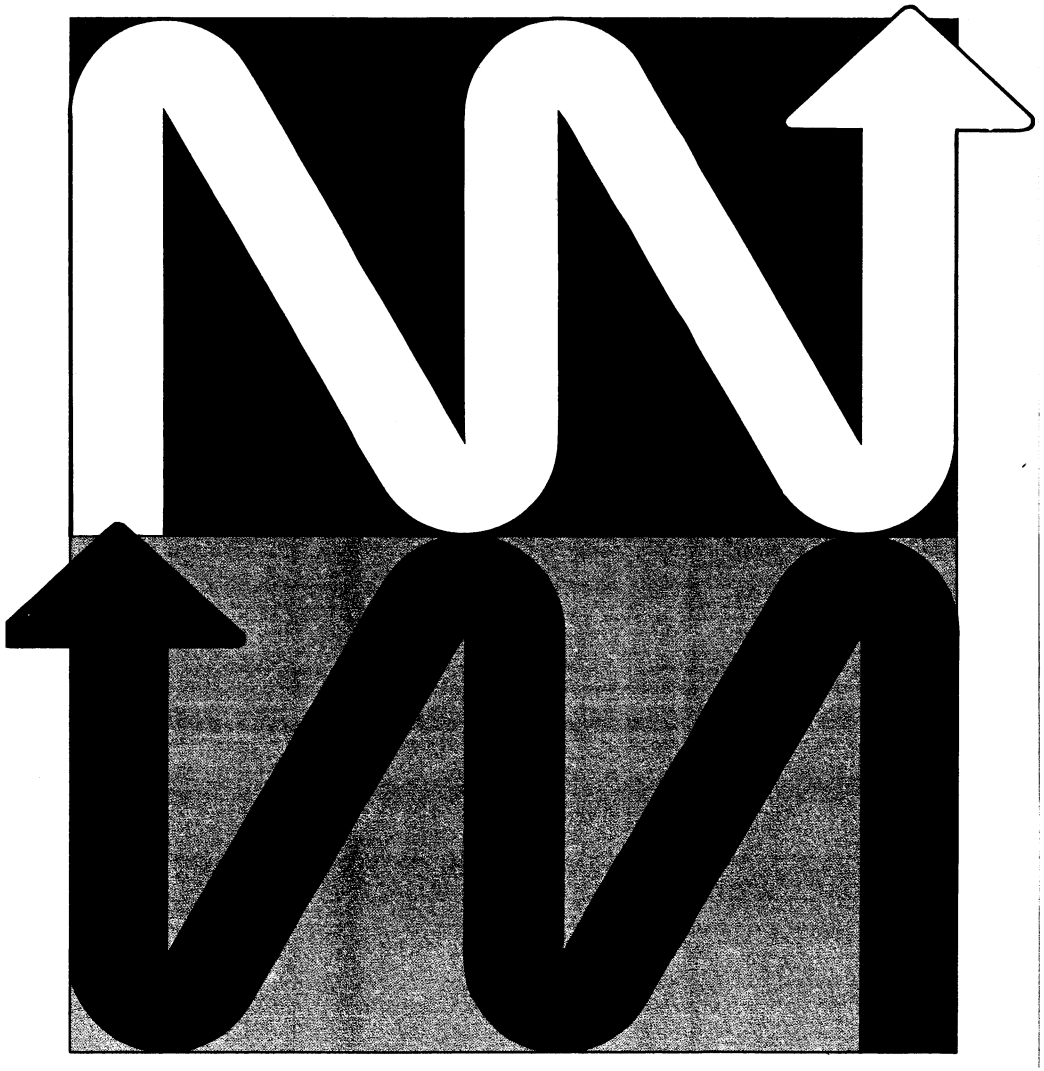
RECENTE ONTWIKKELINGEN
IN HET MARKTONDERZOEK
JAARBOEK '91-92
van de
Nederlandse Vereniging van Marktonderzoekers



2. *Multiniveau Analyse voor Markt- en Opinie-onderzoek*

J.J. HOX

SAMENVATTING

Sociaalwetenschappelijk onderzoek en marktonderzoek leiden beide regelmatig tot gegevens met een hiërarchische structuur. Voorbeelden zijn gegevens van respondenten verzameld op verschillende lokaties of op verschillende tijdstippen (respondenten genest binnen lokaties of tijdstippen), en herhaalde metingen (metingen genest binnen respondenten, zoals bijvoorbeeld conjuncte analyse). Deze zgn. *multiniveau* gegevens kunnen niet met standaard analysemodellen geanalyseerd worden. Recent ontwikkelde multiniveau-modellen en de bijbehorende software maken een adequate analyse van multiniveau gegevens eenvoudig. In deze bijdrage wordt het statistische model beschreven en een voorbeeld gepresenteerd van herhaalde gegevens bij een vignet-onderzoek.

1. INLEIDING

In sociaalwetenschappelijk en marktonderzoek worden onderzoekers regelmatig geconfronteerd met gegevens die een hiërarchische structuur vertonen. Dat wil zeggen; de gegevens kunnen worden gezien als een meertrapssteekproef uit verschillende hiërarchische niveaus, waarbij de variabelen op elk van deze niveaus gemeten kunnen zijn. Een duidelijk voorbeeld van hiërarchische gegevens is te vinden in het onderwijs: leerlingen bevinden zich in klassen, die weer deel uitmaken van scholen. Onderwijsonderzoek heeft vaak tot doel om verbanden te analyseren tussen school-, klas- en leerlingvariabelen, en het gebruik van *multiniveau modellen* ligt dan voor de hand. Na het verschijnen van een aantal basis-publikaties (De Leeuw & Kreft, 1986; Raudenbush & Bryk, 1986; Goldstein, 1987; Longford, 1987), en vooral van efficiënte computerprogramma's (Bryk et al., 1988; Longford, 1990; Prosser et al., 1990) heeft er binnen het onderwijs-onderzoek een explosie plaatsgevonden van toepassingen van multiniveau-modellen, die nog het meest doet denken aan de explosieve toename van het gebruik van structurele modellen na de publikaties van Jöreskog en het verschijnen van LISREL.

Behalve het schoolsysteem worden vele andere sociale systemen gekenmerkt door een hiërarchische structuur. Individuen zijn op verschillende wijzen georganiseerd in groepen, die weer onderdeel zijn van grotere (bijvoorbeeld geografische) eenheden. Verschillende vormen van steekproeftrekken produceren zulke gegevens, zoals het ondervragen van voorbijgangers in een beperkt aantal winkelstraten in een beperkt aantal steden. Zulke steekproeven kunnen niet beschouwd worden als eenvoudige aselecte

steekproeven, en dienen eigenlijk op een speciale wijze te worden geanalyseerd (zie het boek over complexe survey designs van Skinner et al., 1989 voor een algemene bespreking, en Goldstein & Silver's bijdrage daarin voor een aanpak via multiniveau analyse). Verder zijn er verschillende onderzoekopzetten die leiden tot multiniveau gegevens. Een goed voorbeeld is panel onderzoek, waar sprake is van een reeks van herhaalde waarnemingen genest binnen individuele respondenten (vgl. Goldstein, 1986, 1989; Bryk & Raudenbush, 1987). Dit betekent dat multiniveau modellen ook gebruikt kunnen worden bij andere onderzoektechnieken, zoals vignetten-onderzoek (Hox et al., 1991) en conjuncte analyse (vgl. Vriens & Wittink, 1990), waar de waarnemingen immers ook genest zijn binnen de respondenten. Een ander voorbeeld is gegevens van survey interviews, waarbij respondenten genest zijn binnen interviewers, en interviewers binnen survey organisaties. Nu er met het oog op de eenwording van Europa steeds meer surveys worden uitgevoerd in verschillende landen tegelijk, zouden multiniveau modellen gebruikt kunnen worden om systematische verschillen tussen interviewers, survey-organisaties en landen te analyseren en tot op zekere hoogte te controleren (vgl. Hox et al., 1991; Hox, 1992).
Hieronder volgt allereerst een korte beschrijving van het statistische model, daarna een uitgewerkt voorbeeld, en in de conclusie wordt teruggekomen op de toepasbaarheid binnen het marktonderzoek.

2. HET STATISTISCHE MODEL

Bij de analyse van multiniveau gegevens doen zich twee soorten problemen voor. Allereerst is er geen sprake van een enkelvoudige aselecte steekproef, maar van een meertraps- of cluster steekproef. Bij zulke steekproeven zijn de individuele waarnemingen doorgaans niet geheel onafhankelijk, zeker niet wanneer de clusters bestaande groepen zijn (zoals scholen of organisaties) waarbinnen mensen een gemeenschappelijke geschiedenis hebben. De standaard statistische toetsen zijn in dit soort gevallen niet van toepassing, en er zouden technieken moeten worden gebruikt die rekening houden met de complexe steekproeftrekking. Die technieken bestaan, maar omdat ze niet zijn ingebouwd in pakketten als SPSS en SAS worden ze weinig gebruikt.
Het tweede probleem ontstaat wanneer onderzoekers variabelen die op verschillende niveaus zijn gemeten met elkaar in verband willen brengen. De klassieke oplossing is disaggregatie: alle gegevens worden naar het laagste niveau gebracht en vervolgens 'gewoon' geanalyseerd. Dit is een slechte oplossing, zoals gemakkelijk valt in te zien wanneer we ons realiseren wat er gebeurt wanneer bijvoorbeeld een interviewervariabele wordt gedisaggregeerd naar het respondentniveau. Daarbij krijgt iedere respondent voor die interviewervariabele de waarde van zijn of haar interviewer, en daarna is het eenvoudig om de correlatie te berekenen van die interviewervariabele met een respondentvariabele. Maar dit behandelt die interviewervariabele feitelijk alsof er evenveel interviewers in het onderzoek betrokken zijn als er respondenten zijn, met iedere respondent zijn of haar eigen interviewer. Aangezien dit niet het geval is, zullen statistische toetsen het aantal onafhankelijke waarnemingen bij de interviewervariabele ernstig overschatten, en dus veel te kleine standaardfouten opleveren (voor een gedetailleerde bespreking van deze en andere bezwaren zie Kreft, 1987).
Het meest gebruikte en best uitgewerkte analysemodel voor multiniveau analyse is het multiniveau-equivalent van multipele regressie analyse. Er zijn ontwikkelingen in de

richting van multiniveau analyse van structurele modellen (Muthén, 1989; Muthén & Satorra, 1989), maar deze zijn nog onvoldoende uitgewerkt. Het hieronder gepresenteerde model is het multiniveau regressie model, dat in de literatuur bekend staat onder een aantal verschillende namen, zoals 'hierarchical linear model', 'random coëfficiënt model', en 'random component model'.

Het multiniveau regressie model neemt aan dat de afhankelijke variabele op het laagste niveau is gemeten, en dat er predictoren zijn op elk van de aanwezige niveaus. Conceptueel kan het model het best gezien worden als een hiërarchisch systeem van regressievergelijkingen. Bijvoorbeeld, stel dat gegevens zijn verzameld in J scholen, met in elke school een (verschillend) aantal leerlingen. We hebben op het leerlingniveau de afhankelijke variabele 'Schoolprestatie' (Y) en de predictor 'Sociale achtergrond' (X), en op het schoolniveau de predictor 'Schooltype' (Z), een schoolvariabele. Voor het verband tussen de afhankelijke variabele Y en de (leerlingniveau) predictor X kan per school een aparte regressievergelijking worden opgesteld, als volgt:

$$\mathbf{Y}_{ij} = \mathbf{\beta}_{0j} + \mathbf{\beta}_{1j}\,\mathbf{X}_{ij} + \mathbf{e}_{ij} \qquad (1)$$

In deze regressievergelijking (en alle volgende) worden alle random (stochastische) variabelen **vet** weergegeven. $\mathbf{e}_{ij}$ is de gebruikelijke error-term; net als in gewone multipele regressie wordt aangenomen dat de error term onafhankelijk is met een verwachting van nul en een variantie σ^2_j. Doorgaans wordt aangenomen dat de variantie van de error term **e** gelijk is in alle scholen, ofwel $\sigma^2_j = \sigma^2$. De intercept $\mathbf{\beta}_0$ en de hellingshoek $\mathbf{\beta}_1$ (slope) van de regressievergelijking zijn random over groepen, dat wil zeggen; we nemen aan dat de verschillende scholen verschillende regressiecoëfficiënten kunnen hebben (vandaar de naam 'random coëfficiënt model'). In ons voorbeeld houdt dit in dat we aannemen dat scholen een verschillend algemeen prestatieniveau hebben (de intercept $\mathbf{\beta}_0$ varieert), en dat we aannemen dat het verband tussen de sociale achtergrond van een leerling en zijn of haar schoolsucces op de ene school niet hetzelfde hoeft te zijn als op de andere school. Sommige scholen worden gekenmerkt door een grote regressiecoëfficiënt $\mathbf{\beta}_1$; in deze scholen heeft de sociale achtergrond van de leerling relatief veel invloed op de schoolprestatie, we zouden deze scholen kunnen omschrijven als selectief. Andere scholen worden gekenmerkt door een kleine regressiecoëfficiënt $\mathbf{\beta}_1$; in deze scholen heeft de sociale achtergrond van de leerling relatief weinig invloed op de schoolprestatie, en we zouden deze scholen kunnen omschrijven als egalitair.
De regressiecoëfficiënten $\mathbf{\beta}$ hebben dus een verdeling over de scholen, met een gemiddelde en een spreiding. De volgende stap in het model behandelt deze regressiecoëfficiënten in feite alsof het afhankelijke variabelen zijn in een nieuwe regressie vergelijking op het schoolniveau. Voor het verband tussen de random coëfficiënten $\mathbf{\beta}$ en de (schoolniveau) predictor Z kunnen de volgende twee regressie vergelijkingen worden opgesteld:

$$\mathbf{\beta}_{0j} = \gamma_{00} + \gamma_{01}\,\mathbf{Z}_j + \mathbf{U}_{0j} \qquad (2)$$

$$\mathbf{\beta}_{1j} = \gamma_{10} + \gamma_{11}\,\mathbf{Z}_j + \mathbf{U}_{1j} \qquad (3)$$

Vergelijking (2) stelt dat het algemene prestatieniveau van een school (de intercept $\mathbf{\beta}_{0j}$)

te voorspellen is uit het schooltype (Z). Vergelijking (3) stelt dat het *verband* tussen de sociale achtergrond van een leerling en zijn/haar schoolprestatie te voorspellen is uit het schooltype (Z). Met andere woorden, of een school selectief (grote regressiecoëfficiënt β_1) of egalitair (kleine regressiecoëfficiënt β_2) is wordt voorspeld uit de schoolvariabele 'schooltype'.
De $\mathbf{U}_j$ termen in vergelijking (2) en (3) zijn error termen op het schoolniveau; aangenomen wordt dat zij een verwachting hebben van nul en ongecorreleerd zijn met de individuele error term $\mathbf{e}_{ij}$. De variantie van $\mathbf{U}_0$ is σ^2_{U0}, en de variantie van $\mathbf{U}_1$ is σ^2_{U1}. In het algemene model is de covariantie tussen $\mathbf{U}_0$ en $\mathbf{U}_1$ *niet* nul. De (co)varianties van de error termen $\mathbf{U}_j$ op het tweede niveau zijn elementen van de covariantiematrix $\Omega_{(2)}$, waarbij het subscript (2) aangeeft dat dit een covariantiematrix is op het tweede niveau. (Bij drie- niveau modellen is er dus ook een covariantiematrix $\Omega_{(3)}$.)
Dit geheel kan worden geschreven als één complexe regressievergelijking door de vergelijkingen (2) en (3) te substitueren in vergelijking (1):

$$\mathbf{Y}_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} Z_j + \gamma_{11} Z_j X_{ij} + \mathbf{U}_{0j} + \mathbf{U}_{1j} X_{ij} + \mathbf{e}_{ij} \qquad (4)$$

In vergelijking (4) bevat het gedeelte $\gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} Z_j + \gamma_{11} Z_j X_{ij}$ alleen gefixeerde coëfficiënten; dit wordt dan ook het gefixeerde gedeelte genoemd. Het gedeelte $\mathbf{U}_{0j} + \mathbf{U}_{1j} X_{ij} + \mathbf{e}_{ij}$ bevat de error termen, en wordt daarom het random gedeelte genoemd.
Wanneer er veel predictoren zijn, en meer dan twee niveaus, dan wordt het bijzonder ingewikkeld om het statistische model te blijven zien als een opeenvolging van regressiecoëfficiënten op lagere niveaus die steeds weer verklaard worden door regressievergelijkingen met predictoren van de hogere niveaus. Veel onderzoekers zullen het eenvoudiger vinden om model (4) direct te interpreteren in termen van predictoren van verschillende niveaus en interacties tussen specifieke predictoren van verschillende niveaus. De gamma's in vergelijking (4) kunnen geïnterpreteerd worden als gewone regressiecoëfficiënten behorende bij de variabelen X en Z, en bij ZX, de interactie tussen X en Z. In ons voorbeeld: **Schoolprestatie** = γ_{00} + γ_{10} Sociale achtergrond + γ_{01} Schooltype + γ_{11} Schooltype * Sociale achtergrond + **errorterm**. Daarbij moet wel bedacht worden dat vergelijking (4) weliswaar lijkt op het gewone multipele regressie model, maar dat de errorterm veel ingewikkelder is, met in principe voor iedere regressiecoëfficiënt ß op het laagste niveau (en dus voor iedere predictor op het laagste niveau) een variantiecomponent **U** op het hogere niveau. Om statistisch correcte standaardfouten en significantietoetsen te krijgen is het belangrijk deze errorstructuur correct te specificeren.
Het statistische model voor vergelijking (4) levert schattingen op voor de standaardfouten van de parameters γ en voor de standaardfouten van de covarianties in de covariantiematrix $\Omega_{(2)}$. Het is dus mogelijk te toetsen of de verschillende effecten significant zijn. Net als bij multipele regressie zijn de verschillende effecten conditioneel op de predictoren in het model; wanneer een predictor wordt toegevoegd kan dit de hele vergelijking veranderen. Vergeleken met 'gewone' multipele regressie analyse zijn er nogal wat coëfficiënten en variantiecomponenten die geïnterpreteerd moeten worden. Verschillende analysestrategieën en de bijbehorende interpretaties worden beschreven door Kreft & Hox (1991). Een goede analysestrategie is om het model langzaam op te bouwen vanaf het individuele niveau, en bij elke stap te inspecteren welke coëfficiënten significant zijn, hoeveel variantie verklaard wordt, en hoeveel variantie in latere stappen nog verklaard kàn worden.

Zo'n strategie ziet er als volgt uit:

Stap 1. Analyseer een model *zonder predictoren* (het zgn. 'intercept only' model). In vergelijking (4) worden dan alle termen weggelaten waarin een predictor X of Z optreedt:

$$\mathbf{Y}_{ij} = \gamma_0 + \mathbf{U}_{0j} + \mathbf{e}_{ij} \qquad (5)$$

Het 'intercept only' model levert ons een schatting van de algemene intercept γ_{00} en twee variantieschattingen: σ^2 op het leerlingniveau en σ^2_{U0} op het schoolniveau. Uit deze gegevens kunnen we een schatting berekenen van de intraclass correlatie: $\eta=\sigma^2_{U0}/(\sigma^2_{U0}+\sigma^2)$. De intraclass correlatie geeft aan welk deel van de variantie van de afhankelijke variabele op het schoolniveau ligt.

Stap 2. Analyseer een model met alleen predictoren op *het laagste niveau* (doorgaans het individuele niveau), maar zonder de bijbehorende variantiecomponenten van de regressiegewichten $\mathbf{U}_{ij}$. In vergelijking (4) wordt dit:

$$\mathbf{Y}_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \mathbf{U}_{0j} + \mathbf{e}_{ij} \qquad (6)$$

Dit model levert schattingen op van de regressiecoëfficiënten voor de X variabelen. Het berekenen van de intraclass correlatie is niet zinvol meer, want de varianties σ^2 (op het individuele niveau) en σ^2_{U0} (op het schoolniveau) zijn conditioneel op de ingevoerde predictoren. Wel kunnen we de σ^2 in dit model vergelijken met de σ^2 van het model in de eerste stap, en hieruit zien hoeveel van de variantie op het leerlingniveau door de predictoren X verklaard wordt. Ook kunnen we de σ^2_{U0} in dit model vergelijken met de σ^2_{U0} van het model in de eerste stap, en hieruit zien hoeveel van de variantie op het schoolniveau door de leerlingvariabelen X verklaard wordt. Verschillen tussen scholen hoeven immers niet noodzakelijk af te hangen van schoolvariabelen. Een deel van de verschillen tussen scholen kan ook het gevolg zijn van verschillen tussen de scholen wat betreft de leerlingkenmerken.

Stap 3. Analyseer vervolgens een model met alleen predictoren op *het laagste niveau* (doorgaans het individuele niveau), maar nu mèt de bijbehorende variantiecomponenten $\mathbf{U}_{ij}$. In vergelijking (4) wordt dit:

$$\mathbf{Y}_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \mathbf{U}_{0j} + \mathbf{U}_{1j} X_{ij} + \mathbf{e}_{ij} \qquad (7)$$

met voor elke X_{ij} een regressiecoëfficiënt, en tevens voor elke X variabele een nieuwe variantieschatting $\mathbf{U}_j$, namelijk de variantie over de scholen van de bijbehorende regressiecoëfficiënt. Wanneer een bepaalde regressiecoëfficiënt of variantiecomponent niet significant is, kan deze uit het model worden verwijderd [1] (regresiecoëfficiënten die zelf niet significant zijn, maar wèl een significante variantiecomponent hebben, moeten in het model gehandhaafd blijven). Wanneer de variantiecomponenten in model (7) significant zijn, wordt de interpretatie van de hoeveelheid 'verklaarde variantie' onduidelijk, omdat er nu meerdere **U** termen zijn, en omdat de variantieschatting $\mathbf{U}_{0j}$ dan niet onafhankelijk is van de schaling van de predictoren. Ik ga daar hier niet verder op in (zie o.a. Bosker & Snijders, 1990; Kreft & Hox, 1991).

Stap 4. Voeg de predictoren van het schoolniveau toe, zonder interacties met predictoren van het leerlingniveau. Dit geeft ons de vergelijking:

$$\mathbf{Y}_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} Z_j + \mathbf{U}_{0j} + \mathbf{U}_{1j} X_{ij} + \mathbf{e}_{ij} \qquad (8)$$

Opnieuw kunnen we onderzoeken in hoeverre het toevoegen van de schoolniveau predictoren Z significante effecten heeft. De vraag is daarbij niet alleen of de regressiecoëfficiënt γ_{01} significant is, maar ook of de variantiecomponenten σ^2 en σ^2_u (zowel voor de intercept als voor de regressiecoëfficiënten van de leerlingvariabelen) kleiner zijn geworden, en wellicht vergeleken met de resultaten van stap 3 nu niet langer significant zijn. Opnieuw kunnen niet-significante regressiecoëfficiënten of variantiecomponenten verwijderd worden (waarbij regresiecoëfficiënten die zelf niet significant zijn, maar wel een significante variantiecomponent hebben, weer moeten worden gehandhaafd).

Stap 5. Wanneer het model in stap 4 nog significante variantiecomponenten bevat voor de *hellingen*, kunnen interacties tussen schoolvariabelen en leerlingvariabelen ($Z_j X_{ij}$) worden ingevoerd om deze te verklaren (vgl. vergelijking 3). We hebben dan het volledige model uit vergelijking (4). Als er na deze stap nog steeds significante variantiecomponenten over zijn, kunnen we alleen concluderen dat er nog onverklaarde systematische variatie tussen de scholen is (wat betreft intercept en/of hellingen) maar dat we in het betreffende onderzoek op het schoolniveau kennelijk niet over de juiste predictoren beschikken om deze te verklaren.

3. EEN VOORBEELD VAN EEN MULTINIVEAU ANALYSE

Het voorbeeld dat ik hier gebruik is een analyse van gegevens uit een *vignet-onderzoek* (Hermkens, 1983). Een vignet-onderzoek lijkt op de onderzoeksopzet van de conjuncte analyse (vgl. Vriens & Wittink, 1990). Aan de respondenten worden beschrijvingen van hypothetische objecten, personen of situaties voorgelegd, vignetten genaamd, en zij moeten die beoordelen op een aangegeven dimensie. De vignetten worden gekenmerkt door een aantal relevant geachte attributen, die door de onderzoeker gevarieerd worden. In dit voorbeeld gaat het om een onderzoek naar de rechtvaardigheid van inkomens van gezinnen. De taak van de respondent is om aan ieder vignet een rechtvaardig geacht gezinsinkomen toe te kennen. In de vignetten wordt een hypothetisch gezin beschreven, waarbij de omschrijvingen variëren wat betreft de attributen leeftijd, gezinsgrootte, opleiding, bron van inkomen (baan, uitkering), aantal inkomensbronnen (een- of tweeverdieners), en SES. Dat zijn zes attributen, en bij drie mogelijke waarden per attribuut zijn dat al 729 mogelijke vignetten. Het is dus niet haalbaar om alle mogelijke vignetten aan de respondenten voor te leggen. Bij vignet- onderzoek wordt dit probleem opgelost door aan iedere respondent een steekproef voor te leggen uit de mogelijke vignetten. Iedere respondent krijgt daarbij dus een andere verzameling vignetten voorgelegd. Het voordeel van deze opzet is dat een grote variëteit aan vignetten kan worden aangeboden, en dat het mogelijk is om zowel hoofd- als interactie-effecten te toetsen. [2]

Het zal duidelijk zijn dat hier sprake is van hiërarchische gegevens; er is een steekproef van respondenten, met voor iedere respondent een steekproef van vignetten. De vignetten zijn genest binnen de respondenten. Een multiniveau regressiemodel is op zijn plaats, waarbij in dit geval de vignetten het laagste niveau zijn (niveau 1) en de respondenten het hoogste (niveau 2).
In het voorbeeld zijn er zes vignet-variabelen. Er zijn zeven respondent-variabelen: leeftijd, gezinsgrootte, burgelijke status (gehuwd/samenwonend of alleen), opleiding, bron van inkomen, SES, en de hoogte van het eigen inkomen. Om de regressiecoëfficiënten vergelijkbaar te maken, zijn alle predictoren omgezet in standaardscores. Er zijn 6x7=42 mogelijke interacties. Om het aantal interacties in te perken, neem ik hier alleen de interacties op tussen *overeenkomstige* variabelen, zoals bijvoorbeeld vignet-opleiding en opleiding van de respondent. Deze hebben een eenvoudige interpretatie, namelijk dat de respondent zijn eigen situatie gebruikt als een 'ankerpunt' voor de beoordeling van de vignetten. (De interactie-variabelen zijn niet verder gestandaardiseerd, zie Jaccard et al., 1990). De afhankelijke variabele is het aan het vignet toegekende rechtvaardige inkomen, uitgedrukt in eenheden van duizend gulden.
Er zijn 106 respondenten, die ieder 16 vignetten hebben beoordeeld. Een aantal resultaten van de stapsgewijze modellering van hierboven staan in de volgende Tabel:

Tabel 1. Stapsgewijze resultaten.

Model:	σ^2	σ^2_{U0}	deviantie
1. Alleen Intercept	.80	.24	4591
2. + Vignet var., fixed	.51	.23	3864
3. + Vignet var., random	.30	.22	3495
4. + Resp. var.	.30	.18	3472
5. + Alleen sig. resp. v.	.30	.19	3478
6. + Interactie	.30	.19	3478

(σ^2 is de vignet variantie, σ^2_{U0} is de respondent variantie)

In de eerste stap blijkt dat de variantie van het toegekende inkomen kan worden opgesplitst in 77% variantie op vignet niveau en 23% op respondent niveau. In stap twee worden de vignet variabelen toegevoegd. Deze variabelen blijken alle significant te zijn; tezamen verklaren ze ,80-.51=.29 variantie. Aangezien de totale variantie op vignetniveau .80 is, kunnen we dit ook formuleren als: de vignetvariabelen verklaren 36% van de variantie die op vignetniveau aanwezig is. De vignetvariabelen verklaren nauwelijks variantie op het respondent niveau (4%). Dat is vanzelfsprekend, want de vignetten zijn over de respondenten gerandomiseerd, en er zijn dus nauwelijks systematische verschillen tussen de respondenten wat betreft het soort vignetten dat zij voorgelegd kregen. In de derde stap worden de vignet variabelen random gemaakt, dat wil zeggen dat de bijbehorende regressiecoëfficiënten mogen variëren tussen de respondenten. De deviantie daalt dan sterk, en alle variantiecomponenten blijken significant te zijn. De variantie op vignetniveau (σ^2) lijkt opnieuw sterk af te nemen. Dit is echter gezichtsbedrog. De correcte interpretatie is dat een deel van de variantie van de vignetten blijkt te bestaan uit verschillen in de hellingen van de regressielijnen tussen

verschillende respondenten. Er is als het ware alleen (error)variantie verplaatst; wanneer we voorspellingen zouden gaan uitrekenen zouden we ontdekken dat we in het model van stap drie niet beter kunnen voorspellen dan in stap twee. De voorspelling zou pas beter worden wanneer we de variaties in de hellingen tussen de respondenten zelf zouden kunnen voorspellen.
In stap vier worden de respondentvariabelen toegevoegd. In dit voorbeeld zijn deze niet alle significant. Daarom voegen we nog een vijfde stap toe, waarin alleen de significante respondentvariabelen zijn opgenomen. De significante respondentvariabelen (burgelijke status, gezinsgrootte en eigen inkomen) verklaren 13% van de aanwezige respondent-variantie. In de laatste stap worden de interacties opgenomen, deze blijken echter niet in staat de verschillende hellingen te verklaren.
Het uiteindelijke model bevat dus alle zes vignetvariabelen, drie respondentvariabelen, en géén interacties. Van de vignetvariantie wordt 36% verklaard (door de vignetvariabelen). Van de respondentvariantie wordt 17% verklaard: 4% door verschillen in de voorgelegde vignetten en 13% door de respondentvariabelen. Alle vignetvariabelen hebben significant verschillende hellingen, wat betekent dat de respondenten verschillende gewichten hechten aan de diverse attributen. We zijn met de hier beschikbare respondent-variabelen echter niet in staat deze laatste verschillen te verklaren.
De parameters van het model uit stap vijf staan in tabel 2:

Tabel 2. Resultaten model 5.

Variabele:	b-gewicht	var. comp.	sigma
Intercept	2.64	.19	.44
Vig.-leeftijd	.06	.01	.11
Vig.-gez. grootte	.14	.01	.10
Vig.-opleiding	.22	.06	.24
Vig.-bron ink.	.11	.01	.11
Vig.-aant. ink. bron	.24	.05	.22
Vig.-SES	.19	.04	.21
Resp.-gez. grootte	-.11		
Resp.-burg. status	.16		
Resp.-inkomen .18			

De eenheid van de afhankelijke variabele is in duizenden guldens. De predictoren zijn alle gestandaardiseerd, waardoor de regressiegewichten onderling eenvoudig vergeleken kunnen worden. Tabel 2 laat zien dat van de vignetvariabelen de opleiding en het aantal inkomensbronnen de belangrijkste invloed hebben. De waarde van de intercept van 2.64 betekent dus dat de respondenten aan het gemiddelde vignet 2640 gulden toekennen (netto). Eén standaard deviatie in leeftijd ouder betekent 60 gulden meer. Bij de respondentvariabelen is het eigen inkomen de belangrijkste voorspeller; respondenten met een één standaard afwijking hoger inkomen kennen gemiddeld 180 gulden méér toe.
De variantiecomponenten in Tabel 2 kunnen ook nog geïnterpreteerd worden. Zoals gebruikelijk kan daarvoor beter naar de standaard afwijking gekeken worden; deze staat in Tabel 2 onder 'sigma'. In alle gevallen is de standaard afwijking aanzienlijk vergeleken met de grootte van de bijbehorende regressiecoëfficiënt. Dat betekent dat de

beoordelingsstrategieën van de verschillende respondenten sterk uiteenlopen. In feite geldt voor alle vignetvariabelen dat, hoewel de gemiddelde regressiecoëfficiënt over alle respondenten positief is, er een enkele respondent kan zijn die zelfs een negatief regressiegewicht heeft!

4. CONCLUSIE

Veel onderzoeksgegevens hebben een hiërarchische structuur, en kunnen dus alleen met multiniveau modellen adequaat geanalyseerd worden. Het best ontwikkelde multiniveau model is het multiniveau regressie model, waarvoor ook goede software beschikbaar is. [3] Het gebruik van multiniveau analyse in plaats van disaggregatie van groepsvariabelen naar het individuele niveau gevolgd door gewone regressie analyse heeft twee belangrijke voordelen. Het eerste voordeel zou op zich al doorslaggevend moeten zijn: bij disaggregatie gevolgd door een 'gewone' analyse zijn de statistische toetsen volstrekt onbetrouwbaar (de literatuur geeft hier vele voorbeelden van, zie bijv. Kreft & Hox, 1991). Het tweede voordeel is dat het uitvoeren van een multiniveau analyse een goede manier is om interacties op het spoor te komen.
Hierboven is een voorbeeld gegeven van een analyse op vignetten data. Omdat iedere respondent daarbij een andere steekproef van vignetten krijgt voorgelegd, ligt multiniveau analyse hier voor de hand. Bij een 'gewone' conjuncte analyse, waarbij elke respondent dezelfde vignetten krijgt voorgelegd kan multiniveau analyse ook gebruikt worden, maar dan zijn de voordelen minder evident.
Multiniveau analyse opent een aantal mogelijkheden die langs andere weg moeilijker te verwezenlijken zijn. In de inleiding werd al gewezen op de analyse van complexe steekproeven, waaronder cluster- en tweetrapssteekproeven vallen, maar ook interviewer-onderzoek en gegevens afkomstig uit groepsgewijs afgenomen individuele interviews. Een van de aardigste toepassingen is het gebruik van een drie-niveau analyse bij herhaalde metingen. Stel dat een onderzoek wordt uitgevoerd bij een winkelketen op een beperkt aantal lokaties. Op elke lokatie wordt drie maal, in drie achtereenvolgende maanden, een aantal respondenten ondervraagd. Dit kan worden ondergebracht in een drie-niveau analyse. Het laagste niveau wordt daarbij gevormd door de respondenten. Het tweede niveau zijn de tijdstippen, en het derde niveau de lokaties. Het aardige van deze opzet is dat hier sprake is van herhaalde metingen op het tweede niveau; bij de achtereenvolgende gelegenheden wordt steeds teruggekeerd naar dezelfde lokaties, maar er worden wel steeds andere respondenten ondervraagd!

NOTEN

1. In plaats van alle random componenten tegelijk toe te voegen, kunnen ze ook een voor een worden toegevoegd en elk apart worden getest voor significantie. Bij grote modellen is dit overzichtelijker, en het rekenwerk verloopt sneller.
2. Bij conjuncte analyse wordt dit probleem gewoonlijk opgelost door de vignetten in te delen volgens een fractionele factoriële proefopzet, zodat in ieder geval de hoofdeffecten van de attribuut-variabelen geschat kunnen worden.
3. De bekendste programma's zijn HLM van Raudenbush en Bryk, ML3 van Goldstein, en VARCL van Longford. Van VARCL bestaat een publiek domein versie voor de PC.

LITERATUUR

Bosker, R.J. en Snijders, T.A.B. (1990). Statistische aspecten van multi-niveau onderzoek. *Tijdschrift voor Onderwijsresearch*, 15, 317-329.

Bryk, A.S. en Raudenbush, S.W. (1987). Applying the hierarchical linear model to measurement of change problems. *Psychological Bulletin*, 101, 147-158.

Bryk, A.S., Raudenbush, S.W., Seltzer, M. & Congdon, R.T. (1988). An introduction to HLM. Computer program and user's guide.

De Leeuw, J. en Kreft, G.G. (1986). Random coefficient models. *Journal of Educational Statistics*, 11,1, 55-85.

Goldstein, H. (1986). Efficient statistical modelling of longitudinal data. *Annals of Human Biology*, 13, 129-141.

Goldstein, H. (1987). *Multilevel Analysis in Educational and Social Research*. London: Griffin.

Goldstein, H. (1989). Efficient statistical modelling of longitudinal data. In: D. Bock (ed.). *Multilevel Analysis of Educational Data*. San Diego: Academic Press.

Goldstein, H. en Silver R. (1989). Multilevel and multivariate models in survey analysis. In: Skinner, C., Holt, D., en Smith, F. (eds). *The Analysis of Complex Surveys*. New York: Wiley.

Hermkens, P.L.J. (1983). Oordelen over de rechtvaardigheid van inkomens. Utrecht: Rijksuniversiteit Utrecht, dissertatie.

Hox, J.J. (1992). Multilevel models for interviewer and respondent effects. *Sociological Methods & Research* (in voorbereiding).

Hox, J.J., Kreft, G.G. en Hermkens, P.L.J. (1991). The analysis of factorial surveys. *Sociological Methods & Research*, 19, 493-510.

Hox, J.J., De Leeuw, E.D. en Kreft, G.G. (1991). The effect of interviewer and respondent characteristics on the quality of survey data: a multilevel model. In: P. Biemer et al. (eds). *Measurement Errors in Surveys*. New York: Wiley.

Jaccard, J., Turrisi, R. en Wan, C.K. (1990). *Interaction Effects in Multiple Regression*. Newbury Park, CA: Sage.

Kreft, G.G. (1987). Methods and Models for the Measurement of School Effects. Universiteit van Amsterdam: dissertatie.

Kreft, G.G. en Hox, J.J. (1991). *Using Hierarchical Linear Models*. Nijmegen: ITS (ter perse).

Longford, N.T. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika*, 74, 817-827.

Longford, N.T. (1990). VARCL. Software for Variance Component Analysis of data with Nested Random Effects. Princeton, NJ: Educational testing Service.

Muthén, B.O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika*, 54, 557-585.

Muthén, B.O. en Satorra, A. (1989). Multilevel aspects of varying parameters in structural models. In: Bock, D. (ed) (1989). *Multilevel Analysis of Educational Data*. San Diego: Academic Press.

Prosser, R., Rasbash, J. en Goldstein, H. (1990). ML3. Software for Three-Level Analysis. Users Guide. London: University of London, Institute of Education.

Raudenbush, S.W. & Bryk, A.S. (1986). A hierarchical model for studying school effects. *Journal of Educational Statistics*, 10, 1-17.

Skinner, C., Holt, D., en Smith, F. (eds) (1989). *The Analysis of Complex Surveys*. New York: Wiley.

Vriens, M. en Wittink, D.R. (1990). Conjuncte analyse in het marktonderzoek. *Jaarboek Nederlandse Vereniging van Marktonderzoekers, 1990-91*. Haarlem: de Vrieseborch.