



UvA-DARE (Digital Academic Repository)

Burgers tegen beleid : een analyse van dynamiek in politieke betrokkenheid

Verhoeven, I.

Publication date

2009

Document Version

Final published version

[Link to publication](#)

Citation for published version (APA):

Verhoeven, I. (2009). *Burgers tegen beleid : een analyse van dynamiek in politieke betrokkenheid*. [, Universiteit van Amsterdam]. Aksant.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



BIJLAGE 1: METHODOLOGISCHE OVERWEGINGEN CLUSTERANALYSE

1. Databewerkingen clusteranalyse

Voor de empirische analyse van de vier stijlen van politieke betrokkenheid in hoofdstuk 1 heb ik gebruikgemaakt van gegevens uit het databestand Culturele Veranderingen in Nederland (CV) voor het jaar 1996. Clusteranalyse negeert respondenten met *missing values*. Omdat de techniek analyseert op antwoordpatronen van respondenten, is het voor representatieve uitspraken belangrijk om niet te veel respondenten te verliezen. Het CV bestand bevatte nogal wat *missing values*, waardoor het nodig was om databewerkingen uit te voeren. Ik heb een *missing value analysis* gedaan. Die ertoe leidde dat 31 respondenten met zeer veel *missing values* werden verwijderd uit de database (oorspronkelijke N= 2255, 18-74 jarigen). Vervolgens heb ik bij andere respondenten met *missing values* vervangende waarden berekend met gebruik van de *replace missing values* module. Ik alle variabelen nog een keer gehercodeerd tot twee categorieën, omdat clusteranalyse variabelen met meer antwoordcategorieën zwaarder laat wegen in de analyse (Hair et al. 1998: 489).

2. De validiteit van de vier gevonden clusters

De bewerkte variabelen zijn geanalyseerd met een hiërarchische clusteranalyse. Hierbij heb ik gebruikgemaakt van de Ward methode met *squared euclidian distance*, zonder aanvullende datatransformaties. Ik heb niet gewerkt met gestandaardiseerde variabelen (bijvoorbeeld met z-scores), omdat daarmee het ordinale karakter van de data verloren zou gaan. Dit wilde ik juist voorkomen, om goed interpreteerbare resultaten over te houden.

Clusteranalyse staat bekend om zijn descriptieve en tamelijk subjectieve karakter. Tamelijk subjectief, omdat de onderzoeker veel meer keuzevrijheid heeft in de analyse waardoor theoretische overwegingen veel zwaarder wegen. Om toch iets meer rechtvaardiging te geven voor een gevonden clusterverdeling dan alleen een theoretische, kan

er worden gewerkt met een zogenaamde *stopping rule*. Als er in het clusteringsproces bij een klein overgebleven aantal clusters een aanzienlijke stijging optreedt in de agglomeratiecoëfficiënt (de coëfficiënt van de samenvoeging van clusters), dan ligt het voor de hand om uit te gaan van de clusterverdeling die in de stap daaraan voorafgaand is gevormd. De reden daarvan is dat er dan geen belangrijke onderscheidende informatie tussen de clusters verloren gaat (Hair et al. 1998: 499). De *stopping rule* kan nog verder worden gevalideerd door deze eerst toe te passen op de gehele steekproef en vervolgens – bij wijze van experiment – door de oorspronkelijke steekproef te splitsen en dezelfde analyse daarop te draaien (Hair et al. 1998: 501). De bevindingen over de toepassing van deze procedure staan in tabel B1.1.

Tabel B1.1 Veranderingen hiërarchische clusteranalyse

	Aantal clusters	Agglomeratie coëfficiënt	Percentage verandering in de coëfficiënt #
Totale steekproef N=2223	6	7292,0	2,95
	5	7507,4	3,69
	4	7784,6	5,12
	3	8183,5	7,38
	2	8787,5	9,19
	1	9595,1	-
Gesplitste steekproef 1 N=1103	6	3620,3	3,16
	5	3738,3	3,64
	4	3874,3	4,81
	3	4060,5	8,27
	2	4396,2	8,45
	1	4768,0	-
Gesplitste steekproef 2 N=1119	6	3746,3	3,01
	5	3859,0	3,02
	4	3975,4	5,23
	3	4183,7	6,85
	2	4470,1	7,85
	1	4821,3	-

Het percentage van verandering in de agglomeratiecoëfficiënt is bijvoorbeeld als volgt berekend: $\frac{\text{coëfficiënt uit stap 5} - \text{coëfficiënt uit stap 6}}{\text{coëfficiënt uit stap 6}} * 100$.

Bron: Culturele veranderingen in Nederland (1975-2002).

De analyse laat zien dat op basis van de *stopping rule* voor de gehele steekproef de verdeling in vier clusters de meest optimale is. De verschuiving in het percentage is tussen stap 4 en stap 3 het grootst (van 5,12% naar 7,38%). Dit indiceert dat bij de stap van 4 naar 3 clusters een substantiële vermindering in onderscheidende informatie ontstaat. De test op de gesplitste steekproef levert voor de eerste steekproef een bevestiging van de verdeling in vier clusters op, terwijl voor de tweede steekproef een verdeling in 5 clusters optimaler zou zijn. Ondanks dit ene afwijkende resultaat kies ik voor een verdeling in

vier clusters, omdat deze voor de gehele steekproef het meest aannemelijk is en omdat deze overeenkomt met mijn theoretische overwegingen. De keuze voor een bepaalde clusterverdeling blijft uiteindelijk een arbitraire kwestie waarmee een onderzoeker greep probeert te krijgen op de diversiteit binnen de onderzoekspopulatie, in dit geval de diversiteit in politieke opvattingen en handelingen onder de Nederlandse bevolking.