

DEEL III

SCHOOLEFFECTEN OP SCHOOLLOOPBANEN

Ita Kreft

Vakgroep AVO/UvA

Jan de Leeuw

Vakgroep Datatheorie FSW/RUL

1 INLEIDING

In schoolloopbaanonderzoek spelen twee discussies een rol. In de eerste discussie ligt de nadruk op de vraag in hoeverre de ongelijkheid van het behaald eindniveau in het onderwijs tot uitdrukking komt in een ongelijkheid van maatschappelijke positie en omgekeerd hoe de ongelijkheid in sociaal economische status van de ouders van de leerlingen tot uiting komt in het bereikte eindniveau van de leerling zelf. De tweede discussie betreft de vraag naar de effectiviteit van het onderwijs. In het volgende verslag gaan we slechts op deze laatste discussie in, echter met enige beperking, omdat we in ons onderzoek gebruik hebben gemaakt van een databestand met slechts individuele gegevens (het *SMVO-cohort*). Schoolkenmerken ontbreken. We beantwoorden daarom alleen de vraag of scholen verschillen in effectiviteit, en niet de vraag waardoor dit wordt veroorzaakt. De effectiviteit van de scholen is meetbaar met behulp van de vele individuele gegevens die in ons bestand voorhanden zijn. Uit deze over een periode van zes jaar verzamelde gegevens stellen we een effectmaat samen die we de naam *schoolloopbaan* geven. Deze maat bevat informatie omtrent de duur, structuur, en kwaliteit van de schoolcarrière van een leerling. Omdat er vele combinatiemogelijkheden zijn van deze drie schoolloopbaankarakteristieken over zes jaar hebben we gezocht naar een vorm van datareductie, die ons inziens het meest efficiënt weergeeft wat er gebeurt, zonder dat we daarbij vooraf assumpties hoeven te maken over bijvoorbeeld de belangrijkheid van niveau van onderwijs (is HAVO 3 hoger dan MAVO 3 ?), of van aantal jaren scholing (is 4 jaar MAVO hoger dan 3 jaar HAVO ?). In tegenstelling tot het Nederlandse onderzoek tot nu toe nemen wij vooraf geen beslissing over de waarde van schooldiploma's of het aantal jaren scholing.

Ook in de Amerikaanse literatuur wordt *Educational Attainment* gedefiniëerd als het aantal jaren scholing gecombineerd met het niveau van die scholing. Jencks

merkt hierover op, dat een dergelijke maat beperkingen heeft. "First it treats each year of school or college as if it were exactly as valuable as the next. This is clearly an oversimplification. An extra year in college increases a man's learning power more than an extra year of high school. Completing the last year of either high school or college brings more economic benefits than completing any one of the three preceding years. A second major limitation of our approach is that we make no qualitative distinctions between types of certification or types of institution." (Jencks, 1973, pagina 136). We zullen in dit paper een voorbeeld geven van een operationalisatie zoals ook Jencks gebruikt, en waarop dan ook dezelfde bezwaren van toepassing zijn. Daarnaast suggereren we een oplossing voor de door Jencks geschetste problemen. In onze operationalisatie van schoolsucces zien we inderdaad een verschil tussen posities, waarbij het verschil tussen de ene positie en de andere soms klein is en soms groot.

2. PROBLEMEN VAN LONGITUDINALE AARD

Voordat we beginnen met een verslag van eigen onderzoek naar de vraag of scholen verschil maken, gaan we na hoe, in Nederlands onderzoek, het begrip *schoolloopbaan* geoperationaliseerd is. We bespreken daartoe kort de verschillende wijzen waarop dit begrip wordt gedefinieerd, gecodeerd, en vervolgens eventueel wordt gekwantificeerd. Men doet dit laatste met behulp een theorie (kwalitatieve reductie), of met behulp van een techniek (kwantitatieve reductie). Dat een dergelijke discussie niet nieuw is, blijkt uit een artikel van Pedhazur (1975). Hij bespreekt hier een door o.a. Mayeske gebruikte techniek om nominale categorieën van variabelen zo te schatten dat ze als numerieke variabelen in regressievergelijkingen te gebruiken zijn. De techniek, *criterion scaling*, wordt ontleend aan Beaton (1969), die de volgende omschrijving geeft. "Criterion scaling does not attempt to produce an absolute

scale for a factor but instead to scale the factor with reference to an external criterion.' Dit betekent, dat de categorieën van een variabele waarden toegekend krijgen die afhangen van het gehanteerde criterium, bijvoorbeeld het homogeniteitscriterium bij de door ons toegepaste techniek. We komen hierop terug in latere paragrafen.

2.1 De schoolloopbaan

Het gebruik van de schoolloopbaan als een longitudinaal gemeten variabele is te beschouwen als een stap vooruit in het schoolloopbaanonderzoek. Wij baseren deze konklusie op verschillende uitspraken van schooleffect onderzoekers, die kritiek uitoefenden op de gebruikelijke effectmaten, meestal gestandariseerde taal- of rekentests. Om verschillende redenen noemt men 'procesvariabelen' betere effectmaten. Gestandariseerde tests, volgens de kritiek,

- a. meten niet wat er wordt geleerd in de school,
- b. gaan voorbij aan de verschillende doelstellingen van de school,
- c. hebben geen directe relatie met schoolsucces,
- d. meten geen proces maar slechts een enkel moment.

Ook Spady wijt de verschillende, en soms elkaar tegensprekende, resultaten van schooleffecten onderzoek in zijn review artikel uit 1973 aan de gebruikte effectmaten, die vaak slechts benaderingen zijn van wat men eigenlijk wil meten. 'The result is *research by proxy*' (l.c., page 172). Ook zijn advies luidt de gestandariseerde toets door geobserveerde procesvariabelen te vervangen.

Dit is geen eenvoudige raad. Een schoolloopbaan kunnen we opvatten als een proces, dat zich binnen een, door het onderzoek bepaald, tijdsinterval (0,T) voltrekt. De periode 1977-1982 uit ons onderzoek is zo'n tijdsinterval. Meestal onderscheidt men binnen een dergelijk proces een eindig aantal vaste meetmomenten. Bijvoorbeeld overgangen van de ene klas naar de volgende klas gemeten over zes jaren. Dit zijn voorbeelden van een verzameling *toestanden S*,

die elkaar opvolgen in de tijd. De schoolloopbaan van leerling i is daarmee een functie f_i die $(0,T)$ afbeeldt in S . Hierbij is $f_i(t)$ de toestand waarin leerling i zich op tijdstip t bevindt. Het aantal mogelijke toestanden in S , dat zo gedefinieerd binnen een bepaald schoolsysteem bestaat, is in principe eindig, maar toch in de praktijk buitengewoon groot. Want behalve de *duur* van de loopbaan kunnen we ook nog de *structuur* (opstroom, afstroom, uitstroom) en de *kwaliteit* onderscheiden. Dit levert in ons eigen (beperkte) onderzoek op scholengemeenschappen al reeds 25 toestanden op voor iedere van de zes tijdstippen. Dat wil zeggen, dat we ruimte hebben voor 25^6 , ofwel ongeveer 250 miljoen verschillende schoolloopbanen. In de praktijk komt een groot aantal hiervan niet voor, wat niet wegneemt, dat we toch ruim 800 verschillende patronen vinden, waarvan een groot aantal uniek is. Dat datareductie noodzakelijk is, is duidelijk. Hoe stop je 800 verschillende loopbanen in een beperkt aantal *categorieën*, of in bovenstaande termen: hoe definieer je n functies f_i zodat deze *optimaal* zijn? Wij gebruiken daarvoor *homogeniteitsanalyse*, waarbij het optimaliteitscriterium *homogeniteit* is. Homogeniteit wil zeggen, dat we eisen dat de tussengroeps-variantie groot is ten opzichte van de totale variantie. Onze oplossing valt daarmee in de categorie kwantitatieve datareductie. Dat ook andere oplossingen mogelijk zijn, komt later aan de orde.

2.2 Datareductie

Dat het probleem van datareductie in schoolloopbaanonderzoek groter en urgenter begint te worden heeft een duidelijke oorzaak. Aanvankelijk vond dit onderzoek plaats op de lagere school, of vaker nog op één enkel tijdstip, het einde van de lagere schoolperiode. Op het moment van de overgang van het LO naar het VO bestudeerde men de voor de hand liggende gegevens: het advies, de CITO-toetsscores, en de uiteindelijk gekozen vorm van voortgezet onderwijs. Het huidige schoolloopbaanonderzoek vindt meer en meer plaats in het

vervolgonderwijs. Niet in de laatste plaats is dit mogelijk gemaakt door het ter beschikking komen van nieuwe datasets, waarbij leerlingen in het voortgezet onderwijs langere tijd gevolgd worden, zoals bij het door het CBS verzamelde cohort 1977 en het Groningse cohort van Meijnen (1983). Deze verschuiving van het onderzoeksterrein is ook theoretisch van belang, omdat in ons onderwijssysteem de differentiatie eerst echt begint in het secundaire onderwijs. Een stelling waar Jungbluth (1985) het overigens niet mee eens zal zijn. Dit betekent tegelijkertijd echter een vergroting van het aantal mogelijk te onderscheiden loopbanen, omdat niet alleen de duur (zoals in het LO), maar ook de structuur en de kwaliteit van de loopbanen van elkaar gaan verschillen.

De noodzaak tot datareductie wordt hierdoor duidelijk. Er blijft natuurlijk altijd de mogelijkheid om op grond van theoretische overwegingen slechts één aspect van de schoolloopbaan te kiezen als zijnde relevant voor het onderzoek, bijvoorbeeld de 'hoogte' van het behaalde eindexamen. Deze keuze voldeed tot nog toe in het schoolloopbaanonderzoek. De hoogte van het 'advies' of de 'keuze van het secundair onderwijs' als maat voor het succes van de leerling aan het einde van de lagere schoolperiode heeft steeds centraal gestaan. Onze bedoeling is om vooralsnog deze keuze niet te maken, maar na te gaan hoe terecht een dergelijke keuze is. Als zittenblijven, opstroom, of afstroom binnen het secundaire onderwijs belangrijke factoren zijn, dan zullen deze tussentijdse wisselingen tot uitdrukking moeten komen in de gebruikte maat voor schoolsucces.

Omdat wij van mening zijn, dat er nog geen aanwijzingen zijn dat zittenblijven op een VWO-school even ernstige gevolgen heeft als zittenblijven op een MAVO of dat één jaar meer scholing op elk niveau evenveel waarde toevoegt aan de schoolloopbaan (vergelijk de opmerking van Jencks, geciteerd in de inleiding), en ook omdat we geen idee hebben of zittenblijven evenveel waarde heeft voor het uiteindelijk bereikte niveau van onderwijs als niet-zittenblijven maar

afstromen naar een lager niveau (zoals in een steekproef als de onze een reële mogelijkheid is) maken we een dergelijke keuze vooraf niet. We bespreken in de volgende paragraaf welke keuzen in Nederlands onderzoek zoal zijn gemaakt, en welke kritiek daarop valt te formuleren. Onze uiteindelijke conclusie is, dat een alternatieve vorm van datareductie met zo weinig mogelijk assumpties vooraf het schooleffect in het secundair onderwijs beter weer kan geven. Althans voor zolang theoretische overwegingen op basis waarvan voor een meer verantwoorde vorm van datareductie gekozen kan worden niet voorhanden zijn.

3. DE SCHALING VAN DE SCHOOLLOOPBAAN

In deze paragraaf houden we ons bezig met de operationalisatie van de afhankelijke variabele, de schoolloopbaan, zoals we die in verschillende Nederlandse onderzoeken tegenkomen. Het valt op dat elke operationalisatie het doel heeft de vele kenmerken van het ruwe materiaal te reduceren tot een kleiner aantal. Concreet komt dit neer op groeperen, clusteren, of schalen van verschillende categorieën in één enkele categorie. Het criterium daarbij is steeds die loopbanen bij elkaar te voegen die veel op elkaar lijken. Wat eenmaal bij elkaar gevoegd is, kan later niet meer onderscheiden worden. Met andere woorden: loopbanen die na reductie in eenzelfde categorie komen, worden verder behandeld als identiek. Het verschil tussen de benaderingen in het Nederlandse onderzoek ligt daarom uitsluitend in het criterium dat men gebruikt voor reductie. In de volgende samenvatting gaan we daarom na wie of wat het criterium is geweest waarop de toewijzing van verschillende loopbanen aan identieke categorieën is gebaseerd. Op basis hiervan komen we tot een tweedeling. In de eerste benadering baseert men de indeling op het oordeel van experts, waarbij het aantal experts en de soort experts wisselt van onderzoek tot

onderzoek. In een tweede benadering baseert men de indeling op een criterium dat afhangt van de gebruikte schalingstechniek. Deze laatste methode lijkt in principe op de criterion-scaling techniek, waarover we het eerder hebben gehad. Een meer gedetailleerde versie van dit deel van ons verhaal staat overigens in De Leeuw en Kreft (1985).

3.1 Het oordeel van experts

Deze vorm van data reductie komt in zeer verschillende vormen voor. We beginnen met de meest duidelijke vorm, die we voor de eerste keer tegenkomen bij Van Weeren (1960). Een groot aantal veel voorkomende schoolloopbanen (102 stuks) worden hier verbaal beschreven en aan 100 experts (leraren) voorgelegd. Men werd gevraagd deze 102 loopbanen in te delen in 11 categorieën, die oplopen in zwaarte. Via schaalmethoden werden deze oordelen omgezet in schaalwaarden. Een soortgelijke procedure vinden we bij Cremers (1980). We maken hierbij wat kritische kanttekeningen. De verbale omschrijvingen van loopbanen houden rekening met de duur én met de kwaliteit van de loopbaan, maar in feite weinig of helemaal niet met de structuur, dat wil zeggen met de diverse overgangen tijdens het secundair onderwijs. Als dit laatste bij deze methode wél in overweging wordt genomen, neemt het aantal mogelijke loopbanen snel toe. Meester en De Leeuw (1983) vinden bijvoorbeeld in het SMVO-cohort méér dan 1600 loopbanen. Het is duidelijk dat geen enkele expert een dergelijke omvangrijke beoordelingstaak aankan. Dit nog afgezien van de subjectiviteit van het oordeel van een specifieke groep beoordelaars (leraar-experts). Het betrekken van externe criteria (hier expertise) bij de beschrijving lijkt ons niet relevant, omdat het mogelijk is de loopbaan geheel intern te beschrijven.

Een geheel andere benadering vinden we bij Bosker e.a. (1985), en in een aangepaste versie bij Roeleveld e.a. (1985). In deze benadering wordt de loopbaan in termen van twee verschillende variabelen beschreven: tussentijds

niveau (einde lager onderwijs) en bereikt eindniveau (einde middelbaar onderwijs). De uiteindelijke loopbaanvariabele is relatief eindniveau, wat het verschil is tussen bereikt eindniveau en tussentijds niveau. Aan de niveaus worden getallen toegekend van 1 t/m 10, waarbij doubleren geen punten oplevert, overgaan wel. De procedure wordt in het vervolgonderwijs uitgebreid met punten voor opstroom en afstroom. Het resultaat is, dat uiteindelijk iedere leerling een getal krijgt toegekend dat de waarde uitdrukt van zijn/haar gevolgde schoolloopbaan. Afhankelijk van het aantal doublures en/of opstroom- en afstroomgebeurtenissen in de schoolcarrière kan een leerling in LBO-5 een gelijke eindscore krijgen als een leerling in LBO-4, MAVO-3, HAVO-2, of (een zéér late leerling) in VWO-1. Dat dergelijke leerlingen in een zelfde categorie terechtkomen van de variabele 'eindniveau' en vervolgens als identiek worden behandeld is natuurlijk vreemd en moeilijk te beargumenteren.

De onderzoekers hebben hier weliswaar als 'experts' de loopbaan gekwantificeerd, maar hun uiteindelijke oordeel is niet gebaseerd op overeenstemming, zoals bij Van Weeren en Cremers, maar op de veronderstelling van gelijke intervallen en van additiviteit van type en jaarklas. Deze a priori overwegingen zijn aanvechtbaar. Dat ze in de praktijk tot niet al te grote vertekeningen hoeven te leiden is te danken aan het feit dat voorbeelden zoals we die eerder gaven in het bestand weinig voorkomen. Wisselingen tussen niveaus zijn vrij zeldzaam, en komen meestal ná het behalen van een diploma voor. Zittenblijven betekent vaak niets anders dan dat men één jaar later het gewenste diploma haalt. Dit blijkt uit analyses met de door ons gebruikte schoolloopbaanvariabele.

Roeleveld e.a. (1985) geven zoals gezegd een aangepaste versie van de methode Bosker. Hun methode wordt toegepast op dataset die bijna identiek is aan degeen die wij gebruiken, uit het SMVO-cohort 1977 van het CBS. In het onderzoek van Roeleveld e.a. krijgt iedere leerling op tijdstip t een score tussen

de 10 en de 90. Voor de aanvang in respectievelijk het LBO, MAVO, HAVO, en VWO worden 10, 20, 30, en 40 punten verstrekt. Elk jaar op dezelfde schoolsoort geeft 10 punten extra, mits men overgaat, anders krijgt men er geen punten bij. Theoretisch zijn er 14 verschillende scores mogelijk, op elk der meetpunten (tijdstippen). Men houdt bij deze scoring rekening met de duur, structuur en de kwaliteit van de loopbaan. Duur en structuur worden echter niet expliciet onderscheiden. Een jaar blijven zitten levert hetzelfde op als overgaan en afstroom naar een hoger leerjaar. Het belangrijkste criterium is de kwaliteit.

Deze interessante vorm van a priori kwantificatie is opgebouwd rond het idee dat iemand, die uit een lager naar een hoger schooltype wil, in het Nederlandse systeem één jaar verliest. Daarom is HAVO-2 hetzelfde als MAVO-3 en als VWO-1. Dit zou men een theoretische overweging kunnen noemen, maar een beetje reflectie laat direct de uiterst beperkte toepasbaarheid van dit soort 'theorie' zien. Vooral in het begin van de loopbaan hoeven overgangen van HAVO naar VWO geen jaar te kosten. Het is evenmin zo, dat men een jaar wint wanneer men van een hoger naar een lager schooltype gaat. De diverse vormen van beroepsonderwijs zijn vrijwel niet in de theorie te passen, evenmin de verschillende waarden van het diploma (A, B, C-niveau van het eindexamen LBO, of het gekozen vakkenpakket op het middelbaar onderwijs). Het gebruik van *hoger* en *lager*, en het gelijkstellen van de intervallen in de codering, suggereren veel meer a priori kennis dan er in feite is. Blijven zitten is hetzelfde als zonder vertraging doorstromen naar een lager type: VWO-3 en dan VWO-3 is hetzelfde als VWO-3 en dan HAVO-4 (het is ook hetzelfde als MBO-1 - MBO-1, of als HAVO-4 - MBO-1). Rest ons te vermelden dat Roeleveld e.a. doorcoderen na vertrek. Wanneer er vertrokken is met diploma krijgen de leerlingen 10 punten extra, wanneer ze ziek, overleden, of geëmigreerd zijn, worden ze missing. Een dergelijke codering introduceert een extra homogeniteit in de loopbaan, met mogelijke gevolgen voor de uitkomsten wanneer men deze variabele als afhankelijke variabele gebruikt.

We bespreken nu enkele andere benaderingen, die met wat moeite ook onder hetzelfde hoofdje behandeld kunnen worden. Een beperkte keuze uit de verschillende kenmerken die aan een schoolloopbaan te onderscheiden zijn, is in de strikte betekenis immers ook te beschouwen als een 'oordeel van experts'. Het verschil met de voorgaande benaderingen is echter groot. Men maakt dikwijls geen keuzen op grond van 'expertise', maar eerder op grond van praktische overwegingen. Zo vinden we in de rapportage van het Van Jaar tot Jaar onderzoek (Kropman en Collaris, 1974, Collaris en Kropman, 1978) en in CBS (1982a) overgangsmatrixen en stroomdiagrammen, met het doel de gegevens zo duidelijk mogelijk weer te geven. Bij Diederens (1983) vinden we een vijftal keuzes van variabelen uit het enorme aantal mogelijke keuzen, nu op grond van een theorie. Een korte beschrijving van deze coderingen volgt.

Bij het gebruik van stroomdiagrammen wordt een individuele loopbaan beschreven in een reeks van schooltypen, bijvoorbeeld LO-LBO-ULO-STOP of LO-VHMO-WO-STOP. Het stroomdiagram laat op redelijk inzichtelijke wijze zien welke reeksen er voorkomen en met welke frequenties. Door alleen op de soorten reeksen te letten, verwaarlozen we natuurlijk de tijd, dat wil zeggen de duur van de loopbaan. Ook merken we daardoor bepaalde kenmerken van de structuur niet op, zoals zittenblijven. In CBS (1982a) vinden we binnen schooltypen wel stroomdiagrammen, waaruit de duur te bepalen is, maar dit gaat ten koste van de structuur, omdat uit deze diagrammen weer niet af te leiden is welke veranderingen van schooltypen er optreden. In CBS (1982a, Tabel 27) staan de veranderingen van type weergegeven. Algemene uitspraken over de schoolloopbaan zijn uit deze tabellen zonder nadere analyse lastig.

Diederens (1983) heeft met behulp van theorieën, die aangeven welke aspecten van de loopbaan belangrijk zijn voor de latere verwerving van inkomen een aantal variabelen uit de loopbaan afgeleid. Van de aspecten duur, structuur, en

kwaliteit, die allemaal een rol spelen, kan men een groot aantal variabelen van het type dat Diederer koos construeren. De keuze van een vijftal moet daarom theoretisch onderbouwd zijn. Wij hebben deze keuze laten liggen, omdat wij, klaarblijkelijk in tegenstelling tot anderen, geen sociaal-wetenschappelijke theorie voorhanden hadden om deze keuze te onderbouwen.

3.2 Schalingstechnieken in plaats van experts

Hoewel men de kop van de paragraaf zo uit zou kunnen leggen, is het bij het gebruik van schalingstechnieken zeker niet mogelijk dat de onderzoeker alles aan de techniek overlaat. Er moeten nog steeds keuzen gemaakt worden, en dit kan nog steeds meer of minder verstandig gebeuren. Zowel de keuze van de techniek als de wijze van codering van de schoolloopbaan, die beiden vooraf gaan aan de toepassing van de analysetechniek, zijn cruciaal voor de uitkomst van de analyse. Immers als men voor de beschrijving $f_1(t)$ slechts een relatief klein aantal meetpunten gebruikt, die in de tijd het verloop van de individuele loopbaan aangeven, dan neemt men in feite aan dat de loopbaan tussen de meetpunten constant blijft. Het gaat er vervolgens om hoe men deze stappen gaat coderen. Het uiteindelijke doel is getallen toe te kennen aan deze codes.

Nu zijn kwantificering en codering twee verschijningsvormen van hetzelfde proces. Bij beiden reduceer je het aantal verschillende waarden dat een variabele aanneemt. Dit betekent, dat je eenzelfde waarde toekent aan verschillende loopbanen, waardoor je deze in het vervolg van het onderzoek verder als identiek opvat. Het streven bij reductie is homogene blokken van loopbanen te aggregeren tot categorieën, waarbij homogeen gedefinieerd zal worden aan de hand van een criterium. Wat homogeen genoemd gaat worden, wordt dus geheel of gedeeltelijk bepaald met behulp van de vraagstelling en de theoretische uitgangspunten. Enkele voorbeelden. Bij Diederer (1983) en CBS (1982b) vinden we een (kwalitatieve) reductie, waarbij de gehele loopbaan in een klein

aantal variabelen wordt weergegeven. Bij Bosker e.a. (1985), Bosker en Van der Velden (1985), Koopman e.a. (1985), Roeleveld e.a. (1985), Kreft en Bronkhorst (1985), en Tesser (1986) vinden we reducties, waarbij tegelijkertijd rekening wordt gehouden met schooltype, leerjaar, uitstroom, en eindexamen. Zeker vergeleken met eerdere pogingen, zijn deze laatste voorbeelden van een weinig reducerend type. Binnen dit laatste type vinden we een markant verschil met betrekking tot het uitgangspunt tussen Tesser en Kreft en Bronkhorst enerzijds en de rest. De eersten gebruiken een *a posteriori* kwantificatie, wat door sommige, terecht, als *theoriearm* wordt beschouwd. Deze omschrijving geeft ons inziens echter tevens het belangrijkste voordeel van de methode weer, omdat het inderdaad niet nodig is vooraf (*a priori*) veel beslissingen te nemen.

Voor de constructie van de in ons onderzoek gebruikte afhankelijke variabele gebruiken we een bepaalde schalingsmethode, homogeniteitsanalyse (computerprogramma PRIMALS), met als criterium homogeniteit. We doen dit in navolging van Kreft en Bronkhorst (1985) en Tesser (1986), die deze methode beiden, onafhankelijk van elkaar, gebruikt hebben. Deze methode wordt hier slechts summier besproken, omdat ze later uitgebreid aan de orde komt. In beide genoemde onderzoeken wordt de schoolloopbaan in de eerste fase gecodeerd in een groot aantal categorieën voor ieder van de meetmomenten (tijdstippen). In deze fase is dus niet zo zeer sprake van datareductie, zoals in de *a priori* methodes wel al gebeurt. Er komt natuurlijk wel reductie, maar later, in de analytische fase. De homogeniteitsanalyse, met als criterium 'homogeniteit binnen groepen' is daarbij het hulpmiddel. Wij zouden in het algemeen voor dit soort uitstel van reductie een pleidooi willen houden. Hoewel het onvermijdelijk is, dat in schoolloopbaanonderzoek vooroordelen een relatief grote rol spelen, is het raadzaam deze in een zo laat mogelijk stadium op te laten draven. Waarmee we overigens niet willen beweren, dat homogeniteitsanalyse een procedure is, die superieur is aan *a priori* kwantificatie, of zelfs aan kwalitatieve beschrijvingen van typische of frequente loopbanen. Als alternatieve methode,

uitsluitend gebaseerd op de gegevens zelf, lijkt het ons bij de huidige stand van zaken een goede aanpak.

De werkwijze van datareductie met homogeniteitsanalyse is als volgt. De eerste fase, die van de codering, resulteert in een $n \times m$ matrix van de vorm $f_i(t_j)$, waarin voor alle n leerlingen op ieder van de m tijdstippen wordt aangegeven waar ze zich bevinden. Hierop wordt in de tweede fase homogeniteitsanalyse toegepast met als uitkomst een kwantificatie van alle toestanden over alle m tijdstippen. Voor de bespreking van homogeniteitsanalyse in het algemeen verwijzen we naar appendix A van De Leeuw en Kreft (1985). We zullen ons hier beperken tot de specifieke toepassing op schoolloopbaanonderzoek. Het is resultaat is dan, dat een rij van m kwalitatieve labels wordt vervangen door m in principe verschillende getallen. Dit gebeurt voor alle leerlingen. Met deze getallen zijn de rekenkundige bewerkingen en analyses mogelijk, die op numerieke variabelen in het algemeen toepasbaar zijn. Een enkele score voor de gehele loopbaan is bijvoorbeeld eenvoudig te berekenen door het gemiddelde te nemen van de m scores over de m tijdstippen van meting.

Het is duidelijk dat de resultaten van homogeniteitsanalyse vrijwel geheel gebaseerd zijn op de structuur van de loopbaan en in feite op de structuur van alle loopbanen in het cohort of bestand. De kwaliteit van de loopbaan telt in principe in het geheel niet mee, alle informatie die we aan het programma geven zijn de labels van de loopbanen, zonder dat we daar van te voren één of andere ordening in aanbrengen. De ordening wordt als het ware aangebracht door de structuur. Het is van belang om op te merken dat de wijze van coderen van de diverse stadia in de loopbaan natuurlijk wel van belang is.

3 LOOPBANEN IN EEN SMVO-DEELBESTAND

3.1 Het SMVO-bestand

Het CBS-SMVO bestand is een gestratificeerde twee-traps cluster steekproef uit alle leerlingen die deelnamen aan het eerste jaar van het Nederlandse voortgezet onderwijs in het jaar 1977. De steekproef bevatte 37.280 leerlingen. Ze werden tijdens hun schoolloopbaan gevolgd. We maken gebruik van de versie van het bestand waarin informatie aanwezig is tot en met september 1982. Informatie over het SMVO-cohort en het SMVO-bestand is te vinden in CBS (1982b, 1982c). Deze publicaties behandelen de schoolloopbaan tot en met het onderwijselement 1980. Informatie over de later toegevoegde variabelen onderwijselement 1981 en 1982 staat in CBS (1983, 1984).

Het aantal secundaire analyses op het CBS-SMVO bestand is ondertussen al te groot om op te noemen. Wij noemen slechts de analyses in de CBS-publicaties zelf, en een aantal publicaties, die zich speciaal met het verloop van de schoolloopbaan bezig houden. Dit zijn Kreft en Bronkhorst (1985), Faasse e.a. (1985), Roeleveld e.a. (1985), Tesser (1986). Al deze artikelen zijn besproken, of ten minste genoemd, in het vorige hoofdstuk.

3.2 Selectie van het deelbestand

We selecteerden in eerste instantie op onderwijselement 1977, dit moet gelijk zijn aan brugklas. Er bleven 15.329 leerlingen over. Vervolgens, via de schoolnummers, op scholengemeenschappen voor MAVO-HAVO-VWO. Er bleven 5.464 leerlingen op 114 scholengemeenschappen over. De redenen voor een dergelijke selectie zijn het verkrijgen van een zo homogeen mogelijke groep ten aanzien van hun beginsituatie in het voortgezet onderwijs met daarna een zo breed mogelijke niveaukeuze. Dat we hierbij niet gekozen hebben voor bredere scholengemeenschappen, bijvoorbeeld met LBO, heeft meerdere redenen.

- a. Een vergelijking tussen een scholengemeenschap met en zonder LBO-afdeling leek ons niet correct.
- b. Het aantal leerlingen op een dergelijke brede scholengemeenschap is beperkt (n=284).
- c. Er blijken belangrijke aanvangsverschillen te bestaan tussen diverse soorten scholengemeenschappen (zie ook de tabellen in 2.2).
- d. Een eenvoudige maat voor de afhankelijke variabele is hierdoor mogelijk. Deze is eenvoudiger omdat we categorieën van LBO-afdelingen en klassen niet bij de constructie van de schoolloopbaan hoeven te betrekken.

Dat we ook niet gekozen hebben om smallere scholengemeenschappen, zoals HAVO-VWO scholen, erbij te betrekken heeft gedeeltelijk dezelfde redenen (zie a en c). Bovendien is ook de kans dat een leerling niet wordt aangenomen op een HAVO-VWO scholengemeenschap groter. De beginselectie is strenger en de keuzemogelijkheden na de brugperiode zijn beperkter dan bij een brede scholengemeenschap. Van de 15.329 leerlingen zitten er overigens maar liefst 7.509 op een HAVO-VWO scholengemeenschap.

3.3 Codering van achtergrondvariabelen

Als achtergrondvariabelen gebruikten we in de analyses van De Leeuw en Kreft (1985) geslacht, CITO-score taaltoets, CITO-score rekenoets, provincie waarin de school gevestigd is, sociaal milieu, en advies van de onderwijzer in de zesde klas LO omtrent de meest geschikte vorm van voortgezet onderwijs. In de analyses in dit paper gebruiken we uitsluitend de CITO-totaalscore. De resultaten van Kreft en Bronkhorst (1985) laten zien dat het riskant is om bij regressieanalyse binnen scholen te veel predictoren te gebruiken, omdat de regressiecoëfficiënten dan wel erg instabiel worden.

Om de representativiteit van ons deelbestand te onderzoeken gebruiken we wat andere achtergrondvariabelen. In ons deelbestand zitten voor 48.2% jongens en

voor 51.8% meisjes. Voor het gehele SMVO-cohort is dit, respectievelijk, 48.5% en 51.5%. De verdeling op provincie, sociaal milieu, en advies onderwijzer vergt wat meer studie. Voor details verwijzen we De Leeuw en Krefth (1985). We vatten hun bevindingen hier kort samen. De belangrijkste verschillen tussen ons deelbestand en het gehele CBS-cohort treden op in advies. Dat is natuurlijk ook logisch, omdat in ons bestand vrijwel geen mensen met LBO-advies zitten. Alleen diegenen, die ondanks een LBO-advies toch naar een MAVO-HAVO-VWO scholengemeenschap zijn gegaan, en dat zijn er slechts 53. Wat de milieusamenstelling betreft, gebeurt er ook min of meer wat we mogen verwachten. De employees zijn oververtegenwoordigd, hoe hoger hoe meer oververtegenwoordigd zelfs. Arbeiders zijn ondervertegenwoordigd, en natuurlijk vooral ongeschoolde arbeiders. De categorie 'overigen', die nader beschreven wordt in CBS (1982b, hoofdstuk 7) is eveneens ondervertegenwoordigd. Dit komt overeen met eerdere bevindingen, dat deze categorie zich aan de lage kant van de sociaal-economische schaal bevindt. Bij de verdeling over provincies valt op dat Friesland en Limburg sterk ondervertegenwoordigd zijn. Drente, Zuid-Holland en Zeeland zijn oververtegenwoordigd. De reden is, dat in Friesland en Limburg slechts 7.7% en 6.9% voor de schoolsoort 'brugklas met MAVO' kiezen, terwijl dit in Zuid-Holland, Zeeland en Drente respectievelijk 27.5%, 27.1% en 24.0% is. Vergelijk CBS (1982b, staat 5.7, pagina 87). De CBS-cijfers zijn berekend over de populatie, waarin zo'n 17% voor deze schoolsoort kiest. In de SMVO-steekproef is dit ongeveer 20%.

Samenvattend kunnen we zeggen dat onze selectie van MAVO-HAVO-VWO scholengemeenschappen tot gevolg heeft gehad dat de achtergrondvariabelen in het deelbestand aanzienlijk anders verdeeld zijn dan die in de CBS-steekproef. In ons deelbestand is een fikse oververtegenwoordiging van de 'hogere' milieus, van de VWO-HAVO adviezen, en van provincies als Zuid-Holland, waar blijkbaar veel van dit soort scholengemeenschappen zijn. Dit betekent

vanzelfsprekend, dat alle uitspraken in dit artikel gaan over een geselecteerde deelpopulatie van de generatie 1977, en zeker niet over de gehele populatie van leerlingen in het secundaire onderwijs. Wanneer we ons deelbestand vergelijken met de deelbestanden die gebruikt zijn in het andere, meest direct vergelijkbare, schoolloopbanenonderzoek, dan vinden we toch nog aanzienlijke verschillen.

In Roeleveld e.a. (1985) worden in principe dezelfde leerlingen uit het SMVO bestand geselecteerd. Kleine scholen, en scholen waarin minder dan 15% van de leerlingen een CITO-toets heeft gedaan, worden echter verwijderd. Er blijven daardoor slechts 4142 leerlingen over. De marginalen op de achtergrondvariabelen verschillen weinig van onze marginalen. Koopman e.a. (1985) gebruiken het Amsterdamse VOCAL-bestand, jaargang 1975, en wel allen leerlingen op kategoriale MAVO's. Van enige vergelijkbaarheid is natuurlijk geen sprake. Faasse e.a. (1985) gebruiken alleen de Noord-Brabantse leerlingen uit SMVO, omdat ze hun resultaten willen vergelijken met die van een ouder Noord-Brabants cohort. Bosker e.a. (1985), vergelijk ook Bosker en Van der Velden (1985), gebruiken het Groningen-cohort van Meijnen. Bij dit soort regionale cohorten is het natuurlijk niet helemaal reëel om representativiteit te onderzoeken. Tesser (1985) gebruikt, zoals we gezien hebben, ook het SMVO-cohort. Zijn deelbestand is gedefinieert door een nogal ingewikkelde trapsgewijze steekproeftrekking, met daar overheen nog een niet al te succesvolle machtigingsprocedure. Het resultaat zijn 4 deelbestanden van ieder ongeveer 250 leerlingen, met als eerste bekende schooltype LBO, MAVO, HAVO en VWO. Bij representativiteitsonderzoek constateert Tesser een oververtegenwoordiging van de 'hogere' milieus in zijn deelbestand. Een deel van deze oververtegenwoordiging komt door de inmiddels beruchte categorie 'overig', waarvan de exacte betekenis en codering eigenlijk niet meer te achterhalen is. Kreft en Bronkhorst (1985) tenslotte gebruiken een steekproef uit het deelbestand waar wij mee werken. Zij kiezen willekeurig 2293 leerlingen uit onze 5464. De marginalen in hun onderzoek en het onze komen goed overeen.

3.4 Codering van de loopbaanvariabele

We volgen de procedure van Kreft en Bronkhorst (1985). Dat wil zeggen dat we drie belangrijke coderingsbeslissingen nemen. In de eerste plaats coderen we terug, dat wil zeggen dat we de brugperiode coderen met gebruikmaking van informatie uit latere jaren. Een loopbaan brugjaar 1, brugjaar 2, HAVO 3 wordt dus HAVO 1, HAVO 2, HAVO 3. Het effect van de codering is dat het aantal mogelijke loopbaanpatronen aanzienlijk teruggebracht wordt. De codering bevordert de homogeniteit van de loopbanen, omdat het schooltype na de brugperiode gebruikt wordt om te doen als of dat de leerling tijdens de brugperiode ook continu op dit schooltype gezeten heeft. In een latere paragraaf zullen we trachten uiteen te zetten wat voor invloed dit soort coderingsbeslissingen hebben op de kwantificatie van de schoolloopbaan. Een tweede belangrijke coderingsbeslissing is, dat we de toestand in leerjaar 1 niet coderen. Iedereen zit in het eerste brugjaar, dat wil zeggen feitelijk in dezelfde toestand. Terugcoderen van het latere schooltype zou natuurlijk wel differentiatie aanbrengen, maar die zou geheel op codering berusten, te meer natuurlijk omdat ook feitelijk het eerste brugjaar dikwijls ongedeeld is. Dit zou ons inziens te veel homogeniteit in de loopbaan suggereren.

De derde beslissing is om individuen die uitstromen vanaf het moment van uitstroom te coderen als missing. Voor een deel is dit gemakkelijk te verdedigen, die zijn inderdaad zoek geraakt en niet verder gevolgd. Per september 1982 zijn er van de oorspronkelijke CBS-steekproef 13.000 mensen vertrokken, dus ongeveer 35%. De redenen staan, kort aangeduid, in CBS (1984). Van degenen die vertrokken zijn is 4% overleden of naar het buitenland of ziek, 56% is werkend of werkzoekend, 30% is om een andere reden uitgevallen, en van 10% is de reden niet bekend. In CBS (1984, Tabel 4) vinden we dat van de leerlingen die in een brugklas beginnen in 1982 19% uitgevallen is. Het gaat daarbij natuurlijk om een paar bredere en om veel smallere scholengemeenschappen dan

onze MAVO-HAVO-VWO scholen. Het is daarom niet verwonderlijk dat ons bestand een uitval heeft van 21% in 1982. We hebben geen informatie over de redenen van vertrek in onze analyse betrokken.

Nog even een korte vergelijking met de methode gevolgd door Tesser (1986). Ook Tesser codeert terug, hij hoeft het echter in veel mindere mate te doen omdat hij in principe het meest in categoriale scholen geïnteresseerd is. Om dezelfde reden neemt Tesser het eerste leerjaar wel op als onderdeel van de schoolloopbaan. Het grootste verschil tussen zijn benadering en de onze zit hem in het feit, dat Tesser doorcodeert. Nadat een leerling vertrokken is, houdt hij op de overblijvende meetmomenten zijn laatste code. Dus HAVO-1, HAVO-2, verdwenen wordt HAVO-1, HAVO-2, HAVO-2, HAVO-2, HAVO-2, HAVO-2. Het is bij Tesser daarom zuiverder om niet te spreken van de schoolloopbaan, maar van het hoogst bereikte onderwijsniveau. Of, beter nog, van het laatst bekende onderwijsniveau. Het spreekt min of meer vanzelf dat deze vorm van doorcoderen kwalitatief veel homogeniteit in de loopbaan introduceert die er eerst niet was, en die eigenlijk in veel gevallen nogal aanvechtbaar is. We weten vaak niet waarom een leerling verdwijnt, of hij nog onderwijs volgt, misschien nog cursussen of vakopleidingen volgt, enzovoorts. Het lijkt ons dat de compactheid van de loopbanen die geïntroduceerd wordt met doorcoderen aanzienlijk groter, en aanzienlijk problematischer, is dan de compactheid geïntroduceerd met terugcoderen (vooral als men het eerste leerjaar weglaat).

Gegeven deze codeerbeslissingen ligt het nogal voor de hand welke toestanden overblijven om de loopbaan te beschrijven. We hebben MAVO-1 t/m MAVO-4, HAVO-1 t/m HAVO-5, VWO-1 t/m VWO-6. Dit zijn de 15 basistoestanden. Daarnaast hebben we categoriën nodig voor leerlingen die vanuit de scholengemeenschappen doorstromen (men zegt ook wel afstromen) naar het LBO. Dit zijn er niet erg veel, en we onderscheiden daarom alleen LBO met nog

geen diploma en LBO met diploma. Bij de MAVO maken we onderscheid tussen MAVO-3+ en MAVO-3-, en tussen MAVO-4+ en MAVO-4-, naar gelang de leerling al dan niet een diploma gehaald heeft op het desbetreffende tijdstip. Ook Tesser (1986) gebruikt het examenresultaat bij de codering van de loopbaan. Een extra probleem zijn de leerlingen die doorstromen naar één of andere vorm van beroepsonderwijs. We onderscheiden kort MBO, MBO-1 t/m MBO-3, en HBO-1 als mogelijke toestanden in ons bestand. Gezamenlijk leidt dit tot 25 toestanden. De frequenties van voorkomen op de diverse meetmomenten worden gegeven in De Leeuw en Kreft (1985, Tabel 3). Ze verschillen overigens weinig van de overeenkomstige frequenties in Kreft en Bronkhorst (1985).

4. EEN MULTILEVEL ANALYSE

4.1 Doel van de kwantificering

Het uiteindelijke doel van het kwantificeringsproces van de schoolloopbaan is deze geschikt te maken voor een vergelijking tussen scholen. Reducties, zoals gebruikt door bijvoorbeeld het CBS, zijn voor ons doel niet geschikt, omdat er niet mee doorgerekend kan worden. Onze schoolloopbaanmaat kan wél gebruikt worden als een numerieke effectmaat in bijvoorbeeld regressievergelijkingen. Vordat we verslag zullen doen over de regressie zullen we eerst nog enkele andere rekenkundige bewerkingen met de nieuwe schoolloopbaanvariabele bespreken.

Daarvoor moeten we hem natuurlijk eerst construeren. De resultaten staan uitvoerig besproken in De Leeuw en Kreft (1985). We gebruiken het programma PRIMALS, dat een homogeniteitsanalyse uitvoert op de 5464 loopbanen die op 5 tijdstippen in 25 categorieën gecodeerd zijn. Dit geeft gekwantificeerde schoolloopbaanvariabelen op ieder van de vijf tijdstippen, en

een globale maat voor de schoolloopbaan dien het gemiddelde is van deze vijf kwantificaties. Tesser (1986) gebruikt vooral de vijf jaar-variabelen, Kreft en Bronkhorst (1985) gebruiken de samenvattende maat. De kwantificaties van de 25 categorieën, op ieder van de vijf tijdstippen, worden gegeven in De Leeuw en Kreft (1985, Tabel 4). Vergelijk ook Kreft en Bronkhorst (1985, Tabel 9) en Tesser (1986, Tabel 4.10). In De Leeuw en Kreft (1985, Tabel 5) worden de jaar-variabelen nader onderzocht door ze te kruisen met de achtergrond van de leerling. Correlatie-ratios worden berekend met provincie, beroep vader, geslacht, en advies van de onderwijzer. De samenhang met geslacht en provincie is zwak, met beroep vader iets sterker, en met advies onderwijzer zeer sterk. De multiële correlatie met de CITO-scores is evenwel nog iets hoger. We zien in het algemeen de correlaties en correlatie-ratios iets afnemen in de tijd. Dat is niet verwonderlijk. 'The predictive relations between test results and criterion measures are a function of the period of time between them. In other words, the correlations tend to decrease over time.' (Jansen, 1977, p 19). Bovendien speelt bij onze metingen de toenemende selectie een rol. Door de uitval en uitstroom krijgen de variabelen in de loop van de tijd een steeds sterkere 'restriction of range', met als gevolg ook lagere correlaties. Dit maakt het voorlopig het minst problematisch om de globale schoolloopbaanmaat van Kreft en Bronkhorst te gebruiken. Gebruik van schoolloopbaanmodellen om de vijf loopbaanvariabelen aan elkaar te relateren wordt geïllustreerd in Roeleveld e.a. (1985), De Leeuw en Kreft (1985), en Tesser (1986).

4.2 Multilevel analyse: vaste effecten

De schoolloopbaanvariabele wordt geconstrueerd op basis van de loopbanen van 5464 leerlingen. We elimineren vervolgens alle leerlingen die geen CITO-totaal score hebben. Er blijven 3590 leerlingen over, verdeeld over 106 scholen. In Tabel 1 staan, per school, de aantallen leerlingen, de gemiddelden van CITO-score en schoolloopbaan, de varianties van CITO-score en schoolloopbaan, en de correlaties tussen CITO-score en schoolloopbaan. In

Figuur 1 wordt gemiddelde schoolloopbaan tegen gemiddelde CITO-score geplot.

Op basis van Figuur 1 en Tabel 1 kunnen we gemakkelijk constateren dat scholen verschillen. We kunnen echter niet direct vaststellen in hoeverre deze verschillen op basis van toeval alleen te verklaren zijn. Tabel 2 doet een eerste poging om de verschillen wat systematischer te beschrijven. We bekijken het eenvoudige regressiemodel

$$y_j = a_j + b_j x_j + \epsilon_j. \quad (1)$$

De y_j zijn de n_j loopbaanscores op school j , en de x_j zijn de n_j CITO-scores. De ϵ_j zijn n_j storingstermen, met gemiddelde nul en variantie s_j^2 . Iedere school heeft dus een eigen intercept a_j , een eigen helling b_j , en een eigen foutvariantie s_j^2 . In vergelijking (1), en in de rest van dit paper, worden stochastische variabelen steeds onderstreept. In Tabel 2 vinden we achtereenvolgens, per school, het intercept, de helling, en de foutvariantie. We komen eigenlijk weer niet verder dan de konstatering dat scholen inderdaad verschillen, maar we weten nog steeds niet of de grote variabiliteit toeval of systematiek is. Op ongeveer hetzelfde punt waren ook Kreft en Bronkhorst (1985) met hun analyses beland.

Om na te gaan of de verschillen nu 'echt' of 'niet-echt' zijn moeten we gebruik maken van de statistiek. We veronderstellen dat er een meer beperkend model dan (1) opgaat, en we gaan na in hoeverre we deze veronderstelling vol kunnen houden. De eerste nadere specificatie die we onderzoeken is het covariantieanalyse model. Hiervoor gaan we ervan uit dat de hellingen en de foutvarianties gelijk zijn, maar dat de intercepten mogen verschillen. De geschatte helling is .0622, en de geschatte foutvariantie is .5335. De geschatte intercepten staan in de vierde kolom van Tabel 2.

In alle in dit paper gepresenteerde analyses gebruiken we de maximum likelihood methode om parameters te schatten. Dit is een uniform toepasbare methode, die over het algemeen goede resultaten oplevert, en die bovendien min of meer automatisch chi-kwadraat toetsen construeert. Met behulp van een chi-kwadraat toets kunnen we bijvoorbeeld nagaan of hellingen en foutvarianties inderdaad gelijk zijn. De chi-kwadraat is 280.16, met 210 vrijheidsgraden niet of nauwelijks significant. Wanneer we willen toetsen of ook nog de intercepten gelijk zijn, dan vinden we voor deze additionele specificatie een chi-kwadraat van 437.4 met 105 vrijheidsgraden. Deze is te groot, en dus kunnen we nu zeggen dat scholen inderdaad verschillen, maar alleen in intercept. Er is geen systematisch verschil in helling, of, anders gezegd, een punt extra op de CITO-toets is op alle scholen evenveel waard. Maar leerlingen met een bepaald niveau op de toets doen het zeker niet op alle scholengemeenschappen even goed.

We willen vervolgens nagaan of andere statistische modellen de conclusies uit de covariantieanalyse kunnen aanvullen of verbeteren. Voor dit doel is het handig uit te gaan van Tabel 3. Daarin vinden we de variantie van CITO-score x en de variantie van loopbaan y , alsmede de covariantie van x en y , opgesplitst in een deel *binnen-scholen* en een deel *tussen-scholen*. De tabel bevat ook drie schattingen van de helling, in alle gevallen de covariantie van x en y gedeeld door de variantie van x , en drie schattingen van de foutvariantie. De schattingen uit de covariantieanalyse vinden we terug in de kolom *binnen*, de schattingen gebaseerd op het model dat zowel hellingen als intercepten gelijk zijn in de kolom *totaal*. De kolom *tussen* geeft schattingen gebaseerd op het geaggregeerde regressiemodel voor de 106 scholen. Dit veronderstelt dat

$$y_j = a + bx_j + \epsilon_j \quad (2)$$

De vette symbolen zijn nu schoolgemiddelden. De storingstermen ϵ_j hebben gemiddelde nul, en variantie $\eta_j^{-1}s^2$. Model (2) is onder andere van belang omdat

het in veel schooleffectiviteitsonderzoek gebruikt wordt. In Nederland bijvoorbeeld in het proefschrift van Van Marwijck Kooy (1984). We zien duidelijk dat (2) niet echt gebruikt kan worden om schoolverschillen te schatten. Schoolverschillen worden immers gedefinieerd in termen van effecten van scholen op de loopbanen van individuele leerlingen, en die individuele leerlingen zijn uit het model verdwenen. Zoals bekend zeggen samenhangen tussen gemiddelden weinig over samenhangen tussen verdelingen (Kreft en Van der Eeden, 1984), en kan men vanuit (2) niets zeggen over (1). Ook uit Tabel 3 blijkt weer dat gebruik van schoolgemiddelden een zeer onbetrouwbare manier is om de parameters van multilevel modellen te schatten.

We hebben, met ons covariantieanalyse model, verschillen tussen scholen geconstateerd, maar niet verklaard. De contextuele analyse (Boyd en Iversen, 1979, Van den Eeden en Saris, 1984) doet een poging om tot verklaring te komen. We bekijken hier een eenvoudige vorm van het contextuele model, waarin we voor de intercepten veronderstellen dat

$$a_j = p + qx_j. \quad (3)$$

De hellingen b zijn allemaal gelijk, net zoals in covariantieanalyse. Vergelijking (3) zegt, dat het intercept van school j een lineaire functie is van het gemiddelde op de CITO-toets. Als 'verklaring' zou kunnen gelden, dat de kwaliteit van de school bepaald wordt door het gemiddelde niveau van de leerlingpopulatie. Schatten van de parameters van dit contextuele model is niet moeilijk, zie De Leeuw (1986). We vinden dat de schatting van b hetzelfde is als in het ANCOVA model, dus b is de schatting van de regressiecoëfficiënt gebaseerd op de binnen-groep statistieken. De schatting van q is gelijk aan het verschil van de tussen-groepen b en de binnen-groepen b . Uit Tabel 3 vinden we dus $b = .0622$ en $q = -.0163$. De schatting van de foutvariantie is $s^2 = .5996$, en $p = -2.0820$.

Ongelukkigerwijs past de additionele specificatie (3) slecht. De chi-kwadraat die

(3) binnen het covariantieanalyse model toetst is 418.95 met 104 vrijheidsgraden. We zouden meer ingewikkelde contextuele modellen kunnen uitproberen, met name door ook rekening te houden met de heterogeniteit van de instroom (de variantie van CITO-score per school), of door ook de hellingen af te laten hangen van schoolkenmerken. Het CBS-cohort bevat echter weinig echte schoolkenmerken, en onze eerdere analyses hebben laten zien dat er geen noemenswaardige heterogeniteit van de hellingen is.

4.3 Multilevel analyse: random effecten

Covariantieanalyse blijft voorlopig het meest geslaagde model. Het heeft echter het nadeel dat het wat veel parameters gebruikt, te weten één intercept voor iedere school. Bovendien worden scholen beschouwd als deelpopulaties, en niet als een steekproef uit de beschikbare scholen. Dit heeft als nadeel dat er geen basis is om over scholen in het algemeen uitspraken te doen. Een bijkomende complicatie is, dat de storingstermen moeilijk geïnterpreteerd kunnen worden. De storingstermen immers worden geacht de invloeden samen te vatten van een groot aantal variabelen die wel de loopbaan beïnvloeden, maar die niet direct gerelateerd zijn aan de CITO-score. Het ligt in de rede om te veronderstellen dat deze overige variabelen op eenzelfde school gedeeltelijk hetzelfde effect zullen hebben. Dit zou betekenen dat we uit moeten gaan van gecorrleerde storingstermen binnen scholen. De genoemde nadelen van het covariantieanalyse model kunnen, althans gedeeltelijk, opgevangen worden via variantie componenten of random coëfficiënten modellen (Aitkin en Longford, 1986; De Leeuw en Kreft, 1986).

We nemen aan

$$y_j = a_j + bx_j + \epsilon_j \quad (4)$$

Hierbij heeft de storingsterm ϵ_j weer variantie s_j^2 , net zoals eerst, maar de intercept a_j is nu stochastisch. Naar analogie met (3) nemen we aan

$$a_j = p + qx_j + u_j, \quad (5)$$

waarbij u_j een storingsterm is op schoolniveau, die gemiddelde nul heeft, ongecorreleerd is met ϵ_j , en die variantie $n_j^{-1}w^2$ heeft. Merk op dat de variantie van het intercept afhangt van de grootte van de school. Wanneer we (5) invullen in (4) dan vinden we

$$y_j = p + qx_j + bx_j + (u_j + \epsilon_j), \quad (6)$$

en de correlatie tussen de storingstermen van leerlingen op dezelfde school is daardoor $w^2/(w^2 + n_j s^2)$. Hoe groter de school, hoe dichter deze correlatie bij nul ligt. En ook: hoe groter de individuele storingsterm is ten opzichte van de schoolstoringsterm, hoe dichter de correlatie bij nul ligt. Merk ook op dat de vaste effecten specificatie (3) een speciaal geval is van (5) met $u_j = 0$, dat wil zeggen met $w_j^2 = 0$.

Schattingen van de parameters van dit model zijn gedeeltelijk hetzelfde als de schattingen onder model (3). We vinden een zelfde b en q , en een zelfde p , maar natuurlijk andere schattingen van de varianties van de storingstermen. Voor de individuele foutvariantie vinden we $s^2 = .5498$, en voor de storing op schoolniveau $w^2 = 1.6868$. Binnen dit model kunnen we nu toetsen dat $w^2 = 0$, dat wil zeggen dat model (3) opgaat. Dit levert een chi-kwadraat op van 162.58, met één vrijheidsgraad, wat duidelijk onaanvaardbaar is. We kunnen ook toetsen of $q = 0$. Dit betekent dat de intercepten een willekeurige steekproef zijn uit een populatie van intercepten, er is geen effect van gemiddelde CITO-score op het intercept. De hypothese $q = 0$ levert een chi-kwadraat van 4.45 op, met één vrijheidsgraad niet significant. Bijbehorende schattingen van de overige structurele parameters zijn $b = .0616$, $s^2 = .5497$, $w^2 = 1.8012$, en $p = -2.8224$. Het random coëfficiënt model met $q = 0$ en $w^2 \neq 0$ levert een goede beschrijving, en gebruikt slechts 4 parameters in plaats van de 108 die

voor het gewone ANCOVA model nodig zijn.

Het lijkt erop alsof we de variabiliteit van de intercepten kunnen 'verklaren' door binnen scholen gecorreleerde storingstermen aan te nemen. De correlatie tussen de storingstermen van leerlingen op een zelfde school varieert van .05 (voor scholen met 60 leerlingen) tot .35 (voor scholen met 6 leerlingen). Anders gezegd: voor grote scholen kunnen we zeggen dat de variantie van de storingstermen 90% individueel 'bepaald' is, voor kleine scholen is wel 40% door de school 'bepaald'. Voor de volledigheid hebben we in Tabel 4a nogmaals de schattingen van de structurele parameters van de diverse modellen bij elkaar gebracht, en in Tabel 4b de diverse modeltoetsen.

5. SAMENVATTING

We hebben in dit paper getracht om uit te zoeken of scholen verschil maken, dat wil zeggen of loopbanen van leerlingen met dezelfde achtergrond op verschillende scholen verschillend verlopen. Een beantwoording van deze vraag is waanzinnig moeilijk, en alles wat we in feite kunnen verwachten zijn bepaalde indicaties. Het antwoord zal bijvoorbeeld afhangen van de operationalisering van de schoolloopbaan, en we hebben gezien dat er een groot aantal reducties mogelijk zijn. Vanzelfsprekend hangt het antwoord af van de variabelen die we kiezen om de achtergrond te definiëren, en we hebben laten zien dat het antwoord tenslotte afhangt van de modellen en technieken die we gebruiken om de vraagstelling statistisch te vertalen. Dit laatste aspect zal nader toegelicht worden in Kreft en De Leeuw (1986).

We hebben in dit artikel niet systematisch onderzocht waar de schoolverschillen mee samenhangen. Het is natuurlijk in principe mogelijk de gevonden schooleffecten, waarvoor we de intercepten uit de covariantieanalyse kunnen

gebruiken, te relateren aan schoolgegevens. In Kreft en Bronkhorst (1985) wordt bijvoorbeeld de lengte van de brugperiode hiervoor gebruikt, maar deze variabele heeft natuurlijk het nadeel dat hij per definitie gerelateerd zal zijn aan de geconstrueerde schoolloopbaanvariabele. Interessanter zou het zijn om de regio waarin de school gelegen is, of de milieusamenstelling van de school, te relateren aan de schooleffecten. Dit kan natuurlijk weer op meerdere manieren. We kunnen de schoolkenmerken direct als variabelen in het model opnemen, zoals in contextuele analyse. We kunnen ook later, in een tweede analyse stap, de schoolkenmerken aan de schooleffecten relateren. Een vergelijking van deze twee benaderingen staat in Van den Eeden en Saris (1984) en in De Leeuw en Kreft (1984).

Voorlopig lijkt het ons een belangrijke uitkomst dat er in ieder geval, met behulp van de geconstrueerde loopbaanvariabele en met behulp van eenvoudige lineaire model technieken, enige significante verschillen tussen scholen aan te tonen zijn. Het is daarbij in zekere zin plezierig en in zekere zin verdrietig dat de verschillen vooral in de intercepten en niet in de hellingen zitten. Modellen met variabele hellingen zijn methodologisch aanzienlijk gecompliceerder, zowel technisch als uit het oogpunt van interpretatie. Inhoudelijk gezien zijn variabele hellingen echter aanzienlijk interessanter, omdat ze op een meer duidelijke manier de interactie tussen leerling- en schoolvariabelen laten zien. Het door ons uiteindelijk verkozen model, met een konstante helling, met alle variabiliteit in de intercepten toegeschreven aan random variatie rond een gemiddeld intercept, is theoretisch even teleurstellend als de één-factor modellen van Tesser (1986).

Referenties

- Aitkin, M. & Longford, N. (1986). **Statistical Modelling Issues in School Effectiveness Studies.** *Journal of the Royal Statistical Society*, in press.
- Beaton, A. (1969). **Some Mathematical and Empirical properties of Criterion-scaled Variables.** In G.W. Mayeske et al., **A Study of our Nation's Schools.** Washington, D.C.: U.S. Office of Education.
- Bosker, R., Hofman, A., & Velden, R. van der (1985). **Een Generatie Geselecteerd.** Groningen: RION.
- Bosker, R., & Velden, R. van der (1985). **Onderwijspositie en Selectie.** Paper gepresenteerd op de ORD-85.
- Boyd, L.H. & Iversen, G.R. (1979). **Contextual Analysis: Concepts and Statistical Techniques.** Belmont, Ca.: Wadsworth.
- CBS (1982a). **Schoolkeuze en Schoolloopbaan bij het Voortgezet Onderwijs. Cohort 1964/'65.** 's-Gravenhage: Staatsuitgeverij.
- CBS (1982b). **Schoolloopbaan en Herkomst van Leerlingen in het Voortgezet Onderwijs. Deel 1: Bestand 1977.** 's-Gravenhage: Staatsuitgeverij.
- CBS (1982c). **Schoolloopbaan en Herkomst van Leerlingen in het Voortgezet Onderwijs. Deel 2: Cohort 1977.** 's-Gravenhage: Staatsuitgeverij.
- CBS (1983) **Schoolloopbaan en Herkomst van Leerlingen in het Voortgezet Onderwijs. (Cohort 1977, peildatum september 1981).** CBS mededeling 7797. Voorburg: CBS.
- CBS (1984). **Schoolloopbaan en Herkomst van Leerlingen in het Voortgezet Onderwijs. (Cohort 1977, peildatum september 1982).** CBS mededeling 7818. Voorburg: CBS.
- Collaris, J.W.M., & Kropman, J.A. (1978). **Van Jaar tot Jaar. Tweede Fase.** Nijmegen: ITS.
- Cremers, P.G.J. (1980). **Konstructie van een Schaal voor Bereikt Niveau van**

- Voortgezet Onderwijs. *Tijdschrift voor Onderwijsresearch*, 5, 80-92.
- Diederer, J. (1983) *Van Jaar tot Jaar. Derde Fase*. Nijmegen: ITS.
- Eeden, P. van den, & Saris, W.E. (1984). Empirisch Onderzoek naar Multilevel Uitspraken. *Mens en Maatschappij*, 59, 165-178.
- Faasse, J.H., Bakker, B., Dronkers, J., & Schijf, H. (1985). *Vergelijking van Individuele Schoolloopbanen in het Voortgezet Onderwijs in Noord-Brabant voor en na de Invoering van de Mammoetwet*. Paper gepresenteerd op de ORD-85.
- Janssen, M.G.H. (1977). *An Application of Bayesian Statistical Methods to a Problem in Educational Measurement*. Dissertatie Groningen.
- Janssen, M.G.H. (1979). Het Simultaan Schatten van Regressievergelijkingen in m Groepen. *Tijdschrift voor Onderwijsresearch*, 4, 159-166.
- Jencks, C. (1973). *Inequality. A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.
- Koopman, P., Eeden, P. van der, & Jong, U. de (1985). *Kategorale MAVO-scholen en Schoolloopbanen in Amsterdam*. Paper gepresenteerd op de ORD-85.
- Kreft, G.G., & Bronkhorst, H. (1985). *Maken Scholen Verschil ? Een Secundaire Analyse op het SMVO-cohort 1977*. Amsterdam: Vakgroep AVO/UvA.
- Kreft, G.G., & Eeden, P. van den (1985). Valkuilen bij het Proefschrift van L. van Marwijk Kooy-Van Baumhauer, en hoe deze uit de weg te gaan. In D.P. Kroes (red), *Studiebijeenkomst over de Dissertatie "Scholen Verschillen"*. SISWO-Publikatie 306. Amsterdam: SISWO.
- Leeuw, J. de, & Kreft, G.G. (1984). *Over Multilevel Analyse*. Rapport RR-84-06. Leiden: Vakgroep Data Theorie FSW/RUL.
- Leeuw, J. de, & Kreft, G.G. (1985). *De Definitie en Kwantificatie van de Schoolloopbaan*. Rapport RR-85-20. Leiden: Vakgroep Data Theorie FSW/RUL.
- Leeuw, J. de, & Kreft, G.G. (1986). *Random Coefficient Models for Multilevel Analysis*. *Journal of Educational Statistics*, in press.

- Leeuw, J. de (1986). **Modelspecifische Estimation of Schooleffects**. Leiden: Vakgroep Data Theorie FSW/RUL.
- Kreft, G.G., & Leeuw, J. de (1986). **Modelspecifieke Schatting van Schooleffecten**. Paper ORD-86, in voorbereiding.
- Marwijk Kooy- Von Baumhauer, L. (1984). **Scholen Verschillen. Een Verkennend Onderzoek naar het Intern Functioneren van Vijftientig Scholengemeenschappen VWO-HAVO-MAVO**. Dissertatie Rotterdam.
- Meester, A., & Leeuw, J. de (1983). **Intelligentie, Sociaal Milieu, en de Schoolloopbaan**. Leiden: Vakgroep Datatheorie FSW/RUL.
- Meijnen, G.W. (1984). **Van Zes tot Twaalf. Een Longitudinaal Onderzoek naar de Milieu- en Schooleffecten op Loopbanen in het Lager Onderwijs**. SVO reeks 79. Harlingen: Flevodruk.
- Pedhazur, E.J. (1975). **Analytic Methods in the Study of Educational Effects. Review of Research in Education**, 3, 243-286.
- Roeleveld, J., Eeden, P. van den, & Jong, U. de (1985). **Scholengemeenschappen met MAVO en Schoolloopbanen**. Paper gepresenteerd op de ORD-85.
- Spady, W.G. (1973). **The Impact of School Resources on Students. Review of Research in Education**, 1, 135-177.
- Tesser, P. (1986). **Sociale Herkomst en Schoolloopbanen in het Voortgezet Onderwijs**. Dissertatie Rotterdam.
- Weeren, P. van (1960). **Project Nationale Differentiatietest. Rapport No 3: Ontwerp Ontwikkeling Criteriumschaal**. Leiden: RUL.

	no	ave x	ave y	var x	var y	cor x,y
1	22	53.41	-.03	44.53	.77	.58
2	17	49.94	.42	34.34	.44	.30
3	23	49.12	-.28	119.77	.64	.52
4	34	50.06	.20	59.96	.68	.52
5	38	43.87	-.01	64.99	.59	.46
6	21	46.67	-.05	59.44	1.01	.51
7	40	50.68	.40	80.21	1.05	.77
8	48	49.98	.25	70.08	1.00	.67
9	8	42.75	.09	102.25	1.08	.69
10	43	46.91	.03	87.78	.72	.68
11	14	44.93	-.42	42.94	.31	.53
12	17	49.29	.22	63.13	1.04	.65
13	38	48.74	.28	84.76	.94	.66
14	36	44.17	-.32	73.42	.70	.36
15	38	46.66	.15	78.69	1.05	.66
16	36	47.03	-.23	74.78	.77	.30
17	12	49.17	.10	63.61	.62	.56
18	38	50.21	.08	68.54	.74	.43
19	35	41.94	-.14	91.03	.80	.60
20	41	45.15	.11	101.30	1.02	.79
21	59	49.36	.30	83.02	.87	.68
22	39	40.28	-.09	115.76	1.12	.68
23	56	51.48	.24	51.62	.98	.45
24	40	47.25	-.23	77.75	.79	.48
25	54	51.30	.16	55.06	.96	.61
26	48	49.21	.43	69.09	.64	.65
27	50	43.12	.30	96.44	1.10	.67
28	39	40.41	-.08	68.42	1.07	.58
29	54	51.30	.16	55.06	.96	.61
30	53	44.28	-.46	80.89	.43	.55
31	33	47.42	.15	121.42	.89	.69
32	35	45.80	.13	76.08	.93	.61
33	9	39.22	-.69	58.46	.42	-.14
34	13	39.15	-.61	69.71	.21	.22
35	15	38.80	-.46	80.89	.43	.55
36	6	52.50	-.72	133.99	.02	.28
37	53	42.38	-.24	104.49	.96	.64
38	51	56.84	.97	37.69	.60	.21
39	33	48.33	.37	109.37	1.11	.56
40	45	42.57	-.49	89.55	.86	.61
41	12	45.08	-.65	28.95	.50	.17
42	22	53.86	.59	44.53	.49	.47
43	23	54.35	.02	51.53	.94	.59
44	22	48.09	.72	60.18	.44	.60
45	53	44.98	-.15	91.37	1.04	.65
46	56	46.09	-.37	88.27	.81	.73
47	43	44.12	-.04	106.26	.77	.79
48	35	47.77	.29	69.53	.72	.53
49	37	52.86	.30	63.84	.81	.64
50	24	46.58	-.02	42.00	.81	.49

	no	ave x	ave y	var x	var y	cor x,y
51	45	49.64	.24	54.85	.85	.51
52	18	50.50	-.14	113.24	.83	.48
53	40	49.87	.29	49.84	.90	.51
54	35	45.40	.66	54.64	.61	.29
55	17	53.23	.25	108.56	.66	.51
56	13	47.23	-.18	90.11	1.37	.91
57	60	48.33	-.04	119.42	.74	.61
58	40	41.20	.09	105.66	1.01	.76
59	25	48.08	.37	99.68	.96	.52
60	15	41.40	-.34	53.34	.44	.64
61	29	42.48	-.35	87.14	.71	.46
62	32	47.78	.43	90.96	.97	.59
63	47	45.13	.10	73.40	1.00	.73
64	56	48.14	.27	66.36	1.30	.71
65	17	43.71	.12	109.58	.85	.76
66	49	41.61	-.22	108.00	.96	.74
67	19	52.89	.08	83.00	1.11	.82
68	34	45.65	-.06	55.33	.67	.61
69	57	41.75	.13	136.57	1.19	.77
70	43	42.86	.08	64.56	.81	.69
71	41	47.56	-.02	102.50	.90	.71
72	41	52.73	.40	135.84	1.02	.82
73	10	45.00	.25	27.22	.90	.16
74	43	48.13	.13	63.14	.84	.46
75	45	44.89	.32	114.05	1.14	.58
76	18	45.39	-.28	58.35	.62	.21
77	17	46.88	-.21	82.99	.96	.83
78	32	45.03	-.14	83.97	1.01	.61
79	25	40.44	-.12	81.96	1.02	.70
80	47	46.30	.19	135.59	1.12	.75
81	19	45.47	.40	63.24	1.06	.58
82	39	42.53	.09	92.62	.88	.62
83	44	49.70	.50	98.31	.77	.61
84	51	49.53	.22	132.83	1.04	.72
85	39	49.44	.33	99.79	1.07	.76
86	43	49.58	-.15	64.40	.69	.59
87	41	51.90	.62	68.51	.70	.49
88	13	50.31	.28	55.01	1.24	.36
89	40	48.88	.38	95.75	1.10	.76
90	42	46.29	.37	98.99	.90	.60
91	24	54.46	.71	35.78	.69	.31
92	40	48.38	.39	85.05	.98	.81
93	16	42.06	-.26	84.61	1.11	.43
94	37	43.16	-.07	75.18	.72	.59
95	20	44.25	.13	48.16	.99	.47
96	39	52.03	-.18	65.52	.69	.69
97	39	44.15	-.38	99.79	.44	.60
98	47	52.51	.21	77.70	.66	.56
99	21	46.90	.01	104.80	.66	.63
100	42	49.36	-.41	70.03	.61	.48

	no	ave x	ave y	var x	var y	cor x,y
101	45	50.69	.05	53.40	.43	.34
102	14	45.43	-.38	114.62	.54	.29
103	30	43.67	-.45	57.16	.60	.54
104	37	49.19	-.10	54.15	.67	.61
105	34	44.15	-.45	50.46	.64	.55
106	13	43.62	.24	117.84	1.06	.80
	3970	47.16	.08	95.13	.94	.60

Tabel 1: Voor iedere school: aantal leerlingen, gemiddelde CITO score, gemiddelde loopbaan, variantie CITO-score, variantie loopbaan, correlatie CITO-score en loopbaan.

model voor iedere school			fixed ANCOVA	random ANCOVA	
intercept	helling	foutvariantie	intercept	intercept	
1	-4.11	.076	.513	-3.35	-1.04
2	-1.27	.034	.396	-2.69	-.53
3	-2.15	.038	.469	-3.34	-1.03
4	-2.58	.056	.500	-2.91	-.71
5	-1.94	.044	.467	-2.74	-.58
6	-3.15	.066	.748	-2.95	-.74
7	-4.07	.088	.429	-2.75	-.58
8	-3.75	.080	.550	-2.86	-.66
9	-2.94	.071	.565	-2.57	-.45
10	-2.86	.062	.388	-2.89	-.69
11	-2.45	.045	.225	-3.21	-.94
12	-3.89	.083	.599	-2.85	-.66
13	-3.10	.069	.528	-2.75	-.58
14	-1.88	.035	.611	-3.07	-.83
15	-3.41	.076	.594	-2.75	-.58
16	-1.66	.030	.701	-3.16	-.89
17	-2.61	.055	.423	-2.96	-.74
18	-2.16	.045	.601	-3.04	-.81
19	-2.51	.056	.515	-2.75	-.58
20	-3.46	.079	.382	-2.70	-.54
21	-3.13	.070	.467	-2.77	-.60
22	-2.78	.067	.600	-2.60	-.47
23	-2.96	.062	.783	-2.96	-.74
24	-2.52	.048	.608	-3.17	-.90
25	-3.98	.081	.604	-3.03	-.80
26	-2.65	.063	.371	-2.63	-.49
27	-2.79	.072	.607	-2.38	-.30
28	-3.02	.073	.713	-2.59	-.47
29	-3.89	.085	.501	-2.74	-.57
30	-2.23	.040	.298	-3.21	-.94
31	-2.66	.059	.468	-2.80	-.62
32	-2.96	.068	.586	-2.72	-.56
33	-.22	-.012	.415	-3.13	-.88
34	-1.09	.012	.202	-3.05	-.81
35	-1.49	.026	.632	-2.87	-.68
36	-.88	.003	.015	-3.99	-1.53
37	-2.84	.061	.568	-2.88	-.68
38	-.53	.026	.570	-2.57	-.44
39	-2.36	.056	.762	-2.64	-.49
40	-3.04	.060	.542	-3.14	-.88
41	-1.66	.022	.487	-3.45	-1.12
42	-2.08	.050	.386	-2.76	-.59
43	-4.31	.080	.611	-3.36	-1.05
44	-1.75	.051	.282	-2.27	-.22
45	-3.27	.069	.601	-2.95	-.74
46	-3.60	.070	.380	-3.24	-.96
47	-3.01	.067	.291	-2.78	-.61
48	-2.28	.054	.517	-2.68	-.53
49	-3.50	.072	.476	-2.99	-.76
50	-3.19	.068	.616	-2.92	-.71

	model voor iedere school			fixed	random
	intercept	helling	foutvariantie	ANCOVA intercept	ANCOVA intercept
51	-2.90	.063	.626	-2.85	-.66
52	-2.22	.041	.642	-3.28	-.99
53	-3.12	.068	.665	-2.81	-.63
54	-.73	.031	.555	-2.16	-.13
55	-1.87	.040	.491	-3.06	-.82
56	-5.49	.112	.236	-3.12	-.86
57	-2.37	.048	.467	-3.05	-.81
58	-2.98	.074	.428	-2.47	-.37
59	-2.08	.051	.700	-2.62	-.48
60	-2.76	.058	.262	-2.92	-.71
61	-2.12	.042	.563	-2.99	-.77
62	-2.48	.061	.632	-2.54	-.42
63	-3.74	.085	.466	-2.71	-.55
64	-4.51	.099	.644	-2.72	-.56
65	-2.80	.067	.359	-2.60	-.47
66	-3.12	.070	.434	-2.81	-.63
67	-4.92	.095	.362	-3.21	-.93
68	-3.12	.067	.420	-2.90	-.70
69	-2.87	.072	.484	-2.47	-.37
70	-3.23	.077	.424	-2.59	-.46
71	-3.18	.066	.446	-2.98	-.76
72	-3.34	.071	.333	-2.88	-.68
73	-1.06	.029	.877	-2.55	-.43
74	-2.43	.053	.666	-2.86	-.67
75	-2.28	.058	.757	-2.47	-.37
76	-1.26	.022	.592	-3.10	-.85
77	-4.40	.089	.299	-3.13	-.87
78	-3.15	.067	.633	-2.94	-.73
79	-3.28	.078	.519	-2.64	-.50
80	-2.97	.068	.490	-2.69	-.54
81	-3.02	.075	.706	-2.43	-.34
82	-2.48	.060	.541	-2.56	-.44
83	-2.19	.054	.486	-2.59	-.46
84	-2.94	.064	.501	-2.86	-.67
85	-3.57	.079	.454	-2.75	-.58
86	-3.18	.061	.449	-3.23	-.95
87	-1.96	.050	.536	-2.61	-.47
88	-2.44	.054	1.081	-2.85	-.66
89	-3.59	.081	.463	-2.66	-.51
90	-2.28	.057	.576	-2.51	-.40
91	-1.64	.043	.626	-2.68	-.52
92	-3.81	.087	.335	-2.62	-.48
93	-2.34	.049	.908	-2.88	-.68
94	-2.56	.058	.469	-2.75	-.59
95	-2.85	.067	.770	-2.62	-.49
96	-3.86	.071	.360	-3.42	-1.09
97	-2.13	.040	.280	-3.13	-.87
98	-2.50	.052	.452	-3.06	-.81
99	-2.33	.050	.396	-2.91	-.70
100	-2.62	.045	.469	-3.48	-1.14

	model voor iedere school			fixed ANCOVA intercept	random ANCOVA intercept
	intercept	helling	foutvariantie		
101	-1.49	.030	.377	-3.10	-.85
102	-1.28	.020	.491	-3.21	-.93
103	-2.87	.055	.427	-3.17	-.90
104	-3.44	.068	.421	-3.16	-.90
105	-3.18	.062	.444	-3.20	-.93
106	-3.06	.076	.380	-2.47	-.37

Tabel 2: Voor iedere school: intercept, helling, en foutvariantie van het model per school, en intercepten voor de covariantieanalyse modellen

	TOTAAL	TUSSEN	BINNEN	T/B
$v(x)$	95.1289	13.6509	81.4780	0.1435
$v(y)$	0.9436	0.0948	0.8488	0.1005
$c(x,y)$	5.6957	0.6265	5.0692	0.1100
b	0.0599	0.0459	0.0622	
s^2	0.6026	2.2355	0.5334	

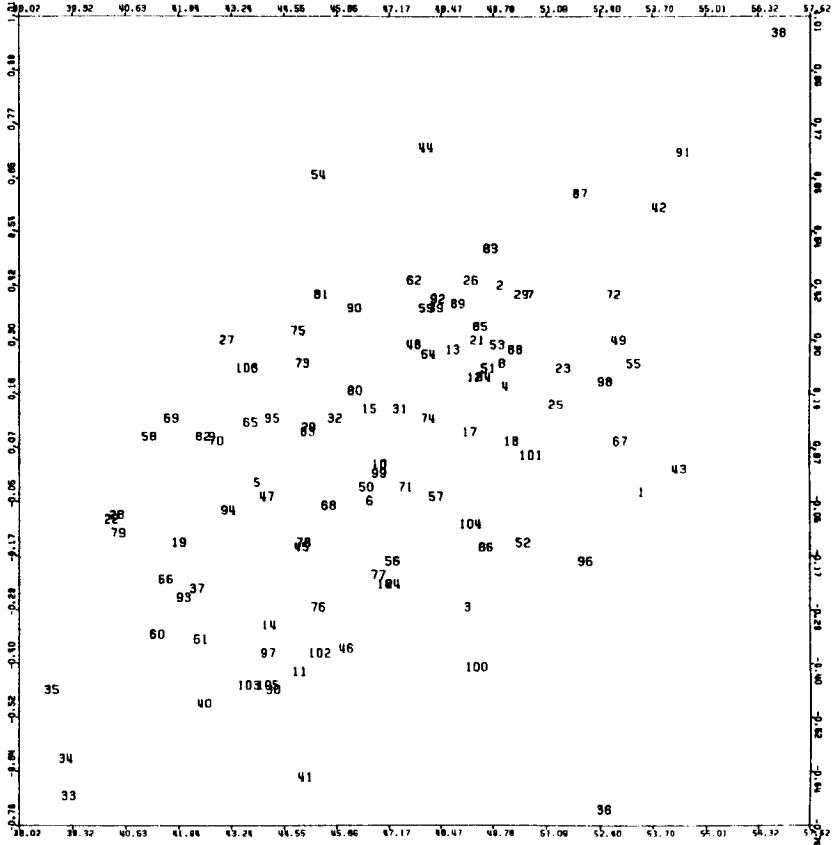
Tabel 3: Partitionering van de varianties en covarianties in tussen- en binnen-componenten.

MODEL	INTERCEPT	HELLING	VARIANTIESCHATTINGEN	
			INDIVIDUEEL	SCHOOL
nulmodel	****	****	****	0
covariantie analyse	****	.0622	.5335	0
gelijke intercepten fixed	-2.7410	.0599	.6026	0
contextueel fixed	****	.0622	.5996	0
contextueel random	****	.0622	.5496	1.6868
gelijke intercepten random	-2.8224	.0616	.5497	1.8012

Tabel 4a: Samenvatting modelschattingen.
nulmodel: intercepten, hellingen, en individuele varianties verschillen per school.
covariantie analyse: intercepten verschillen per school.
contextuele modellen: $p = -2.0820$ en $q = .0163$.

MODEL	# VRIJE PARAMETERS	VERSCHIL LOG-LIKELIHOODS	CHI- KWADRAAT	DFR
(1) nulmodel	318	****	****	****
(2) covariantie analyse	108	(1)-(2)	280.16	210
(3) gelijke intercepten fixed	3	(2)-(3)	437.40	105
(4) contextueel fixed	4	(2)-(4)	418.95	104
		(5)-(4)	162.58	1
(5) contextueel random	5	****	****	****
(6) gelijke intercepten random	4	(5)-(6)	4.45	1

Tabel 4b: Samenvatting modeltoetsen.



Figuur 1:
Gemiddelde CITO-score (x-as) en
gemiddelde schoolloopbaan (y-as)
voor 106 scholen.