

De groepsindeling van de ANP-CBS-beursindex nader bezien

Een clusteranalyse van aandelenkoersen

DRS. W. C. BOESCHOTEN* — PROF. DR. M. M. G. FASE**

Indices van beurskoersen zijn een veel gebruikt middel om de gang van zaken in het bedrijfsleven te beoordelen. De ANP-CBS-beursindex is een gewogen gemiddelde van zes groepen fondsen die naar sector zijn ingedeeld. In dit artikel worden dezelfde fondsen gegroepeerd aan de hand van hun koersverloop in het verleden met behulp van clusteranalyse. Op grond van hun analyse stellen de auteurs de vraag of het CBS de thans gebezigde groepsindeling niet beter kan vervangen door een gewogen indexcijfer van de beschouwde aandelenkoersen.

Inleiding en opzet

De eerste verwezenlijking van het idee om de koersontwikkeling van een groep aandelen in een kengetal weer te geven dateert van 1884. In dit jaar publiceerde Charles Henry Dow in de *Customer's Afternoon Letter*, de voorloper van de hedendaagse *Wall Street Journal*, het eerste beursindexcijfer. De huidige Dow Jones beursindexcijfers in de *Wall Street Journal*, waarvan het dagelijks gepubliceerde Dow Jones Industrial Average het bekendste is, dragen nog steeds de naam van hun geestelijke vader.

De verschillende beursindices waarmee de stemming ter beurze voor bepaalde groepen van aandelen wordt aangegeven, genieten een levendige belangstelling. Van de vele toepassingen noemen we:

- gebruik van de beursindex als nationale of zelfs internationale conjunctuurindicator 1);
- de beursindex als macro-economische indicator voor de waarde van het in de moderne portefeuille-theorie soms gebruikte reële kapitaal;
- de beursindex als voorspeller van de toekomstige ontwikkelingen op de aandelenmarkt;
- vergelijking van het koersbeloop van de eigen effectenportefeuille en dat op de markt uitgedrukt in het beursindexcijfer.

In Nederland is de ANP-CBS-beursindex de bekendste. Deze, thans uit 55 individuele aandelenkoersen samengestelde index, wordt sinds 1 april 1955 dagelijks door het Centraal Bureau voor de Statistiek, in samenwerking met het Algemeen Nederlands Persbureau (ANP) en de Vereniging voor de Effectenhandel, gepubliceerd.

Voor de berekening van deze beursindex worden de beschouwde fondsen verdeeld over zes groepen. In elke groep zijn voor de koersontwikkeling in de desbetreffende groep representatief geachte fondsen bijeengebracht. Daarbij is een sectorindeling gehanteerd die de soort bedrijven tot uitgangspunt heeft.

Een directere classificatiemethode is de clusteranalyse. Met deze, bij uitstek voor classificatie ontworpen methode kunnen de aandelen direct aan de hand van hun koersverloop in het verleden worden gegroepeerd. Dit artikel brengt verslag uit van een clusteranalyse op aandelenkoersen uit de ANP-CBS-beursindex. De analyse is verricht om na te gaan of de resulterende groepering correspondeert met de door het CBS gebezigde classificatie.

De berekening van het ANP-CBS-beursindexcijfer en het doel van ons onderzoek

Tot 1 januari 1975 werden de 54 bijgehouden aandelenkoersen verdeeld over vijf groepen. Met ingang van die datum is door het CBS de basisperiode verlegd van 1963 naar 1970. Tegelijkertijd is de toegepaste groepsindeling herzien. Daarbij werd ten eerste de groep „scheepvaart”, uitgebreid met de KLM tot de groep „scheep- en luchtvaart” waarmee het totale aantal fondsen op 55 kwam. Gelet op de toenemende discrepantie tussen de koersontwikkeling van bankaandelen en verzekeringsaandelen werd voorts met ingang van diezelfde datum, de oorspronkelijke groep „bank- en verzekeringswezen” in twee afzonderlijke groepen gesplitst. Het aantal groepen kwam daarmee op zes.

Tabel 1. Groepsindeling van aandelenkoersen vóór en na 1 januari 1975

Groepen	Aantal fondsen		Gewichten in procenten	
	vóór 1-1-1975	na 1-1-1975	vóór 1-1-1975	na 1-1-1975
Internationale concerns	5	5	40	48
Industrie	25	25	36	39
Scheep- en luchtvaart	4	5	3	4
Bankwezen	8	4	5	5
Verzekeringswezen	4	4	1	1
Handel e.a.	12	12	16	3
Totaal	54	55	100	100

Tabel 1 vermeldt de groepsindeling en de wegingsfactoren van vóór en na 1-1-1975. De algemene index wordt berekend als gewogen gemiddelde van de onderscheiden groepsindices. De wegingsfactoren worden ontleend aan de reële omzetten ter beurze in het basisjaar.

De berekeningswijze van de ANP-CBS-index is niet onbesproken gebleven. Een eerste punt van kritiek is de correctie voor kapitaalmutaties (ten gevolge van een emissie, een bonus

*Medewerker van de groep wetenschappelijke onderzoekingen en econometrie bij De Nederlandsche Bank NV.

**Adjunct-onderdirecteur bij De Nederlandsche Bank NV en buitengewoon hoogleraar aan de Erasmus Universiteit Rotterdam.

1) Vgl. M. M. G. Fase en E. W. van den Berg, Tussen Damrak en de rest van de wereld, *Maandblad voor accountancy en bedrijfshuishoudkunde*, jg., 53 nr. 4, april 1979, blz. 174-189.

of een stockdividend) die de invloed hiervan op de beursindex beoogt te elimineren. Deze correcties kunnen op de lange termijn, aldus het CBS, vertekeningen in de index tweewegbrengen 2). Buijs wijst er op dat dit ook op de korte termijn geldt en stelt, evenals Corver en Out, een alternatieve wijze van corrigeren voor 3).

Een tweede punt van kritiek betreft de gehanteerde wegingsmethode. Zo wekt het volgens Corver en Out bevreemding dat er per groep wel, maar binnen de groep per aandeel afzonderlijk niet gewogen wordt, met als gevolg dat elk aandeel binnen de groep even zwaar meetelt. Buijs merkt echter op dat het nemen van de ongewogen gemiddelden binnen de groep alleszins begrijpelijk is gezien de problemen die de keuze van de gewichten binnen de groepen zou opleveren (moet men kiezen voor beursomzet, eigen vermogen of grootte van het aandelenkapitaal?). Bovendien zouden de gekozen wegingsfactoren wellicht nog meer dan die voor de groepen aan veroudering onderhevig zijn. Verder heeft Fase vastgesteld dat een koersindex waarbij weging van de groepsindices achterwege wordt gelaten, voor de korte termijn slechts in geringe mate afwijkt van een gewogen index 4). Voorts blijkt uit Fase's onderzoek dat de wegingsproblematiek kan worden omzeild door met behulp van principale-componentenanalyse een index te ontwikkelen. Hier is de keuze van de gewichten een uitvloeisel van de analyse.

Een derde en laatste punt van kritiek betreft de gehanteerde groepsindeling. Gelet op de berekeningswijze van het indexcijfer kan men zich afvragen of deze sectorindeling afgemeten aan de variabiliteit voldoende homogeen is. De resultaten die Wolf 5) vindt m.b.v. een factoranalyse duiden op een ontkenkend antwoord. Dit artikel gaat met behulp van clusteranalyse na of de door het CBS gehanteerde sectorindeling steun vindt in de uitkomsten van deze statistische classificatiemethode 6).

Methode van analyse en gebruikte gegevens

Clusteranalyse is een techniek om een verzameling van objecten (onderzoekseenheden, variabelen) op te delen in een niet noodzakelijk van te voren bekend aantal, zo homogeen mogelijke, deelverzamelingen (clusters). De classificatie vindt plaats op grond van kenmerken (eigenschappen, waarnemingen) gemeten of waargenomen aan deze objecten. Men beoogt daarbij een zodanige opdeling te geven dat twee willekeurige objecten binnen een cluster veel meer gemeenschappelijk hebben dan twee objecten uit twee willekeurige, verschillende clusters.

Een klassiek voorbeeld is de classificatie van het planten- en dierenrijk van Linnaeus. Op grond van een groot aantal voor elke plant of dier vastgelegde eigenschappen werden er, aanvankelijk tamelijk impliciet, classificaties gemaakt. In den beginne was taxonomie als het ware meer kunst dan kunde. Gaandeweg werden echter objectievere criteria ontwikkeld. Aldus heeft de biologie veel bijgedragen tot de ontwikkeling van de clusteranalyse. De eerste omschrijving en definitie van wat tegenwoordig onder clusteranalyse wordt verstaan, werden geformuleerd door Tryon in 1939. Sindsdien zijn er verschillende technieken voorgesteld. Deze technieken lopen in hun benaderingswijze nogal uiteen. Zo komt Everitt tot vijf verschillende categorieën van technieken, die bij toepassing lang niet altijd dezelfde uitkomsten blijken op te leveren 7). Daarbij is het zeer moeilijk te bepalen of één methode beter is dan een andere. We beperken ons hier tot een beknopte beschrijving van de (agglomeratieve) hiërarchische clustermethoden, waartoe de in deze studie toegepaste methode van Ward behoort 8).

Een hiërarchische clustermethode gaat uit van een situatie waarin alle te clusteren objecten — in dit geval fondsen — in eerste instantie als afzonderlijke en uit één object bestaande clusters worden beschouwd. Bij iedere stap worden twee op grond van hun waarnemingen als meest nabij of meest gelijk

beschouwde clusters samengevoegd. Bij de eerste stap zijn dit altijd twee objecten, maar bij latere stappen kunnen even goed een object en een uit verscheidene objecten bestaande cluster, of ook twee uit verscheidene objecten bestaande clusters worden samengevoegd. Bij n te clusteren objecten resteren er na de eerste stap nog $(n-1)$ clusters, na de tweede stap $(n-2)$ clusters totdat er na $(n-1)$ stappen één cluster overblijft dat alle n objecten bevat.

Het verschil tussen de verschillende hiërarchische clustermethoden schuilt in de wijze waarop het nabijheids- of gelijkheids criterium wordt geconcretiseerd. Bij Wards methode worden steeds die clusters samengevoegd, wier samenvoeging gepaard gaat met de kleinste toename van de residuele kwadratensom, de som van de afstanden van de objecten tot het zwaartepunt van de cluster waartoe zij behoren. De methode lijkt zeer geschikt voor de classificatie van fondsen aan de hand van hun koerswaarden mits deze eerst op juiste wijze worden getransformeerd. Immers, wil men de ontwikkeling ter beurze van de waarde van een groep aandelen met een indexcijfer weergeven, dan is een minimale variantie binnen die groep gewenst 9).

Na afloop van het clusterproces moet het werkelijke aantal clusters worden bepaald waarover men de n objecten wil verdelen. Daarvoor is geen algemene regel te geven. Vaak wordt de keuze gemaakt aan de hand van de criteriumwaarde (bij Ward is dit, zoals we zagen, de totale residuele kwadratensom binnen de clusters). Men kiest dan de opdeling die voorafgaat aan een plotseling zeer grote toename van deze criteriumwaarde. Het is echter ook mogelijk dat men van te voren al een idee heeft van het aantal clusters.

Het boven beschreven hiërarchische clusterproces kan grafisch worden geïllustreerd met een zogenaamd dendrogram (boomstructuur-diagram). Bij wijze van voorbeeld is in figuur 1 een dergelijk dendrogram voor een fictieve clustering van zeven objecten weergegeven.

Er zijn in dit onderzoek twee deelperioden beschouwd, te weten de periode 1973-1975 en de periode 1977-1979. De daarbij gebezigde reeksen zijn weekmiddenkoersen van de aandelen die ten grondslag liggen aan de algemene ANP-CBS-beursindex met uitzondering van die aandelen waarvoor ten gevolge van discontinuïteiten in de beschouwde perioden

2) Voor vergelijkingen over langere periodes beveelt het CBS aan om de eveneens door het CBS berekende beurswaarde-index te gebruiken. Zie hiervoor CBS, ANP-CBS beursindex op nieuwe basis, *Maandstatistiek Financiering*, jg. 23, nr. 1, januari 1975, blz. 30-33 en ook E. F. Rietzschel, ANP-CBS-beursindex op een nieuwe basis, *Beursplein* 5, januari 1975, blz. 7-8 en blz. 17-18.

3) A. Buijs, Indexcijfers voor de effectenbeurs, *Bank- en Effectenbedrijf*, maart 1974, blz. 95-100. R. Corver en B. Out, Beursindexcijfers, *Bank- en Effectenbedrijf*, september 1980, blz. 293-302.

4) M. M. G. Fase, Een index voor het koersverloop van aandelen: een hoofdcomponentenanalyse, *Bank- en Effectenbedrijf*, november 1976, blz. 409-416.

5) H. W. Wolf, *Koersfluctuaties op de Amsterdamse effectenbeurs en de groepsindeling van het CBS: een factoranalytisch onderzoek*, januari 1978.

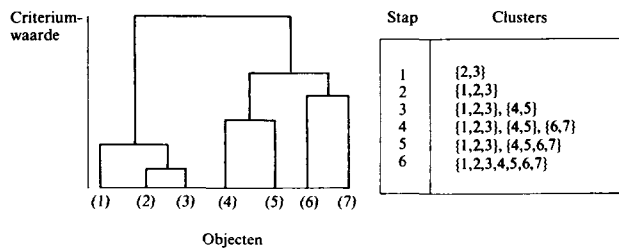
6) Het is ook mogelijk om de clusteranalyse na correctie van algemene invloeden toe te passen. Zo voert King een vorm van clusteranalyse uit op een residuele correlatiematrix die resteert na eliminatie van de door hem genoemde „marktfactor”. Zie hiervoor B. B. F. King, Market and industry factors in stock price behaviour, *The Journal of Business*, januari 1966, blz. 129-190.

7) B. Everitt, *Cluster analysis*, Heinemann Educational Books Ltd, Londen, 1974. Zie ook B. S. Duran en P. L. Odell, *Cluster analysis; a survey*, Berlijn, 1974 en W. D. Fisher, *Clustering and aggregation in economics*, Baltimore, 1969. Zeer verdienstelijk is ook J. G. Bethlehem, *Handleiding voor hiërarchische clusteranalyse*, Rapport SN5/76, Mathematisch Centrum, Amsterdam, 1976.

8) J. H. Ward, Hierarchical grouping to optimize an objective function, *Journal of the American Statistical Association*, jg. 58, 1963, blz. 236-244.

9) Men kan zich afvragen of een methode waarbij gebruik wordt gemaakt van een associatiemaat, zoals bijvoorbeeld een correlatiecoëfficiënt in plaats van een (Euclidische) afstandmaat, uiteindelijk misschien toch niet de voorkeur verdient. Zie hiervoor ook J. G. Bethlehem, op. cit., blz. 50.

Figuur 1. Voorbeeld van een dendrogram



geen volledige reeksen beschikbaar zijn. De gemiddelden en standaarddeviaties van de desbetreffende aandelenkoersen staan voor beide deelperioden in tabel 2 weergegeven. Voor de periode 1973-1975 zijn voor alle 54, oorspronkelijk opgenomen aandelen complete reeksen van koersen beschikbaar. In de drie gevallen dat er vervanging van fondsen heeft

plaatsgevonden, is de reeks van het vervangende fonds op basis van de gemiddelde koersverhouding op de twee voorafgaande woensdagen gekoppeld aan de oude reeks.

De koers van aandelen KLM, pas per 1 januari 1975 in het ANP-CBS-beursindexcijfer opgenomen, is in de eerste deelperiode buiten beschouwing gelaten. Van de oorspronkelijke 55 aandelenkoersen uit de tweede deelperiode vallen er twee af. De koersen van de aandelen Holland-Amerika Lijn en Koninklijke Scholten-Honig zijn namelijk respectievelijk per 23 september 1977 en 6 maart 1978 niet meer in het beursindexcijfer opgenomen, zodat er in de tweede deelperiode uiteindelijk sprake is van een groep van 53 verschillende fondsen.

Uitkomsten van de clusteranalyse

Alvorens tot clustering van de aandelen over te gaan, moeten de koerswaarden eerst zodanig worden getransfor-

Tabel 2. Gemiddelden en standaarddeviaties van de aandelenkoersen

Nr.	Sector	Naam bedrijf	1973 - 1975		1977 - 1979	
			gemiddelde	standaarddeviatie	gemiddelde	standaarddeviatie
1	Internationals	AKZO	55	15	29	3
2	"	Hoogovens	65	11	32	5
3	"	Kon. Ned. Petroleum	93	18	136	9
4	"	Philips	69	24	26	2
5	"	Unilever	113	23	123	5
6	Industrie	ACF	1.147	125	144	41
7	"	Berkel's Patent	303	75	99	11
8	"	Lucas-Bols	1.038	248	69	5
9	"	Bos Kalis Groep	1.033	194	117	10
10	"	Ver. Bedr. Bredero	416	77	267	22
11	"	Bührmann - Tetterode	376	49	69	5
12	"	Centrale Suiker Mij./Nutricia (1-1-1975)	715	140	42	6
13	"	Desseaux Tapijtfabriek	392	107	50	6
14	"	Fokker Vliegtuigenfabriek	177	29	26	3
15	"	Gamma Holding	216	60	34	6
16	"	Gist-Brocades	669	67	42	8
17	"	Océ - Van der Grinten	973	132	159	12
18	"	Heinekens Bierbrouwerij	805	275	102	16
19	"	Holec Samenw. Elec. Fabr.	229	52	135	21
20	"	Holl. Betron Groep	317	67	98	13
21	"	IHC Holland/Caland Holding (19-12-1979)	474	243	16	3
22	"	Kon. Ned. Papierfabriek	287	52	43	6
23	"	Meelfabriek Ned. Bakkerijen	253	85	59	10
24	"	Chem. Fabriek Naarden	567	68	29	10
25	"	Nijverdal - Ten Cate	80	19	51	14
26	"	Rijn-Schelde-Verolme	244	45	58	16
26	"	Kon. Scholten - Honig (tot 6-3-1978)	261	42	—	—
27	"	Twentsche kabelfabriek	383	72	269	19
28	"	Verenigde Machinefabrieken	164	17	51	15
29*	"	Wessanen's Kon. Fabrieken	562	135	65	9
29	Scheep- en Luchtvaart	Holland-Amerika Lijn (tot 23-9-1977)	79	24	—	—
30	"	KNSM Groep	149	30	95	12
31	"	Van Ommeren	312	47	161	27
32	"	Ned. Scheepvaart Unie/Ned. Lloyd Groep	249	37	93	17
33	"	KLM (toegevoegd per 1-1-1975)	—	—	117	24
34	Banken	ABN	299	50	340	22
35	"	AMRO-Bank	354	46	72	4
36	"	Ned. Middenstandsbank	312	46	193	27
37	"	Mees en Hope Groep/Slavenburg (12-10-1975)	314	68	242	8
38	Verzekering	AMEV	582	60	80	14
39	"	AMFAS	462	100	104	11
40	"	Ennia	646	97	131	13
41	"	Nationale Nederlanden	698	94	105	9
42	Handel en andere	Ahold	483	166	98	11
43	"	Ceteco	171	10	220	23
44	"	Deli Mij.	146	31	117	11
45	"	Hagemeijer	504	90	71	18
46	"	Internatio-Müller	292	49	41	6
47	"	Kon. Bijenkorf Beh.	442	88	81	9
48	"	Kluwer Uitgeversmij.	1.034	218	90	10
49	"	Lindeteves-Jacoberg (cert.)/Pont Houhandel (25-5-1975)	258	21	223	26
50	"	O G E M Nat. Bez.	312	76	27	5
51	"	Pakhoek Holding	726	74	52	16
52	"	Ver. Ned. Uitg. Bedr. VB	558	178	92	8
53	"	Wijers	118	21	62	15

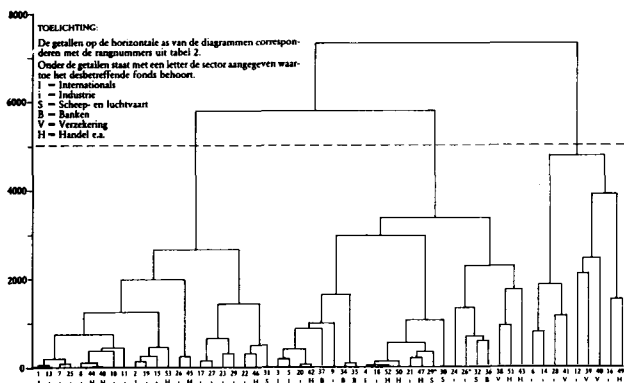
Toelichting bij de tabel:
 — indien een bepaald bedrijf door een ander bedrijf is vervangen wordt dat aangegeven met een schuine streep met daarna de naam van het nieuwe bedrijf en de datum waarop de verandering is ingegaan. In geval van IHC Holland is er uitsluitend sprake van een naamsverandering;
 — bedrijven waarvoor maar voor een deel van de beschouwde periode koersen beschikbaar zijn, worden (voor die periode) buiten beschouwing gelaten.

meer dat hun niveaus vergelijkbaar, d.w.z. van dezelfde orde van grootte zijn 10). Doet men dat niet dan zullen de fondsen voornamelijk op grond van hun gemiddelde koersniveaus worden gegroepeerd. Van de mogelijke transformaties om de invloed van de koersniveaus uit te schakelen zijn er twee beproefd. Ten eerste zijn de koerswaarden gestandaardiseerd. Daarnaast zijn procentuele mutaties (eigenlijk eerste logaritmische verschillen) genomen. Ten einde de stabiliteit van de uitkomsten te beoordelen zijn twee deelperioden, nl. 1973-1975 en 1977-1979 beschouwd (zie tabel 3). De deelperiode 1973-1975 komt overeen met de door Fase 11) beschouwde periode. De tweede, even lange periode 1977-1979 beslaat de drie meest recente jaren waarvoor koersen beschikbaar waren toen met deze studie een begin werd gemaakt.

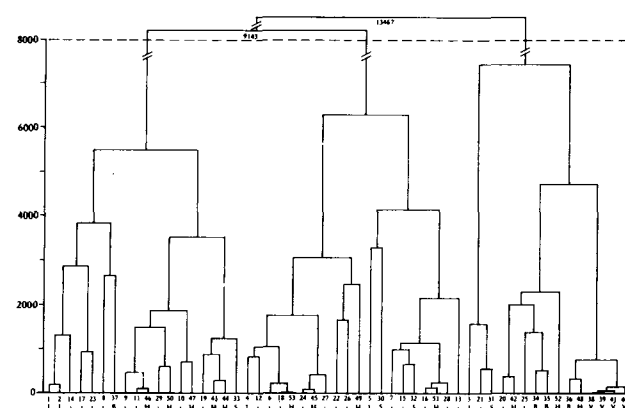
De uitkomsten voor de periode 1977-1979

Bij de clustering van de aandelen aan de hand van hun gestandaardiseerde koerswaarden lijkt de opdeling in drie clusters bij een criteriumwaarde van circa 5.000 het geschiktst (zie figuur 2). De eerste cluster, bestaande uit tweeëntwintig fondsen, bevat veertien van de vijfentwintig bedrijven die het CBS tot de industriector rekent. Daarnaast bestaat deze cluster nog uit twee „internationals” (AKZO en Hoogovens), de scheepvaartmaatschappij van Ommeren en vijf bedrijven uit de sector handel en anderen. De tweede cluster, met drieëntwintig fondsen, bevat alle banken, de verzekeringsmaatschappij AMEV, de „internationals” Koninklijke Nederlandse Petroleum, Philips en Unilever, verder zes bedrijven uit de industriector met een wellicht enigszins internationaal karakter en zes bedrijven uit de sector handel en andere. De derde cluster met negen fondsen omvat drie van de vier verzekeringsmaatschappijen, vijf bedrijven uit de industriector en één bedrijf uit de sector handel en andere.

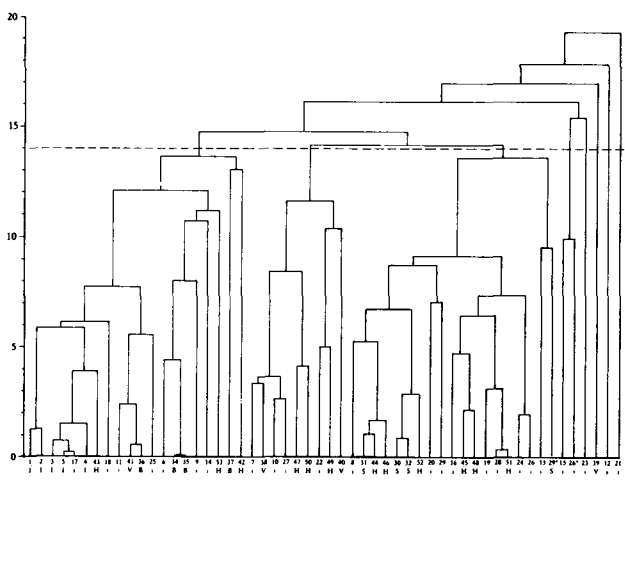
Figuur 2. Dendrogram van de clustering van aandelen op grond van gestandaardiseerde koerswaarden, 1973-1975



Figuur 3. Dendrogram van de clustering van aandelen op grond van gestandaardiseerde koerswaarden, 1977-1979



Figuur 4. Dendrogram van de clustering van aandelen op grond van eerste logaritmische verschillen van koerswaarden, 1973-1975



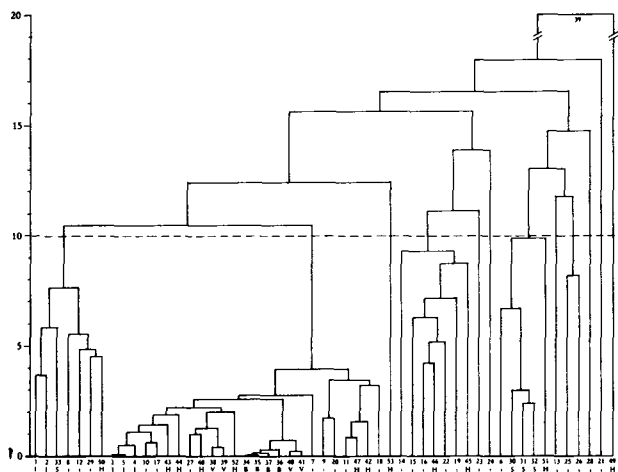
10) Gaarne betuigen we onze dank aan dr. G. A. van der Knaap van de Erasmus Universiteit Rotterdam voor het beschikbaar stellen van de programmatuur voor de hier gebezigde clusteranalyse van Ward.
11) Fase, op. cit., 1976.

Tabel 3. Opdeling van aandelen bij de twee clustervarianten in twee deelperioden

Gestandaardiseerd	1973 - 1975			1977 - 1979		
	I (22)	II (23)	III (9)	I (18)	II (20)	III (15)
Internationals (I)	1,2	3,4,5	6,12,14,16,28	1,2	4,5	3
Industrie (i)	7,8,10,11,13,15,17,19,22,23,25,26,27,29	9,18,20,21,24,26a	6,12,14,16,28	8,9,10,11,14,17,19,23,29	6,7,12,13,15,16,18,22,24,26,27,28	20,21,25
Scheep- en luchtvaart (S) .	31	29a,30,32		33	30,32	31
Banken (B)		34,35,36,37		37		34,35,36
Verzekering (V)		38	39,40,41			38,39,40,41
Handel e.a. (H)	44,45,46,48,53	42,43,47,50,51,52	49	43,44,46,47,50	45,49,51,53	42,48,52

Procentuele mutaties	1973 - 1975				1977 - 1979				
	I (20)	II (9)	III (17)	uitschieters (8)	I (7)	II (25)	III (8)	IV (5)	uitschieters, (8)
Internationals (I)	1,2,3,4,5				1,2	3,4,5			
Industrie (i)	6,9,11,14,17,18,25	7,10,22,27	8,16,19,20,24,26,28,29	12,13,15,21,23,26a	8,12,29	7,9,10,11,17,18,20,27	14,15,16,19,22,23	6	13,21,24,25,26,28
Scheep- en luchtvaart (S) .			30,31,32	29a	33				30,31,32
Banken (B)	34,35,36,37					34,35,36,37			
Verzekering (V)	41	38,40		39		38,39,40,41			
Handel e.a. (H)	42,43,53	47,49,50	44,45,46,48,51,52		50	42,43,44,47,48,52	45,46	51	49,53

Figuur 5. Dendrogram van de clustering van aandelen op grond van eerste logaritmische verschillen van koerswaarden, 1977-1979



Bij de clustering aan de hand van *procentuele mutaties* van de koersen blijkt iedere stap een nogal gelijkmatige toename van de bijbehorende criteriumwaarde op te leveren. De afwezigheid van sprongen in de criteriumwaarde doet vermoeden dat er van „natuurlijke” clusters nauwelijks sprake is (zie figuur 3). Bovendien is er een groot aantal uitschieters: aandelen van bedrijven die niet duidelijk tot een bepaalde cluster behoren. Indien niettemin een opdeling wordt verlangd, lijkt die bij een criteriumwaarde van circa 14 het aanvaardbaarst. Er resulteren dan drie grote clusters met daarnaast nog zes uitschieters. In de *eerste cluster* van twintig fondsen bevinden zich alle „internationals” en banken. In tegenstelling tot de „internationals” komen de banken, met name de Nederlandse Middenstandsbank en de Mees en Hope Groep, pas in een laat stadium bij elkaar in één groep. In deze eerste cluster bevinden zich verder zeven bedrijven uit de industriële sector, drie bedrijven uit de sector handel en andere en de Nationale Nederlanden als enige verzekeringsmaatschappij. De AMEV en ENNIA vormen met vier industrieën en drie bedrijven uit de sector handel en andere sectoren de *tweede cluster* van negen fondsen. De *derde cluster*, bestaande uit zeventien fondsen, bevat drie van de vier scheepvaartmaatschappijen, acht bedrijven uit de industrie en zes bedrijven uit de categorie handel en andere. De uitschieters zijn zes industriebedrijven waaronder KSH, de HAL en de verzekeringsmaatschappij AMFAS. Twee van deze uitschieters, namelijk de fondsen HAL en KSH, zijn vanaf respectievelijk 1977 en 1979 niet meer in het algemene beursindexcijfer opgenomen.

Bij vergelijking van de twee gevonden groeperingen onderling blijkt er met de sectorindeling van het CBS nauwelijks overeenkomst te zijn. Alleen de banken komen in beide gevallen in één groep terecht. Voor het overige lijken de twee gevonden clusterstructuren geen duidelijke systematiek te bezitten. Het feit dat bedrijven uit een zelfde sector meestal over twee of meer clusters verspreid zijn, weersprekt de door het CBS aangehouden sectorindeling. Behalve de banken die bij beide varianten in één cluster terechtkomen, vormen de „internationals”, bij de tweede variant een gedeeltelijke uitzondering.

De uitkomsten voor de periode 1977-1979

Bij clustering van de aandelen naar gestandaardiseerde koerswaarden lijkt een opdeling in drie groepen bij een criteriumwaarde van circa 8.000 de meest aangewezen (zie figuur 4). De *eerste cluster* telt achttien fondsen: de „internationals” AKZO en Hoogovens, Slavenburg als enige bank, KLM, een groep van negen bedrijven uit de industriële sector en ten slotte vijf bedrijven uit de sector handel en andere. De *tweede cluster*, met twintig fondsen, bevat Philips en Unilever,

de scheepvaartmaatschappijen KNSM Groep en de Ned-Lloyd Groep, twaalf voor een groot deel overwegend binnlands georiënteerde industrieën en vier eveneens overwegend binnlands georiënteerde bedrijven uit de sector handel en andere. De *derde cluster* met vijftien fondsen bevat alle vier de verzekeringsmaatschappijen, drie van de vier banken, de „international” Koninklijke Nederlandse Petroleum, de scheepvaartmaatschappij Van Ommeren en verder drie bedrijven uit de industriële sector en drie bedrijven, waaronder de twee uitgeverijmaatschappijen uit de sector handel en andere.

De clustering aan de hand van de *procentuele mutaties* levert, net als voor de periode 1973-1975, geen duidelijke clusterstructuur op. Niettemin is een groepering gekozen bij de criteriumwaarde 10. Dit resulteert in vier clusters met daarenboven nog zes uitschieters. De *eerste cluster*, met zeven fondsen, bevat opnieuw AKZO en Hoogovens, KLM en daarnaast drie industriebedrijven en een bedrijf uit de sector handel en andere. De banken en verzekeringsmaatschappijen vallen, te zamen met de drie resterende „internationals”, acht bedrijven uit de industriële sector en zes bedrijven uit de sector handel en andere in de *tweede cluster* met vijftientig fondsen. De *derde cluster*, bestaande uit acht fondsen bevat uitsluitend bedrijven uit industrie en de sector handel en andere. De *vierde cluster*, met vijf fondsen, herbergt de drie scheepvaartmaatschappijen samen met ACF en Pakhoed Holding. De uitschieters worden gevormd door acht bedrijven uit de sectoren industrie en handel en andere.

Nadere analyse van de uitkomsten

Vergelijking van de uitkomsten uit 1977-1979 met die voor de periode 1973-1975 leert dat de clusterstructuren voor beide benaderingswijzen niet stabiel zijn tussen de perioden. Uitgaande van procentuele mutaties vormen de banken een hechte groep. Bij de indeling aan de hand van gestandaardiseerde koersen is dat in veel mindere mate het geval, met name omdat Slavenburg, voor het CBS de opvolger van Mees en Hope, buiten de „bankencluster” valt (zie figuur 4). Het verzekeringswezen daarentegen komt nu, anders dan in de eerste periode, bij beide meetwijzen in één cluster terecht, te zamen met respectievelijk drie van de vier en alle banken. De in 1975 doorgevoerde opsplitting van het bank- en verzekeringswezen was voor de periode 1977-1979, gezien de resultaten, niet zo nodig geweest. Evenmin wordt de toevoeging van de KLM aan de sector scheepvaart gerechtvaardigd door onze bevindingen.

Een interessante vraag is hoe gevoelig de gepubliceerde index is voor verandering van de gekozen classificatie. Om dit te onderzoeken is voor de periode 1977-1979 het verloop van een vijftal indices berekend, die gebaseerd zijn op verschillende classificaties 12).

Tabel 4 toont een uiteenlopende mate van paralleliteit tussen de verschillende indices. De voor de weging gekozen indeling blijkt van belang te zijn voor de beschrijving van het koersverloop op de korte termijn. Met name de (cluster)

Tabel 4. Correlatiematrix van enkele op basis van verschillende classificaties geconstrueerde koersindices, 1977-1979

Classificatiecriterium	1	2	3	4	5	Gemiddelde waarde	
Sectorindeling (ANP-CBS)	.1					70,2	
Clusterindeling o.g.v. gestandaardiseerde koersen	.2	0,89	1			73,0	
Clusterindeling o.g.v. procentuele koersmutaties	..	0,67	0,86	1		75,3	
Gewogen gemiddelde van de koersen4	0,98	0,84	0,61	1	73,8	
Ongewogen gemiddelde van de koersen5	0,82	0,97	0,96	0,76	1	75,7

12) Bij de berekening van de verschillende indices is geen rekening gehouden met kapitaalmutaties. De indices vallen hierdoor aanzienlijk lager uit dan wanneer er voor deze mutaties zou zijn gecorrigeerd.

indeling op grond van procentuele mutaties van de koerswaarden leidt tot een betrekkelijk sterk afwijkende index. Daarentegen valt er een sterke parallelliteit waar te nemen tussen de index op basis van de sectorindeling en de gewogen gemiddelde koers (index) waarbij de koersen van de fondsen afzonderlijk aan de hand van hun omzetten in het basisjaar zijn gewogen.

Conclusie

Bij dit onderzoek is nagegaan of een groepering van de aandelenkoersen door middel van clusteranalyse de door het CBS gehanteerde sectorindeling bevestigt. Dit is niet het geval. Alleen de banken blijken een hechte groep te vormen. Dit betekent dat indelen naar de aard van de bedrijfsactiviteit een andere groepering oplevert dan een mechanische classificatie. Het ontbreken van stabiliteit betekent dat andere factoren dan het vervaardigde eindprodukt wellicht een belangrijke rol spelen bij de groepsindeling. Hierbij kan gedacht worden aan de kwaliteit van de ondernemingsleiding

of de soort afnemers. Op grond van onze uitkomsten rijst in elk geval de vraag of het CBS de groepsindeling niet beter kan laten varen.

Een soortgelijke conclusie is indertijd ook getrokken door Wolf. Hij heeft, voortbouwend op eerder onderzoek van Fase, een factoranalyse uitgevoerd op de aandelenkoersen uit de periode 1975-1976¹³). Hij kwam tot de slotsom dat de sectorindeling van het CBS geen steun vindt bij zijn uitkomsten en dat het KLM-fonds ten onrechte samen met de scheepvaartfondsen in één groep is ingedeeld. Anderzijds vindt bij Wolf de indeling van het bankwezen in één groep geen bevestiging.

De invloed van de voor de weging gekozen indeling is over het algemeen vrij aanzienlijk. Er blijkt echter weinig verschil te zijn tussen het verloop van de index op basis van de sectorindeling en de index die direct uit het gewogen gemiddelde van de afzonderlijke aandelenkoersen wordt berekend.

W. C. Boeschoten
M. M. G. Fase

13) Wolf, op. cit., 1978; Fase, op. cit., 1976.