bron: SCP (CV'04).

Figuur G.3: Leefsituatie-score en een indexscore gebaseerd op 'geengew5' voor enkele sociale groepen (2004)



Bron: SCP (CV'04).