LOGLINEAIRE MODELLEN

MET LATENTE VARIABELEN

EN MISSING DATA

Scriptie voor het vak Statistiek en
Methoden en Technieken van Onderzoek

Begeleider: prof. dr. J.A.P. Hagenaars

Jeroen Vermunt
Tilburg, juni 1988.

VOORWOORD.

Deze afstudeerscriptie is geschreven ter afsluiting van het vak Statistiek en Methoden en Technieken van Onderzoek. Ik denk in deze scriptie de doeleinden te hebben gerealiseerd die staan vermeld in de afstudeernota van de vakgroep SMT.

Bijna een jaar lang ben ik met veel plezier bezig geweest met dit onderzoek. Het thema van deze scriptie, loglineaire modellen en nonresponse, is afkomstig van mijn afstudeerbegeleider Jacques Hagenaars, die zich al enige tijd met dit onderwerp bezig houdt. Ik wil hem op deze plaats bedanken voor zijn intensieve en stimulerende begeleiding. Verder wil ik Jaap Dronkers bedanken voor het ter beschikking stellen van de door mij gebruikte dataset.

INHOUDSOPGAVE.

INLEIDING.

Vele onderzoekers hebben te maken met de omstandigheid dat er gegevens van de in de steekproef gevallen onderzoekseenheden ontbreken: het zogenaamde nonresponse-verschijnsel. Deze nonresponse kan van tweeërlei aard zijn. Het kan om item-nonresponse gaan, deze doet zich voor wanneer bij een bepaalde variabele gegevens ontbreken, of om de "gewone" nonresponse wanneer een onderzoekspersoon, door wat voor reden dan ook, niet deelneemt aan het onderzoek. De omvang van de item-nonresponse zal over het algemeen afhangen van het onderwerp van de vragen en van de kwaliteit van de interviewers. Het percentage van de onderzoekseenheden dat helemaal niet deelneemt aan het onderzoek ligt bij een gewone survey meestal tussen 30% en 40%. Bij panelonderzoek ligt dit percentage over het algemeen nog veel hoger: vaak blijft na drie golven nog maar 30% van de oorspronkelijke onderzoeksgroep over (Hagenaars,1984,p.133).

Indien men mag veronderstellen dat het al dan niet deelnemen aan het onderzoek geheel op toeval berust dan zal de nonresponse weinig invloed hebben op de onderzoeksresultaten. Het enige dat er gebeurt is dat de nauwkeurigheid van de schattingen voor de populatieparameters kleiner is dan de beoogde doordat het aantal waarnemingen kleiner is dan het "geplande aantal". Over het algemeen zal er echter wel degelijk een verband bestaan tussen het al dan niet participeren in een onderzoek, of het al dan niet antwoorden op een bepaalde vraag, en de te meten kenmerken. Dit betekent dat wanneer er niets wordt ondernomen de onderzoeksresultaten vertekend zullen zijn; de nonresponse is dan immers geen a-selecte steekproef uit de onderzochte populatie.

Er zijn twee manieren om de vertekening van de onderzoeksresultaten door nonresponse te voorkomen. Ten eerste kan een onderzoeker proberen om alsnog informatie te verkrijgen over de nonresponse-groep, bijvoorbeeld door het inzetten van beter gekwalificeerde interviewers (Segers,1977,p.117). Vaak is dit echter een te kostbare aangelegenheid en bovendien zal er ook na een tweede of derde ondervragingsronde nonresponse overblijven. De tweede strategie die een onderzoeker kan volgen is het "corrigeren" voor nonresponse door middel van speciaal hiervoor ontwikkelde technieken. In de volgende paragraaf zullen enkele van deze correctiemethoden, en de assumpties die ten grondslag liggen aan de diverse methoden, aan de orde komen.

Bij de analyse van categorische data, ofwel gegevens van nominaal meetniveau, met behulp van loglineaire modellen is het erg belangrijk om alle beschikbare informatie te benutten. De cellen van de multivariate frequentietabel die wordt geanalyseerd moeten immers voldoende waarnemingen bevatten om efficiënte schattingen te verkrijgen voor de verbanden tussen de onderzoeksvariabelen, of om een van te voren opgestelt model met enige kans op verwerping op zijn adequaatheid te toetsen. In de tweede paragraaf zal worden ingegaan op de techniek van de loglineaire analyse, zowel met als zonder latente variabelen. Daarna zal de door Fuchs (1982) voorgestelde correctiemethode voor nonresponse bij de analyse van categorische data met behulp van loglineaire modellen en de uitbreiding van Hagenaars naar modellen met latente variabelen aan de orde komen. Deze techniek, die kortgezegd neerkomt op het vinden van de "maximum likelihood" (meest aannemelijke) schattingen voor de "sufficient statistics" (de te reproduceren marginale frequenties), gaat uit van de assumptie dat de nonresponse "missing at random" is: binnen subgroepen, bestaande uit de categorieën van de variabelen waarvan de scores wel bekend zijn, is de nonresponse "toevallig". Bij de "missing completely at random" assumptie daarentegen, dient de nonresponse-groep een a-selecte steekproef uit de totale onderzoekspopulatie te zijn.

In paragraaf 3.2 zal de toetsingsprocedure die Fuchs gebruikt, worden besproken. Deze is alleen toe te passen wanneer de nonresponse "missing completely at random" is, terwijl de correctiemethode zelf uitgaat van de zwakkere "missing at random" assumptie. Dit betekent dat wanneer de nonresponse niet "missing completely at random" is, het gepostuleerde model altijd verworpen zal worden. In dit geval kunnen de populatieparameters echter wel geschat worden (Hagenaars, 1986). Little en Rubin (1987) geven aan dat het bij een "missing at random" nonresponse wel mogelijk is om een hypothetisch model te toetsen via conditionele toetsing. Hagenaars (1986) stelt voor om de toetsingprocedure aan te passen op de wijze die beschreven wordt door Chen en Fienberg (1974). In d~~ez~~e paragraaf ~~zal~~ een algemenere oplossing aangedra-

vier/zal

gen worden die afkomstig is van Fay (1986). Fay neemt zogenaamde response-variabelen, een soort indicator-variabelen, op in het causale model met de oorspronkelijke onderzoeksvariabelen. Ze geven aan of een respondent wel of niet respondeert op een bepaalde onderzoeksvariabele. De oplossing van Fay is algemener dan die van Chen en Fienberg (en dan die van Little en Rubin) omdat het response-mechanisme zowel "ignorable" als "nonignorable" kan zijn. Er is sprake van een "ignorable" response-mechanisme wanneer de nonresponse niet systematisch samenhangt met de variabele waarvan gegevens ontbreken, terwijl bij een "nonignorable" responsemechanisme deze samenhang wel bestaat. In causale modellen waarin response-variabelen zijn opgenomen kan de onderzoeker zelf bepalen of er al dan niet een direct effect wordt veronderstelt tussen een variabele met ontbrekende gegevens en de daarbij horende response-variabele. Dit betekent dat met behulp van deze techniek niet alleen het toetsingsprobleem van Fuchs wordt opgelost, maar dat het bovendien niet meer nodig is om assumpties te maken over de aard van de nonresponse.

Daarnaast zal in paragraaf vier ~~zal~~ de methode die Fay voorstelt uitgebreid worden naar loglineaire modellen met latente variabelen.

In de vijfde paragraaf zal het belang van deze correctiemethode voor nonresponse worden geïllustreerd aan de hand van de exploratie van een reële dataset die veel gebruikt wordt in onderwijs-sociologisch onderzoek: de zogenaamde uitgebreide Mathijssen-Sonnemans cohort.

1. "IGNORABLE" VERSUS "NONIGNORABLE" RESPONSE-MECHANISMEN.

De technieken die gebruikt worden ter "correctie" voor nonresponse gaan uit van twee verschillende assumpties omtrent de aard van het response-mechanisme. Het response-mechanisme kan "ignorable" of "nonignorable" zijn (Little,1982; Hagenaars,1987).

"Ignorable" nonresponse houdt in dat er geen verband bestaat tussen ~~de~~ het al dan niet responderen en het ~~de~~ te meten kenmerken. Wanneer er bijvoorbeeld gegevens ontbreken op de variabele opleiding dan mag er, onder de assumptie dat er sprake is van "ignorable" nonresponse, geen direct verband bestaan tussen het opleidingsniveau en de kans op nonresponse. Dit betekent dat de nonresponse toevallig moet zijn, ofwel binnen de totale populatie ofwel binnen subgroepen uit die populatie. In eerste geval is de nonresponse "missing completely at random" en in het tweede geval is er sprake van "missing at random" nonresponse (Rubin, 1974; Marini e.a.,1979). Bij "missing at random" mag de responsekans op de variabele opleiding bijvoorbeeld per leeftijdscategorie of voor mannen en vrouwen verschillen, terwijl bij "missing completely at random" de responsekans voor alle leeftijdscatogieën, voor beide geslachten en voor alle combinaties van deze kenmerken gelijk dient te zijn. Het zal duidelijk zijn dat "missing at random" een zwakkere assumptie is dan "missing completely at random". Bij de "missing at random" assumptie mag de nonresponse immers direct samenhangen met alle onderzoeksvariabelen behalve met de variabele waarvan sommige gegevens ontbreken, terwijl "missing completely at random" betekent dat de nonresponse met geen enkele onderzoeksvariabele direct samenhangt.

Wanneer het response-mechanisme "nonignorable" is, bestaat er wel een systematisch verband tussen bepaalde onderzoeksvariabelen en het al dan niet responderen op die variabelen. Dit betekent in het hierboven aangehaalde voorbeeld dat de kans om te responderen op de variabele opleiding verschilt voor de diverse opleidingsniveaus, met constanthouding van de andere in het causale model opgenomen variabelen.

De meeste correctiemethoden gaan uit van een "ignorable" response-mechanisme. De belangrijkste reden hiervoor is dat wanneer men een "nonignorable" response-mechanisme verkiest men precies moet aangeven hoe de nonresponse tot stand komt, hetgeen meestal niet mogelijk is omdat er te weinig inzicht bestaat in de oorzaken van de nonresponse.

Voorbeelden van correctiemethoden die uitgaan van de "missing completely at random" assumptie zijn "listwise" en "pairwise deletion of cases". Dit zijn de standaard correctiemethoden waaruit men kan kiezen bij het opstellen van een correlatie-matrix met behulp van SPSS^{x}. In het geval van de loglineaire analyse van categorische data gaat men uit van deze assumptie wanneer men de analyse uitvoert met behulp van de multivariate frequentietabel van degenen die overal op responderen. Bij deze technieken, die alle neerkomen op het in de analysefase verwijderen van de onderzoekselementen die niet responderen op bepaalde variabelen, is er geen sprake van een echte correctie voor nonresponse in de zin dat men op wat voor manier dan ook met behulp van de bekende gegevens iets tracht te zeggen over de ontbrekende data.

Omdat de "missing completely at random" assumptie over het algemeen niet realistisch is, gaan de meeste "echte" correctiemethoden uit van de zwakkere assumptie "missing at random". De nonresponse hoeft dan alleen maar binnen subgroepen van de variabelen zonder ontbrekende scores toevallig te zijn. Voorbeelden van correctie-procedures die uitgaan van deze assumptie zijn: het herwegen, het bijschatten van scores en het schatten van de "sufficient statistics".

Bij het *herwegen* wordt de onderzoeksgroep onderverdeeld in subgroepen of strata op grond van enkele bekende kenmerken. Binnen elke subgroep wordt de kans om te responderen op de variabelen met ontbrekende gegevens bepaald. De scores van de respondenten worden vervolgens gewogen met een wegingsfactor die gelijk is aan de inverse van responsekans. Op deze manier weegt de score van een respondent uit een stratum met een lage responsekans relatief zwaarder mee bij het schatten van de parameters, dan van iemand uit een subgroep met een hoge responsekans.

Ook bij het *bijschatten* van score wordt de onderzoeksgroep verdeeld in subgroepen op grond van de bekende kenmerken. Het verschil met de herwegingsprocedure is dat men geen gewichten toekent aan de respondenten, maar dat de nonrespondenten een score krijgen toegewezen. De wijze waarop de scores worden bijgeschat hangt nauw samen met het meetniveau van de onderzoeksvariabelen. Wanneer de variabelen van intervalniveau zijn kunnen de nonrespondenten als score het gemiddelde van de subgroep krijgen. Ook is het mogelijk om de ontbrekende scores modelmatig te "voorspellen" door middel van een regressievergelijking waarin de variabelen waar de scores wel van bekend zijn als de verklarende variabelen worden opgenomen. Wanneer er sprake is van variabelen van nominaal meetniveau dan kunnen de nonrespondenten ofwel aan de modale klasse worden toegewezen ofwel proportioneel worden verdeeld over de categorieën van de missing variabelen; dit alles natuurlijk binnen de subgroepen waartoe men behoort op grond van de wel gemeten kenmerken.

De derde correctieprocedure die uitgaat van de assumptie "missing at random" is het vinden van de meest aannemelijke schattingen voor de "sufficient statistics". Bij deze techniek wordt alle beschikbare informatie gebruikt bij de analyse. Wanneer de gegevens gemeten zijn op intervalniveau komt dit neer op het opstellen van aparte correlatiematrices voor de diverse subgroepen. Onderzoekspersonen behoren daarbij tot dezelfde subgroep wanneer ze op exact dezelfde deelverzameling van de onderzoeksvariabelen geantwoord hebben. Met behulp van deze correlatiematrices wordt de "maximum likelihood-schatting" voor de "echte" correlatiematrix bepaald (Marini e.a., 1979). Wanneer de gegevens van nominaal meetniveau zijn worden meerdere multivariate frequentieverdelingen opgesteld. Deze worden vervolgens gebruikt om de frequentietabel, die verkregen zou zijn wanneer er geen nonresponse zou zijn geweest, te schatten (Fuchs, 1982). Deze correctieprocedure zal in paragraaf 3 worden besproken.

Het herwegen en het proportioneel bijschatten zijn te verkiezen boven het toewijzen van nonrespondenten aan de modale klasse of het bijschatten van de ontbrekende score met behulp van het groepsgemiddelde of een regressievergelijking. Deze laatste procedures hebben namelijk het nadeel dat ze de sterkte van de verbanden tussen de onderzoeksvariabelen beinvloeden. Over het algemeen zullen, bij gegevens van intervalniveau, de correlatiecoëfficiënten verder van nul af komen te liggen wanneer men scores bijschat met behulp van het groepsgemiddelde of via een regressievergelijking. Bij categorische data verschilt het effect van modale toewijzing van nonrespondenten per geval: in een 2 bij 2 tabel bijvoorbeeld, waarin de cellen (1,1) en (2,2) het best gevuld zijn, zal het verband meestal overschat worden, terwijl het in een 5 bij 5 tabel heel moeilijk te voorspellen is wat er zal gebeuren. Het schatten van de "sufficient statistics" is echter de beste procedure, en wel om de volgende twee redenen: ten eerste levert de toetingsprocedure, waarmee een van te voren opgesteld model getoetst dient te worden, problemen op bij het herwegen en bijschatten van scores. Doordat de scores geen onafhankelijke trekkingen zijn uit de onderzoekspopulatie zal de X^2 benadering vande toetsingsgrootheid, de (log)aannamelijkheidsratio L of de Pearson-X^2, zeer twijfelachtig zijn. Dit probleem doet zich niet voor bij het schatten van de "sufficient statistics", omdat daar geen sprake is van afhankelijkheid tussen de scores van verschillende onderzoekselementen. De tweede reden dat deze correctieprocedure beter is, is dat ze "rekening houdt" met de verwachte relaties tussen de onderzoeksvariabelen. De andere correctieprocedures gaan daarentegen alle uit van het zogenaamde "verzadigde model", waarbij alle interacties tussen de onderzoeksvariabelen worden opgenomen, hetgeen leidt tot niet efficiënte schattingen (Hagenaars, 1987).

Eén van de correctiemethoden voor "nonignorable" response is de procedure die uitgedacht is door Fay (1986). Deze correctieprocedure, die later nog uitgebreid zal worden toegelicht, komt kortgezegd neer op het opstellen van een causaal model waarin "response-variabelen" worden opgenomen. Een "response-variabele" is een variabele die aangeeft of iemand al dan niet respondeert op een bepaalde onderzoeksvariabele. Afhankelijk van de causale verbanden die opgenomen worden in het model, zal het responsemechanisme "ignorable" of "nonignorable" zijn. Wanneer er een directe causale relatie bestaat tussen een variabele met onbrekende gegevens en de bijbehorende "response-variabele" dan is er sprake van "nonignorable" nonresponse. Wanneer de "response variabelen" echter alleen afhankelijk zijn van de "andere" variabelen dan is de nonresponse "ignorable".

2. LOGLINEAIRE ANALYSE.

2.1. Inleiding.

Zoals bij de regressieanalyse getracht wordt de variantie in de scores op de afhankelijke variabele te verklaren vanuit de relaties met de onafhankelijke variabelen, zo probeert men met behulp van de loglineaire analyse de geobserveerde celfrequenties in een multivariate frequentieverdeling te verklaren vanuit de relaties tussen de onderzoeksvariabelen, waarbij het evenals bij regressieanalyse mogelijk is om een causale ordening aan te brengen tussen deze variabelen.

Loglineaire analyse is een zeer geschikte analysetechniek wanneer de onderzoeksvariabelen van nominaal meetniveau zijn. Dit is dan ook meteen een van de belangrijkste voordelen van deze techniek; het meetniveau is immers altijd (tenminste) van nominaal niveau, waardoor de loglineaire analyse vrijwel altijd toe te passen is, zij het met enig "informatieverlies" wanneer zij wordt toegepast op variabelen gemeten op ordinaal-, interval- of rationiveau. Een tweede belangrijke voordeel van deze techniek is dat er, in tegenstelling tot de regressieanalyse, bijzonder zwakke assumpties gemaakt worden omtrent de verdelingen van de variabelen. De enige eis die gesteld wordt is dat de cellen van de multivariate frequentieverdeling een multinominale-, een produktmultinominale- of een poissonverdeling volgen. Een derde voordeel is dat de zogenaamde interactietermen toegestaan zijn, hetgeen bij regressieanalyse niet het geval is. Een "nadeel" is echter dat in loglineaire modellen niet al te veel variabelen kunnen worden opgenomen omdat de cellen van de multivariate frequentieverdeling voldoende "gevuld" moeten zijn om bij de schattings- en toetsingsprocedures niet al te afhankelijk te zijn van toevalsfluctuaties.

Hierboven is bij de vergelijking tussen de loglineaire analyse en de regressieanalyse vooral gelet op hun toepasbaarheid. Men moet echter zeer voorzichtig zijn bij het vergelijken van uitkomsten die resulteren uit analyses met behulp van deze twee technieken. Regressieanalyse gaat namelijk uit van een additief verklaringmodel, terwijl de loglineaire analyse gebruik maakt van een multiplicatief model. Het kan voor de verklaring die gevonden wordt nogal wat uitmaken of zij is gezocht door het optellen van effecten of door vermenigvuldiging van effectparameters.

2.2. Loglineaire analyse zonder latente variabelen.

Wanneer men een multivariate frequentieverdeling heeft van drie variabelen, A, B, en C, dan worden de geobserveerde frequenties aangeduid als f_{ijk}^{ABC} en de verwachte frequenties als F_{ijk}^{ABC}, waarbij geldt dat de verwachte frequenties gelijk zijn aan $N\pi_{ijk}^{ABC}$, ofwel de steekproefomvang maal de kans om tot een bepaalde cel te behoren. In hiërarchische modellen bepalen de relaties tussen de onderzoeksvariabelen die in een model worden opgenomen welke geobserveerde marginalen exact worden gereproduceerd. Via de "iterative proportional fitting-procedure" (Fienberg,1977; Hagenaars,1985,p.35) worden de meest aannemelijke schattingen voor de verwachte frequenties bepaald. Deze "maximum likelihood-schattingen" geeft men weer als $\hat{F}_{ijk}^{ABC}$, waarbij geldt dat $\hat{F}_{ijk}^{ABC} = N\hat{\pi}_{ijk}^{ABC}$.

Wanneer er sprake is van een "<u>verzadigd model</u>" - een model waarin alle parameters zijn opgenomen - ziet het verband tussen de verwachte frequenties en de effectparameters er als volgt uit:

$$F_{ijk}^{ABC} = \eta \tau_i^A \tau_j^B \tau_k^C \tau_{ij}^{AB} \tau_{ik}^{AC} \tau_{jk}^{BC} \tau_{ijk}^{ABC} \qquad (1)$$

waarbij n het overalleffect is; τ_i^A, τ_j^B en τ_k^C de hoofdeffecten zijn: τ_{ij}^{AB}, τ_{ik}^{AC}, τ_{jk}^{BC} de eerste orde interacties zijn; en τ_{ijk}^{ABC} het tweede orde interactie effect is. Wanneer vergelijking (1) weergeven wordt met behulp van loglineaire parameters ziet hij er als volgt uit:

$$G_{ijk}^{ABC} = \theta + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC} \qquad (2)$$

hierbij is:

$$G_{ijk}^{ABC} = \ln F_{ijk}^{ABC},\ \theta = \ln \eta,\ \lambda_i^A = \ln \tau_i^A,\ \lambda_{ij}^{AB} = \ln \tau_{ij}^{AB} \text{ en } \lambda_{ijk}^{ABC} = \ln \tau_{ijk}^{ABC}.$$

De loglineaire parameters sommeren over elk subscript tot 0, en de multiplicatieve parameters zijn vermenigvuldigd over elk subscript gelijk aan 1.

Voor het verzadigde model geldt dat geschatte verwachte frequenties, $\hat{F}_{ijk}^{ABC}$, gelijk zijn aan de geobserveerde frequenties, f_{ijk}^{ABC}. Bij niet-verzadigde modellen is dat over het algemeen niet het geval. Bepaalde lambda's worden dan immers op nul gesteld, ofwel bepaalde tau's worden a priori op één gezet, hetgeen er toe leidt dat niet de celfrequenties exact worden gereproduceerd maar alleen enkele randtotalen.

Er zijn twee soorten "onverzadigde" modellen: hiërarchische en niet-hiërarchische. Er is sprake van een hiërarchisch loglineair model wanneer bij het weglaten van een bepaald effect ook alle hogere orde interacties worden weggelaten. Dat betekent bijvoorbeeld dat wanneer het tweede orde effect τ_{ij}^{AC} niet wordt opgenomen in het model ook het hogere orde effect τ_{ijk}^{ABC} niet mag voorkomen. In niet-hiërarchische modellen komen wel hogere orde interacties voor terwijl lagere orde interacties zijn weggelaten. Wij zullen ons beperken tot de hiërarchische modellen. Deze modellen geven we weer met behulp van de te reproduceren marginale frequenties. Het voorbeeld dat hierboven werd genoemd moet dan worden aangeduid als {AB,BC}, het verzadigde model als {ABC} en het onafhankelijksmodel als {A,B,C}.

Door de geschatte verwachte frequenties die horen bij een bepaald model te vergelijken met de geobserveerde celfrequenties kan een model dat opgesteld is op zijn houdbaarheid worden getoetst. Dit kan zowel geschieden met behulp van de Pearson-X^2 als met (log)-likelihoodratio L, waarbij:

$$X^2 = \sum_i \sum_j \sum_k [(f_{ijk}^{ABC} - \hat{F}_{ijk}^{ABC})^2 / \hat{F}_{ijk}^{ABC}] \tag{3a}$$

$$L = 2 \sum_i \sum_j \sum_k [f_{ijk}^{ABC} \ln(f_{ijk}^{ABC} / \hat{F}_{ijk}^{ABC})] \tag{3b}$$

Beide benaderen asymptotisch de X^2-verdeling. Het aantal vrijheidsgraden van de toets is gelijk aan het aantal celfrequenties min het aantal onafhankelijk te schatten parameters.

Het gebruik van de (log)aannemelijksratio L verdient de voorkeur boven het gebruik van de Pearson-X^2, omdat het schatten van de effectparameters met behulp van het "maximum likelihood-principe" neerkomt op het minimaliseren van toetsingsgrootheid L; de L past dus het beste bij de gekozen schattingsprocedure. Het verstandigste is echter om ter controle beide toetsingsgrootheden tegelijkertijd te hanteren. Wanneer een model goed "fit" mogen deze toetsingsgrootheden immers niet al te veel van elkaar verschillen omdat bij een ware hypothese beide grootheden asymptotisch dezelfde X^2-verdeling volgen.

Het komt vaak voor dat een onderzoeker vooraf geen model opstelt dat getoetst dient te worden met behulp van de beschikbare data. In plaats daarvan worden de data gebruikt ter <u>exploratie</u> van het onderzoeksveld. Hiervoor bestaan twee verschillende procedures: voorwaartse en achterwaartse selectie. Voorwaartse selectie houdt in dat er begonnen wordt met een zeer "spaarzaam" model en dat er telkens effecten aan dat model toegevoegd worden, terwijl bij de achterwaartse selectie begonnen wordt met een model met veel te veel effecten en er steeds één of meer afvallen.

De (log)aannemelijksratio L is een zeer bruikbaar gereedschap bij het exploratieve werk. Door middel van <u>conditionele toetsing</u> van genestelde modellen, waarbij de respectievelijke L-waarden met elkaar vergeleken worden, kan er bekeken worden welk model het beste bij de data past. Twee modellen zijn genesteld wanneer in het uitgebreide model alle effecten opgenomen zijn uit het beperkte model. De modellen {A,B,C}, {AB,BC} en {ABC} zijn genesteld. De conditionele toetsingsgrootheid is:

$$L_{b/u} = L_b - L_u = 2 \sum_i \sum_j \sum_k [f^{ABC}_{ijk} \ln(\hat{F}^{ABC}_{ijk,u} / \hat{F}^{ABC}_{ijk,b})]$$
$$df_{b/u} = df_b - df_u \qquad (4)$$

Door de L-waarde van het beperkte model te verminderen met die van het uitgebreide model kan men nagaan of de weggelaten parameters significant zijn. De conditionele toetsingsgrootheid $L_{b/u}$ volgt asymptotisch een X^2-verdeling met $df_{b/u}$ vrijheidsgraden.

Wanneer er een causaal model wordt opgesteld waarin bepaalde variabelen zowel de rol van afhankelijke als van onafhankelijke variabele vervullen, dan kunnen de parameters van dat loglineaire model niet in één keer worden geschat. Het schatten van de parameters dient dan stapsgewijs te geschieden zoals we dat kennen uit de padanalyse. Er wordt in dit verband daarom ook wel gesproken van "modified path analyses".

Wanneer we bijvoorbeeld een model hebben met vier variabelen - A, B, C en D - waarbij de causale ordening loopt van A naar D, dan moeten we de relatie A-B bekijken binnen de tabel AB (deze tabel verkrijgt men door de celfrequenties te sommeren over C en D). Vervolgens worden de relaties tussen A, B en C geanalyseerd binnen de tabel ABC, waarbij wel het effect AB in het model moet worden opgenomen. En als laatste komt de tabel ABCD aan bod ter bestudering van de relaties tussen A, B, C en D, waarbij dan wel alle effecten tussen A, B en C moeten worden opgenomen (zie Hagenaars, 1985, pag. 55-58).

De L waarden van de modellen die in de verschillende stappen worden gevonden, kunnen evenals het aantal vrijheidsgraden worden gesommeerd. Met behulp van de op deze wijze verkregen "overall-L" kan het cauasale model in zijn geheel op zijn houdbaarheid worden getoets.

2.3. Loglineaire modellen met latente variabelen.

In deze subparagraaf worden de in de vorige subparagraaf behandelde loglineaire modellen uitgebreid naar modellen met latente variabelen: de zogenaamde latente-klasse-analyse.

Stel dat er een latente variabele X is met vier indicatoren: A, B, C en D. De relatie tussen de manifeste variabelen dient te worden verklaard vanuit de achterliggende niet gemeten factor X.

Schematisch ziet dat model er als volgt uit:

$$
\begin{tikzcd}
 & A \\
 & B \\
X \arrow[ur] \arrow[r] \arrow[dr] \arrow[ddr] & C \\
 & D
\end{tikzcd}
$$

De onderzoekspopulatie wordt daartoe verdeeld in T uitputtende en elkaar uitsluitende klassen (zie vergelijking 7). Bovendien wordt de assumptie van "lokale onafhankelijkheid" gemaakt. Dit houdt in dat de manifeste variabelen onafhankelijk van elkaar dienen te zijn bij constanthouding van de latente variabele (zie vergelijking 8). In termen van latente-klasse-analyse parameters kan het model weergegeven worden als:

$$\pi^{ABCD}_{ijkl} = \sum_{t=1}^{T} \pi^{ABCDX}_{ijklt} \tag{7}$$

waarbij geldt dat:

$$\pi^{ABCDX}_{ijklt} = \pi^{X}_{t}\pi^{\bar{A}X}_{it}\pi^{\bar{B}X}_{jt}\pi^{\bar{C}X}_{kt}\pi^{\bar{D}X}_{lt} \tag{8}$$

π^{X}_{t} is hierin de kans om tot klasse t van de latente variabele te behoren en $\pi^{\bar{A}X}_{it}$ is de conditionele kans om op A i te scoren gegevens dat men tot latente klasse t behoort.

Het equivalente loglineaire model is {AX,BX,CX,DX} en dat wordt in termen van loglineaire parameters weergegeven als:

$$F^{ABCDX}_{ijklt} = \eta\tau^{A}_{i}\tau^{B}_{j}\tau^{C}_{k}\tau^{D}_{l}\tau^{X}_{t}\tau^{AX}_{it}\tau^{BX}_{jt}\tau^{CX}_{kt}\tau^{DX}_{lt}$$

Het schatten van de verwachte frequenties F^{ABCDX}_{ijklt} kan nu niet meer geschieden via de "iterative proportional fitting-procedure", zoals dat bij de gewone loglineaire analyse gebruikelijk is, omdat er geen geobserveerde frequenties f^{ABCDX}_{ijklt} zijn. De schattingen worden nu verkregen met behulp van het <u>EM-algorithme</u>: een zeer algemeen algorithme ter bepaling van de meest aannemelijke schattingen wanneer er gegevens ontbreken. Het is een iteratieve procedure die bij elke iteratie twee stappen doorloopt. Ten eerste is er de zogenaamde <u>E-stap</u>. Hierin worden de "expected sufficient statistics" berekend op grond van de in de vorige stap gevonden schattingen voor de parameters van het model. In dit geval betekent dat, dat we met behulp van de $\hat{F}_{ijklt}^{ABCDX}$ uit de vorige iteratie en de geobserveerde frequenties f_{ijkl}^{ABCD} de niet geobserveerde frequenties f_{ijklt}^{ABCDX} schatten op de volgende manier:

$$\hat{f}_{ijklt}^{ABCDX} = (\hat{F}_{ijklt}^{ABCDX} / \hat{F}_{ijkl}^{ABCD}) f_{ijkl}^{ABCD} = \hat{\pi}_{ijklt}^{ABCD\bar{X}} f_{ijkl}^{ABCD} \tag{10}$$

Met behulp van de $\hat{f}_{ijklt}^{ABCDX}$ kunnen de te reproduceren randtotalen $\hat{f}_{it}^{AX}, \hat{f}_{jt}^{BX}, \hat{f}_{kt}^{CX}$ en $\hat{f}_{lt}^{DX}$ worden bepaald. Hiermee worden vervolgens in de <u>M-stap</u>, volgens het "maximum likelihood-principe", nieuwe $\hat{F}_{ijklt}^{ABCDX}$'s verkregen. Dit gebeurt weer met behulp van de "iterative proportional fitting-procedure". Deze schattingen voor de verwachte frequenties worden dan weer in de E-stap gebruik om de niet geobserveerde frequenties opnieuw te schatten enz. Deze prodecure blijft men herhalen totdat de uitkomsten van twee achtereenvolgende iteraties vrijwel identiek zijn. Het algorithme kan ook worden gebruikt indien het model gedefinieerd is in termen van latente-klasse-parameters (zie Hagenaars 1985, p.76-77).

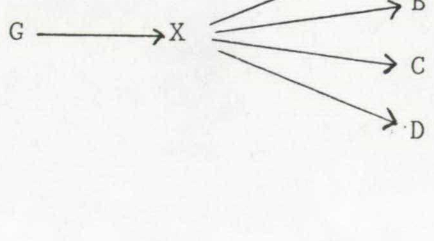
Met behulp van het programma LCAG (Hagenaars en Luijkx,1987) kan de hierboven beschreven procedure uitgevoerd worden. Men heeft daarbij de mogelijk om de eerste iteratie ofwel te laten uitvoeren met random beginschatters ofwel met zelf gekozen startwaarden voor de latente-klasse-analyse parameters π_t^X, $\pi_{it}^{\bar{A}X}$ etc. Als uitvoergegevens kan men zowel de geschatte latente-klasse-analyse parameters als de geschatte verwachte frequenties, $\hat{F}_{ijklt}^{ABCDX}$ verkrijgen. De tau- en lambda-parameters moet men zelf met de hand berekenen. Het is echter ook mogelijk om ze door een programma voor "gewone" loglineaire analyse, bijvoorbeeld ECTA laten berekenen. Daartoe kan men de $\hat{F}_{ijklt}^{ABCDX}$'s laten wegschrijven naar een uitvoerbestand dat als invoerbestand kan dienen voor ECTA.

De toetsingprocedure verloopt bij de modellen met latente variabelen op dezelfde wijze als bij modellen zonder latente variabelen. Met behulp van de toetsingsgrootheden L en X^2 kan er worden nagegaan of de geobserveerde frequenties, f^{ABCD}_{ijkl}, niet te veel afwijken van de geschatte verwachte frequenties, $\hat{F}^{ABCD}_{ijkl}$.

Er bestaat bovendien de mogelijkheid om meerdere latente variabelen op te nemen in een causaal model en om causale relaties te onderzoeken tussen zogenaamde "externe" variabelen en latente variabelen. Externe variabelen zijn buiten het latente-klasse-model gehouden manifeste variabelen. Dit lijkt erg veel op wat er gedaan wordt met behulp van Lisrel met variabelen van intervalniveau. Er wordt dan ook wel gesproken van een "<u>modified Lisrel approach</u>" (Hagenaars,1985 p.103).

Met behulp van LCAG kunnen deze Lisrel-achtige modellen worden geschat en getoetst. Daarbij dienen de externe variabelen als latente variabelen te worden gedefinieerd door gebruik te maken van de restricties die opgelegd kunnen worden aan een model. Eerst wordt daartoe het aantal latente klassen vermenigvuldigd met het aantal categorieën van de externe variabele, dan heeft men het nieuwe aantal latente klassen. Vervolgens worden bepaalde conditionele kansen die betrekking hebben op de externe variabele op 1 gesteld en andere op 0, om er voor te zorgen dat de externe variabele perfecte gerelateerd is aan de nieuwe latente. Daarnaast moeten er gelijkheidsrestricties worden geformuleerd voor bepaalde conditionele kansen voor de categorieën van de manifeste variabelen om er voor te zorgen dat de scores op de manifeste variabelen alleen afhangen van de echte latente variabele (zie Hagenaars en Luijkx,1987).

Ter illustratie nemen we hetzelfde voorbeeld als hierboven maar dan met een externe variabele G. Schematisch ziet het model er als volgt uit:



Wanneer nu alle variabelen dichotomieën zijn, dan zijn er 2x2,ofwel 4 latente klassen. De latente klassen 1 en 3 worden perfect gerelateerd aan categorie 1 van G en de klassen 2 en 4 aan categorie 2 van G. Dit doet men door de conditionele kansen $\pi_{11}^{\bar{G}X}$, $\pi_{13}^{\bar{G}X}$ $\pi_{22}^{\bar{G}X}$ en $\pi_{24}^{\bar{G}X}$ op 1 te fixeren. Daarnaast moeten er gelijkheidsrestricties opgelegd worden om de echte latente variabele terug te krijgen. Hiertoe stelt men: $\pi_{11}^{\bar{A}X} = \pi_{12}^{\bar{A}X}$, $\pi_{23}^{\bar{A}X} = \pi_{24}^{\bar{A}X}$, $\pi_{11}^{\bar{B}X} = \pi_{12}^{\bar{B}X}$, $\pi_{23}^{\bar{B}X} = \pi_{24}^{\bar{B}X}$ en hetzelfde doet men voor C en D. Door voor alle manifeste variabelen de conditionele kansen bij de latente klassen 1 en 2 en die bij 3 en 4 aan elkaar gelijk te maken, wordt ervoor gezorgd dat er per variabele maar twee verschillende waarden voor de conditionele kansen zijn. Dat zijn de conditionele kansen die horen bij de klassen van de oorspronkelijke latente variabele X.

3. FUCHS CORRECTIEPROCEDURE VOOR NONRESPONSE.

3.1. Het vinden van de meest aannemelijke schattingen voor de "sufficient statistics".

Fuchs (1982) geeft een correctiemethode voor nonresponse die uitgaat van de assumptie dat het response-mechanisme "ignorable" is. Daarbij gaat hij wel uit van de zwakste vorm, namelijk de "missing at random" variant.

Met behulp van het EM-algoritme kunnen de meest aannemelijke schattingen voor de "sufficient statistics" en voor de parameters van het loglineaire model worden gevonden. Daartoe worden de onderzoekspersonen verdeeld over elkaar wederzijds uitsluitende subgroepen, waarbij degenen die op exact dezelfde deelverzameling van de onderzoeksvragen responderen tot dezelfde subgroep behoren. Voor elke subgroep wordt vervolgens de multivariate frequentietabel opgesteld voor de variabelen waarvan de scores bekend zijn.

Stel we hebben 3 onderzoeksvariabelen: A, B en C. In dat geval zijn theoretisch de volgende subgroepen te vormen: ABC, AB, AC, BC, A, B en C. Tot subgroep ABC behoren de respondenten die op alle drie de variabelen een score hebben. Tot subgroep AB behoren al degenen die op A en B responderen maar niet op C, enz. Iedereen behoort dus precies tot één subgroep, behalve degenen waarvan helemaal niets bekend is. Meestal zullen bepaalde subgroepen leeg zijn omdat voor geen enkele respondent precies die combinatie van kenmerken bekend is.

Stel dat alleen in de subgroepen ABC, AB, BC en B onderzoekspersonen zitten en dat de andere subgroepen leeg zijn. De geobserveerde frequenties voor deze subgroepen worden aangeduid als: f_{ijk}^{ABC}, f_{ij}^{AB}, f_{jk}^{BC} en f_{j}^{B}. De omvang van de subgroepen geeft men als volg weer: N^{ABC}, N^{AB}, N^{BC} en N^{B}.

Het eerste deel van de E-stap van het EM-algorithme verloopt voor het voorbeeld van hierboven als volgt:

$$\hat{f}^{ABC}_{ijk} = f^{ABC}_{ijk} + f^{AB}_{ij}\frac{\hat{F}^{ABC}_{ijk}}{\hat{F}^{ABC}_{ij+}} + f^{BC}_{jk}\frac{\hat{F}^{ABC}_{ijk}}{\hat{F}^{ABC}_{+jk}} + f^{B}_{j}\frac{\hat{F}^{ABC}_{ijk}}{\hat{F}^{ABC}_{+j+}} \tag{11}$$

$$= f^{ABC}_{ijk} + f^{AB}_{ij}\,\hat{\pi}^{AB\bar{C}}_{ijk} + f^{BC}_{jk}\,\hat{\pi}^{\bar{A}BC}_{ijk} + f^{B}_{j}\,\hat{\pi}^{\bar{A}B\bar{C}}_{ijk}$$

In de E-stap worden eerst de nieuwe schattingen voor de niet-geobserveerde celfrequenties van de volledige tabel, $\hat{f}^{ABC}_{ijk}$, bepaald met behulp van de in de vorige iteratie verkregen schattingen voor de verwachte frequenties $\hat{F}^{ABC}_{ijk}$ (zie vergelijking 11). Daarna worden de te reproduceren marginale frequenties (de "sufficient statistics") berekend door te sommeren over de respectievelijke subschriften. In de M-stap worden vervolgens met behulp van de "iterative proportional fitting-procedure" de geschatte verwachte frequenties aangepast aan de nieuwe te reproduceren randtotalen. Dit blijft men herhalen totdat de uitkomsten van twee opeenvolgende iteraties convergeren. Het maakt in dit geval meestal niet uit welke beginschatters er gekozen worden voor $\hat{F}$. Het eenvoudigste is dan om ze allemaal op één te zetten bij de eerste iteratie. Wel moet er op worden gelet dat er geen nullen voorkomen in de te reproduceren randtotalen van de subgroep met volledige gegevens. Anders zullen, indien door de andere subgroepen wel een bijdrage wordt geleverd aan de cellen die horen bij deze randtotalen, de verkregen schattingen voor $\hat{f}$ en $\hat{F}$ afhankelijk zijn van de beginschatters (zie ook Fuchs,1982).

Wanneer het te toetsen loglineaire model het verzadigde model is en bovendien de nonresponse <u>genesteld</u> is, dan kunnen de $\hat{F}$'s op een niet iteratieve wijze worden verkregen. Genestelde patronen van nonresponse vindt men vooral bij panelonderzoek. Meedoen aan een bepaalde golf impliceert vrijwel altijd dat de respondent ook aan de voorafgaande golven heeft deelgenomen. Wanneer enkel de subgroepen ABC, BC en C zouden bestaan dan zou er in ons voorbeeld sprake zijn van een genesteld patroon. De formule die in deze gevallen gebruikt kan worden voor het schatten van de verwachte frequenties is te vinden bij Fuchs (1982) en Hagenaars (1987).

Hagenaars (1987) breidt het toepassingsgebied van de correctiemethode voor nonresponse van Fuchs uit naar loglineaire modellen met <u>latente variabelen</u>. In dat geval is er namelijk sprake van een dubbel "missing data-probleem". Ten eerste ontbreken de scores op de latente variabele voor alle respondenten en bovendien ontbreken voor sommige onderzoekspersonen de scores op één of meerdere manifeste variabelen. Hagenaars lost dit probleem op door de twee varianten van het EM-algorithme (zie vergelijkingen (10) en (11)) te combineren. In het geval van één latente variabele X, drie manifeste variabelen A, B en C en de subgroepen ABC, AB, BC en C, ziet het eerste deel van de E-stap er als volgt uit:

$$
\begin{aligned}
\hat{f}^{ABCX}_{ijkt} &= f^{ABC}_{ijk} \frac{\hat{F}^{ABCX}_{ijkt}}{\hat{F}^{ABCX}_{ijk+}} + f^{AB}_{ij} \frac{\hat{F}^{ABCX}_{ijkt}}{\hat{F}^{ABCX}_{ij++}} + f^{BC}_{jk} \frac{\hat{F}^{ABCX}_{ijkt}}{\hat{F}^{ABCX}_{+jk+}} + f^{B}_{j} \frac{\hat{F}^{ABCX}_{ijkt}}{\hat{F}^{ABCX}_{+j++}} \\
&= f^{ABC}_{ijk}\, \hat{\pi}^{ABC\bar{X}}_{ijkt} + f^{AB}_{ij}\, \hat{\pi}^{AB\bar{C}\bar{X}}_{ijkt} + f^{BC}_{jk}\, \hat{\pi}^{\bar{A}BC\bar{X}}_{ijkt} + f^{B}_{j}\, \hat{\pi}^{\bar{A}B\bar{C}\bar{X}}_{ijkt}
\end{aligned}
\tag{12}
$$

Het EM-algorithme verloopt op dezelfde wijze als hierboven is besproken.

In het het eerder genoemde programma LCAG is deze correctieprocedure voor nonresponse ingebouwd. Het enige dat men hoeft te doen wanneer men van deze procedure gebruik wil maken, is het definiëren van de subgroepen en het apart inlezen van de geobserveerde frequenties van deze subgroepen.

3.2. De toetsingsprocedure.

Wanneer er gecorrigeerd wordt voor nonresponse op de in de vorige paragraaf beschreven wijze dan kan er met behulp van de toetsingsgrootheden L en X^2 worden nagegaan of het opgestelde loglinaire model met of zonder latente variabelen houdbaar is. Om daarbij het probleem van de afhankelijkheid van waarnemingen, dat zich onder andere voordoet bij het herwegen en bij het bijschatten van scores, te voorkomen mag men echter niet, zoals Hocking en Oxpring (1974) voorstellen, $\hat{f}^{ABC}_{ijk}$ en $\hat{F}^{ABC}_{ijk}$ met elkaar vergelijken. Fuchs stelt daarom, in navolging van Chen en Fienberg (1974) en Chen (1979), voor om de toetsingsprocedure per subgroep uit te voeren. Daartoe worden met behulp van de de geschatte kansen $\hat{\pi}^{ABC}_{ijk}$ en omvang van de subgroepen, N^{ABC}, N^{AB}, N^{BC} en N^{B}, de $\hat{F}$'s voor de cellen van de frequentietabellen van de subgroepen bepaald. Vervolgens wordt voor elke groep de waarde van de (log)aannemelijksratio L en/of de Pearson-X^2 bepaald (zie vergelijking 3). De overall-toets kan worden uitgevoerd door de waarden van de diverse subgroepen te sommeren. Het aantal vrijheidsgraden van deze toets is gelijk aan het totaal aantal cellen min het aantal subgroepen min het aantal te schatten loglineaire parameters in het model (Hagenaars, 1987). De toetsingsgrootheden benaderen ook nu weer asymptotisch de theoretische X^2-verdeling.

Met behulp van de hierboven beschreven toets wordt echter niet alleen maar nagegaan of het opgestelde model houdbaar is, er wordt tegelijkertijd getoetst of de subgroepen uit dezelfde verdeling afkomstig zijn. Anders gezegd: er wordt nagegaan of de nonresponse "missing completely at random" is. Dat is op z'n minst vreemd te noemen omdat de schattingsprocedure uitgaat van de zwakkere "missing at random" assumptie. De consequentie hiervan is dat een model alleen maar op deze wijze kan worden getoetst indien de assumptie "missing completely at random" houdbaar is. Dit laatste is zeer eenvoudig na te gaan door eerst het verzadigde model, dat natuurlijk altijd aanvaard dient te worden, op te stellen. Wanneer dat model geaccepteerd wordt dan is de nonresponse inderdaad "missing completely at random" en kan de toets gebruikt worden bij het zoeken naar een passend model. Maar wordt het verzadigde model daarentegen verworpen dan kan geen enkel ander model worden getoetst met behulp van deze procedure. Volgens Little en Rubin (1897) is het bij een "missing at random" nonresponse toch mogelijk zijn om een hypothetisch model te toetsen, namelijk door middel van conditionele toetsing van dat model ten opzichte van het verzadigde model.

Als oplossing voor dit probleem stelt Hagenaars (1987), in navolging van Chen en Fienberg (1974), voor om een nonresponse model op te stellen, waarin niet alleen de verwachte frequenties $\hat{F}$ geschat worden maar ook de kansen λ. Bij elke cel van elke subgroep hoort en λ. Een λ is de kans om in een bepaalde subgroep terecht te komen, dus om op bepaalde items niet te responderen, gegeven dat de respondent een bepaalde combinatie van kenmerken heeft. In het geval dat de nonresponse "missing at random" is, hebben de λ's verschillende waarden voor de diverse cellen van de subgroepen. Wanneer echter alle λ's aan elkaar gelijk zijn is er weer sprake van "missing completely at random" nonresponse. Chen en Fienberg geven aan hoe de meest aannemelijke schattingen kunnen worden gevonden voor de kansen λ. De geschatte verwachte frequenties per subgroep bepaald men nu door de $\hat{\pi}$'s te vermenigvuldigen met de omvang van de subgroep én met de bij de cel behorende λ. De toetsingsprocedure verloop op dezelfde wijze als hierboven is beschreven, alleen moet men het aantal vrijheidsgraden verminderen met het aantal geschatte λ-parameters.

Een andere oplossing voor het toetsingsprobleem wordt gegeven door Fay (1986). De door Fay voorgestelde aanpak van nonresponse is algemener dan die van Chen en Fienberg omdat er geen assumpties gemaakt hoeven te worden over de aard van het responsemechanisme; dit houdt in dat de nonresponse zowel "ignorable" als "nonignorable" kan zijn. Bovendien sluit deze methode beter aan bij de loglineaire analyse omdat het response-model wordt weergegeven met behulp van de loglineaire parameters (τ en λ) in plaats van de kansen λ. De procedure, die in de volgende paragraaf zal worden beschreven, is veel meer dan alleen maar een oplossing voor het toetsingsprobleem. Ze biedt tevens de mogelijkheid om inzicht te krijgen in de oorzaken van nonresponse en in de gevolgen ervan voor de uitkomsten van het onderzoek.

4. CAUSALE MODELLEN MET RESPONSE-VARIABELEN.

Fay (1986) stelt voor om de nonresponse te modelleren. Daartoe moeten zogenaamde response-variabelen samen met de onderzoeksvariabelen in een causale model worden opgenomen. Een response-variabele is een variabele die aangeeft of een respondent wel of niet respondeert op een onderzoeksvariabele. Het opstellen van zo'n model gebeurt op dezelfde manier als dat altijd geschiedt bij loglineaire modellen.

Er zijn twee typen response-modellen op te stellen. Enerzijds zijn er de modellen waarbij het response-mechanisme "ignorable" is. Daarin wordt aangenomen dat er geen directe verbanden bestaan tussen de response-variabelen en de bijbehorende onderzoeksvariabelen. Zolang het response-model "ignorable" is maakt het voor de schattingen van de parameters, die de sterkte van de relaties tussen de onderzoeksvariabelen weergeven, niet uit welk response-model opgesteld wordt. Wel zullen de schattingen over het algemeen verschillen ten opzichte van de situatie waarin de response-variabelen worden weggelaten en er dus helemaal niet gecorrigeerd wordt voor nonresponse. In het geval van "ignorable" nonresponse doen zich geen extra identificatieproblemen voor. Als het model zonder response-variabelen identificeerbaar is dan, zal ook het model met de response-variabelen dat zijn. Anderzijds zijn er de modellen met een "nonignorable" response-mechanisme. In deze modellen worden wel directe relaties verondersteld tussen de onderzoeksvariabelen met ontbrekende gegevens en de daarbij horende response-variabelen. In het geval van een "nonignorable" response-mechanisme zullen de parameters, die relaties tussen de onderzoeksvariabelen weergeven, wel afhankelijk zijn van het opgestelde causale model voor nonresponse. Bovendien kunnen er identificeerbaarheidsmoeilijkheden optreden bij response-modellen van dit type.

Het schatten van de loglineaire parameters geschiedt in dit geval eveneens met behulp van het EM-algorithme, alleen worden de $\hat{f}$'s op een iets andere manier bepaald. Ter illustratie zullen we hetzelfde voorbeeld als in §3 nemen. Daarin hadden we drie onderzoeksvariabelen A, B en C en vier subgroepen ABC, AB, BC en B, met de daarbij horende frequenties f^{ABC}_{ijk}, f^{AB}_{ij}, f^{BC}_{jk} en f^{B}_{j}. Dat betekent dat er twee response-variabelen opgenomen moeten worden in het model omdat er bij twee variabelen, bij A en bij C, gegevens ontbreken. Het al dan niet responderen op A zullen we R noemen, waarbij score 1 betekent wel responderen en 2 niet responderen, en de response-variabele die hoort bij C noemen we S. Het eerste deel van de E-stap van het EM-algoritme wordt door Fay als volgt gedefiniëerd:

$$\hat{f}^{ABCRS}_{ijk11} = f^{ABC}_{ijk} \qquad (13a) \qquad\qquad \hat{f}^{ABCRS}_{ijk12} = f^{AB}_{ij}\,\frac{\hat{F}^{ABCRS}_{ijk12}}{\hat{F}^{ABCRS}_{ij+12}} \qquad (13b)$$

$$\hat{f}^{ABCRS}_{ijk21} = f^{BC}_{jk}\,\frac{\hat{F}^{ABCRS}_{ijk21}}{\hat{F}^{ABCRS}_{+jk21}} \qquad (13c) \qquad\qquad \hat{f}^{ABCRS}_{ijk22} = f^{B}_{j}\,\frac{\hat{F}^{ABCRS}_{ijk22}}{\hat{F}^{ABCRS}_{+j+22}} \qquad (13d)$$

Zoals uit de vergelijkingen hierboven blijkt bestaat het eerste deel van de E-stap, waarin de niet geobserveerde frequenties $\hat{f}^{ABCRS}_{ijklm}$ worden bepaald, in dit voorbeeld uit vier afzonderlijke delen. Steeds wordt er met behulp van de geobserveerde frequenties uit een bepaalde subtabel een deel van de tabel van niet geobserveerde frequenties geschat. Het vervolg van het EM-algorithme verloopt op dezelfde wijze als we dat eerder gezien hebben. In de E-stap worden eerst nog de "sufficient statistics" bepaald waarna in de M-stap de $\hat{F}^{ABCRS}_{ijklm}$'s op de bekende wijze daaraan worden aangepast. Vervolgens gebruikt men de geschatte verwachte frequenties in de E-stap om de $\hat{f}$'s te verbeteren, enz. Zoals Hagenaars (1987) de procedure van Fuchs uitbreidt naar loglineaire modellen met latente variabelen, zo is de hierboven beschreven procedure eveneens toepasbaar te maken voor modellen met latente variabelen. Wanneer we aan het voorbeeld van hierboven een latente variabele X toevoegen, dan ziet de E-stap van het EM-algorithme er als volgt uit:

$$\hat{f}^{ABCRSX}_{ijk11t} = f^{ABC}_{ABC} \frac{\hat{F}^{ABCRSX}_{ijk11t}}{\hat{F}^{ABCRSX}_{ijk11+}} \quad (14a) \qquad \hat{f}^{ABCRSX}_{ijk12t} = f^{AB}_{ij} \frac{\hat{F}^{ABCRSX}_{ijk12t}}{\hat{F}^{ABCRSX}_{ij+12+}} \quad (14b)$$

$$\hat{f}^{ABCRSX}_{ijk21t} = f^{BC}_{jk} \frac{\hat{F}^{ABCRSX}_{ijk21t}}{\hat{F}^{ABCRSX}_{+jk21+}} \quad (14c) \qquad \hat{f}^{ABCRSX}_{ijk22t} = f^{B}_{j} \frac{\hat{F}^{ABCRSX}_{ijk22t}}{\hat{F}^{ABCRSX}_{+j+22+}} \quad (14d)$$

Met behulp van het door ons gebruikte programma LCAG kan, ondanks de iets andere implementatie van het EM-algorithme (zie vergelijking 12), Fay's techniek worden uitgevoerd. Men moet daarbij gebruik maken van de gewone missing data parameters. Het enige verschil met de eerder beschreven toepassing is dat er één of meerdere response-variabelen moeten worden gecreëerd. Dat doet men door de subgroepen anders te definiëren: de subgroep ABC wordt nu ABCRS, de subgroep AB wordt ABRS, BC wordt BCRS en B wordt BRS. De frequenties in de subgroep ABCRS, de f^{ABCRS}_{IJKLM}'s, zijn voor R=1 en S=1 gelijk aan de f^{ABC}_{ijk}'s en voor de rest gelijk aan 0. In subgroep ABRS zijn de geobserveerde frequenties voor R=1 en S=2 gelijk aan die van subgroep AB en voor de rest gelijk aan nul. Op dezelfde manier worden de frequenties in subgroepen BCRS en BRS met behulp van nullen opgevuld. Nu kan op de normale manier een causaal model worden gedefinieerd waarbij indien nodig de parameters stapgewijs worden geschat. Het EM-algorithme verloop op de wijze als dat in de vorige paragraaf is beschreven. Met behulp van vergelijking (15) worden de $\hat{f}^{ABCRSX}_{ijklmt}$'s bepaald. Dit is dezelfde vergelijking als vergelijking (12) alleen zijn er de variabelen R en S aan toegevoegd.

$$\hat{f}^{ABCRSX}_{ijklmt} = f^{ABCRS}_{ijklm} \frac{\hat{F}^{ABCRSX}_{ijklmt}}{\hat{F}^{ABCRSX}_{ijklm+}} + f^{ABRS}_{ijlm} \frac{\hat{F}^{ABCRSX}_{ijklmt}}{\hat{F}^{ABCRSX}_{ij+lm+}} + f^{BCRS}_{jklm} \frac{\hat{F}^{ABCRSX}_{ijklmt}}{\hat{F}^{ABCRSX}_{+jklm+}} + f^{BRS}_{jlm} \frac{\hat{F}^{ABCRSX}_{ijklmt}}{\hat{F}^{ABCRSX}_{+j+lm+}} \quad (15)$$

Dit zijn dezelfde $\hat{f}$'s als die verkregen worden met behulp van de vergelijkingen (14a) tot en met (14d), die de uitbreiding naar loglineaire modellen met latente variabelen vormen van de door Fay gegeven vergelijkingen (13a) tot en met (13d). Dit is als volgt aan te tonen: het eerste onderdeel aan de rechterzijde van vergelijking (15) komt overeen met vergelijking (14a): alle waarnemingen in de subgroep ABCRS zitten immers in de cellen waarvoor geldt dat R=1 en S=1 en de rest zijn allemaal nullen. Het tweede onderdeel, waarvoor geldt dat alleen bij R=1 en S=2 de cellen gevuld zijn, komt overeen met vergelijking (14b), die betrekking heeft op de subgroep AB. Het derde en vierde deel zijn vrijwel identiek aan (14c) en (14d). In vergelijking (15) worden de bijdragen van de diverse subgroepen aan de geschatte geobserveerde frequenties bij elkaar opgeteld. Dit levert hetzelfde resultaat op als het samenvoegen in een tabel van de uitkomsten uit vergelijkingen (14a) tot en met (14d); de lege cellen leveren immers geen bijdrage aan $\hat{f}^{ABCRSX}_{ijklmt}$.

De houdbaarheid van een model kan, net zoals in modellen zonder response-variabelen, met behulp van de toetsingsgrootheden L en Pearson-X^2 worden <u>getoetst</u>. Daartoe worden per subgroep de verwachte en geobserveerde frequenties met elkaar vergeleken (zie paragraaf 3.2.). Voor subgroep ABC betekent dat dat de geobserveerde frequenties f^{ABC}_{ijk} , ofwel f^{ABCRS}_{ijk11}, met de geschatte verwachte frequenties $\hat{F}^{ABCRS}_{ijk11}$ moeten worden vergeleken. Voor subgroep AB zijn dat f^{AB}_{ij} en $\hat{F}^{ABCRS}_{ij+12}$, voor subgroep BC f^{BC}_{jk} en $\hat{F}^{ABCRS}_{+jk21}$ en voor subgroep B f^{B}_{j} en $\hat{F}^{ABCRS}_{+j+22}$. Het aantal <u>vrijheidsgraden</u> voor deze toets is: het aantal geobserveerde frequenties min het aantal subgroepen min het aantal te schatten loglineaire parameters, met inbegrip van de parameters die nodig zijn bij het opstellen van een response-model.

Met behulp van LCAG kan de hierboven beschreven toetsingsprocedure, zei het op een iets andere manier, worden uitgevoerd. Het programma berekent de toetsingsgrootheden met behulp van de volledige tabellen $\hat{F}^{ABCRS}_{ijklm}$ en $\hat{f}^{ABCRS}_{ijklm}$. Voor de subgroep ABC maakt dat niets uit omdat ook nu $\hat{F}^{ABCRS}_{ijk11}$ en $\hat{f}^{ABCRS}_{ijk11}$, ofwel f^{ABC}_{ijk}, met elkaar worden vergeleken. Bij de andere subgroepen ligt dat anders;

daar zouden de geschatte geobserveerde en de geschatte verwachte frequenties allereerst gesommeerd moeten worden over de variabelen waarvoor de gegevens ontbreken. Voor de subgroep AB betekent dat, dat $\hat{F}^{ABCRS}_{ijkl2}$ en $\hat{f}^{ABCRS}_{ijkl2}$ gesommeerd dienen te worden over C, zodat met behulp van $\hat{F}^{ABRS}_{ijl2}$ en f^{ABRS}_{ijl2} de toetsingsgrootheid L of X^2 berekend kan worden. Het blijkt echter niet nodig te zijn om te sommeren over de variabele met missing data: de waarde van toetsingsgrootheid blijft gelijk. Dit komt omdat de geobserveerde en de verwachte frequenties van de subgroepen, bijvoorbeeld f^{ABRS}_{ijl2} en $\hat{F}^{ABRS}_{ijl2}$, met dezelfde conditionele kansen, $\hat{\pi}^{AB\bar{C}RS}_{ijkl2}$, vermenigvuldigd zijn om de geobserveerde en de verwachte frequenties in de volledige tabel, $\hat{f}^{ABCRS}_{ijkl2}$ en $\hat{F}^{ABCRS}_{ijkl2}$, te schatten. Hier is de subgroep AB als voorbeeld genomen, maar hetzelfde geldt voor de subgroepen BC (R=2 en S=1) en B (R=2 en S=2). De toetsingsgrootheid wordt in LCAG weliswaar goed berekend, het aantal vrijheidsgraden dat het programma geeft is niet juist. Het programma beschouwt het aantal cellen in de tabel ABCRS als aantal geobserveerde frequenties waarmee het aantal vrijheidsgraden dient te worden bepaald. Het werkelijke aantal geobserveerde frequenties wordt natuurlijk verkregen door het aantal cellen van de diverse subgroepen bij elkaar op te tellen. De omvang van de fout die gemaakt wordt bij de berekening van het aantal vrijheidsgraden is exact het verschil tussen deze twee.

Met behulp van de hierboven beschreven procedure toetst men twee dingen tegelijk: de adequaatheid van het response-model en de houdbaarheid van het opgestelde model voor de onderzoeksvariabelen. De beste strategie om deze twee zaken uit elkaar te houden is om eerst alleen het response-model te toetsen door voor de onderzoeksvariabelen het "verzadigde" model op te stellen. Wanneer het response-model geaccepteerd wordt kan vervolgens het model voor de onderzoeksvariabelen op zijn houdbaarheid worden getoetst.

5. EEN VOORBEELD AAN DE HAND VAN REELE DATA.

5.1. De data.

De data, die zullen worden gebruikt ter illustratie van het belang van de in de vorige paragrafen beschreven techniek, zijn afkomstig uit de zogenaamde uitgebreide Matthijssen-Sonnemans-cohort. Dit is een longitudinaal databestand met zowel school- als beroepsloopbaangegevens. In 1952 werden bij 5823 leerlingen uit de zesde klas van het gewoon lagere onderwijs in Brabant schoolvorderingstests en intelligentietests afgenomen. Bovendien vulden de schoolhoofden voor 5387 leerlingen een enquêtelijst in waarin onder andere naar het beroep van de vader werd gevraagd. In 1957 verzamelden Matthijssen en Sonnemans nieuwe informatie, onder andere over de opleiding van de ouders, bij de leerlingen die hoger dan gemiddeld scoorden op de gesommeerde schoolvorderingstests. In 1958 en 1959 werd dezelfde informatie verzameld bij boeren- en arbeiderszonen. Van 2740 onderzoekspersonen had men alleen informatie uit 1952 en van 3059 personen uit zowel 1952 als uit de tweede ondervragingsronde.

tabel 5.1: aantal respondenten bij de diverse ondervragingen.

ondervragingsronde	aantal
eerste golf (1952)	5387
tweede golf (1957/1958/1959)	3059
derde golf (1983)	2587

In 1983 is er door het I.V.A. in opdracht van Hartog en Dronkers door middel van een enquête nieuw materiaal verzameld, met name over het bereikte opleidingsniveau en het beroep van de onderzoekspersonen. Van 2587 respondenten zijn er bruikbare gegevens over deze kenmerken (Dronkers en Bakker, 1986).

Acht variabelen worden in dit voorbeeld gebruikt: de drie intelligentie- en schoolvorderingstests uit 1952, het beroep van de vader, het geslacht van de respondent (beide gemeten in 1952), de opleiding van de vader (bij de intelligente groep gemeten in 1957 en bij arbeiders- en boerenzonen in 1958 en 1959), de opleiding en het beroep van de respondenten (beide gemeten in 1983).

Deze dataset is te beschouwen als een panelstudie met drie golven. We hebben te maken met een aanzienlijke nonresponse in de tweede golf, waarin de opleiding van de vader wordt gemeten, en in de derde golf, waarin naar het beroep en de opleiding van de respondent wordt gevraagd. Hier beperken we ons tot deze paneluitval en laten de nadere bestudering van de item-nonresponse op de diverse variabelen achterwege.

Van de uitval in de tweede golf weten we dat die direct samenhangt met variabele intelligentie, in eerste instantie werd immers alleen bij intelligente kinderen nagegaan wat de opleiding van de ouders was, en met het beroep van de vader omdat later ook bij arbeiders- en boerenkinderen naar de opleiding van de ouders is gevraagd. Er is dus zeker geen sprake van "missing completely at random" nonresponse. Wel mag worden aangenomen dat de nonresponse "ignorable" is. De uitval bij de variabele opleiding vader is immers niet direct afhankelijk van het opleidingsniveau van de vader, maar enkel van andere onderzoeksvariabelen, namelijk de intelligentie van respondent en het beroep van de vader. Het is een stuk moeilijker om een verklaring te geven voor de uitval in de derde golf. Het zou kunnen dat de nonresponse geheel toevallig is, "missing complety at random" dus. Het is echter ook heel goed denkbaar dat het niet responderen op de variabelen opleiding en beroep direct samenhangt met het opleidings- en beroepsniveau van de respondenten. In dat geval zou er sprake zijn van "nonignorable" nonresponse.

Omdat het aantal categorieën van de door ons gebruikte variabelen veel te groot was om loglineaire analyse op uit te voeren, moesten de variabelen opnieuw worden gecodeerd. Daarbij is gestreefd naar zo weinig mogelijk categorieën per variabele zonder een al te groot verlies aan informatie. De drie intelligentietest (A, B en C), die zullen dienen als indicatoren voor de dichotome latente variabele intelligentie (X), zijn alle gedichotomiseerd, waarbij degenen die lager dan de mediaan scoorden zijn ingedeeld in categorie 1 en degenen die hoger dan de mediaan scoorden in categorie 2. De variabele beroep vader (D) heeft 4 categorieën: 1 is het hoger en middelbaar personeel, 2 zijn de zelfstandige ambachtslieden en boeren, 3 zijn de mensen in loondienst en 4 is de rest. Opleiding vader (E) heeft 2 categorieën: 1 is tot en met lagere school en 2 is meer dan lagere school. Opleiding van de respondent (G) is een dichotomie waarbij degenen met een middelbare beroepsopleiding of hoger tot categorie 2 horen en degenen met een lagere opleiding tot categorie 1. Het beroep van de respondendent (H) is gemeten met behulp van de ARBI-functie niveau indeling die 7 niveaus kent. Wij hebben de niveaus 1 tot en met 4 en de niveaus 5 tot en met 7 samengevoegd tot respectievelijk de categorieën 1 en 2. De codering van de variabele geslacht (F) spreekt voor zich. Voor nadere informatie over de codering zie bijlage I.

5.2. De analyse van de "volledige" tabel.

Allereerst wordt er een analyse uitgevoerd met behulp van de gegevens van de respondenten die op alle variabelen hebben gerespondeerd (tabel II.1 van bijlage II). De analyse geschiedt via exploratie van de data zoals dat is beschreven in paragraaf 2.2.. Daarbij wordt zowel gebruikt gemaakt van achterwaartse als voorwaartse selectie van effectparameters. Omdat het de opgestelde modellen stapsgewijs, met behulp van de "modified path analyses", dienen te worden geschat is het van belang om allereerst de causale ordening tussen de onderzoeksvariabelen vast te stellen. Eerst komen de drie achtergrond variabelen beroep vader(D), opleiding vader(E) en geslacht(F), die zijn van gelijk causaal niveau. Daarna volgt de latente variabele intelligentie (X), dan opleiding van de respondent(G) en als laatste in het causale schema komt het beroep van de respondent(H). De analyse is uitgevoerd met behulp van LCAG, omdat het daarin zowel mogelijk is om latente variabelen op te nemen in het model als om de parameters stapgewijs te schatten.

Om een model te vinden dat goed past bij de data gingen we als volgt te werk. Eerst moest worden nagegaan of de relaties tussen de variabelen A, B en C konden worden verklaard vanuit de latente variabele X. Daartoe werd een model opgesteld (model 1) waarin de variabelen A, B en C alleen direct afhankelijk waren van X en waarin verder alle relaties tussen D, E, F, G, H en X werden opgenomen. Dit model past goed bij de data (zie tabel 5.2). Daarna werden achtereenvolgens de relaties in de marginale tabellen DEF, DEFX, DEFGX en DEFGHX onderzocht. Daarbij werd model 1 als basis genomen. De opgestelde modellen verschilden steeds maar in één marginale tabel ten opzichte van dat model. Door conditionele toetsing van deze modellen ten opzichte van model 1 kon de houdbaarheid van het opgestelde model voor die tabel worden nagegaan. De op deze wijze gevonden L-waarde is de L-waarde die hoort bij het opgestelde model voor die marginale tabel. Bovendien zijn er, wanneer meerdere genestelde modellen pasten bij de data, conditionele toetsen uitgevoerd tussen deze alternatieve modellen. In bijlage IV zijn de toetsingsresultaten voor alle modellen en voor de conditionele toetsen te vinden: de modellen 2a t/m 2c hebben betrekking op de "margin" DEF, 3a t/m 3e op DEFX, 4a t/m 4h op DEFGX en 5a t/m 5j op de "margin" DEFGHX.

tabel 5.2: Toetsingsresultaten van de exploratie van de subgroep zonder ontbrekende gegevens.

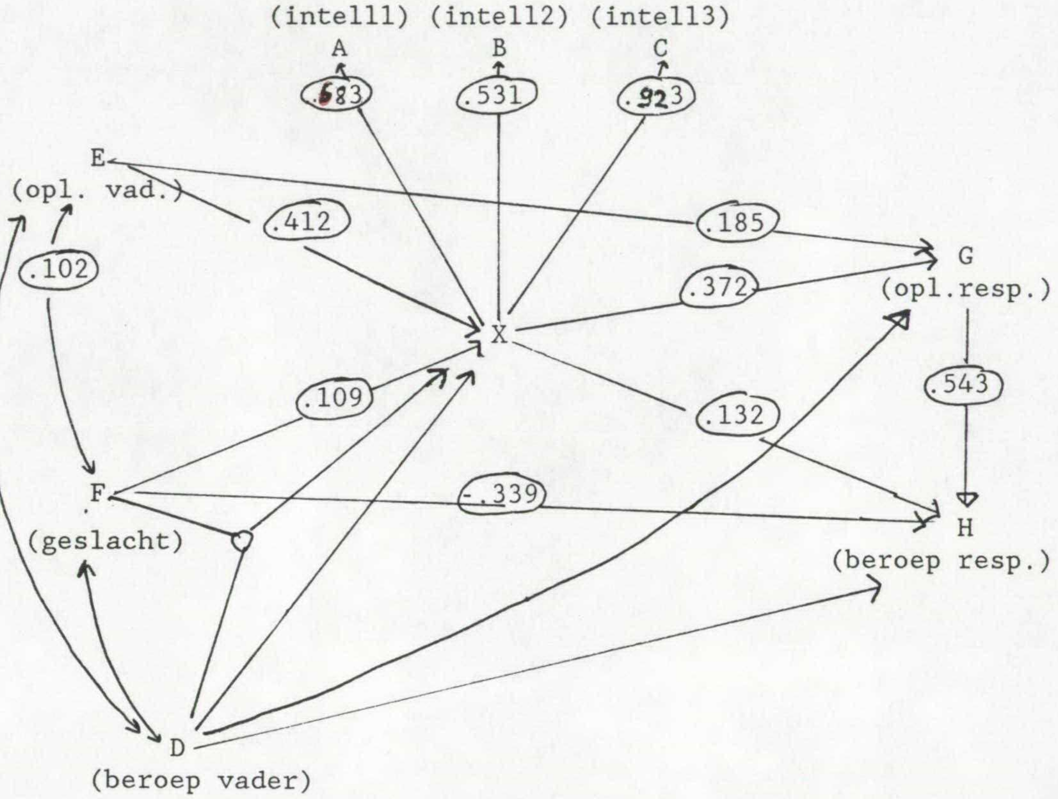
Tabel	Model	L	df	p(x)
ABCDEFGHX	(1) {AX,BX,CX,DEFGHX}	363.6	378	.694
DEF	(2b-1) {DE,DF,EF}	3.8	3	.25
DEFX	(3d-1) {DEF,DFX,EX}	3.0	7	.9
DEFGX	(4a-1) {DEFX,DG,EG,GX}	24.9	26	.5
DEFGHX	(5a-1) {DEFGX,DH,FH,HX}	54.1	57	.3
ABCDEFGHX	(1a) {AX,BX,CX,DEFGHX} {DE,DF,EF}{DEF,DFX,EX} {DEFX,DG,EG,GX} {DEFGX,DH,FH,HX}	451.9	471	.729

In tabel 5.2 wordt het "best passende" model gegeven. Model 1 geeft weer dat de relaties tussen de 3 intelligentietesten verklaard kunnen worden vanuit de latente variabele intelligentie. In de marginale tabel DEF blijkt er een sterk verband te bestaan tussen de opleiding van de vader en het beroep van de vader (zie schema 5.1) en een zwak verband (λ=.102) tussen de opleiding van de vader en het geslacht van de respondent: meisjes hebben vaders met een hogere opleiding dan jongens. Bovendien is er een relatie tussen het beroep van de vader en het geslacht van de respondent, jongens komen relatief vaker uit gezinnen van boeren of zelfstandige ambachtslieden (λ=.143) en meisjes uit de restcategorie van beroep vader (λ=.-230). Deze twee laatste relaties zijn erg vreemd, maar conditionele toetsing wees uit dat ze wel degelijk in het model moeten worden opgenomen. Waarschijnlijk zijn ze het gevolg van een niet geheel toevallige uitval.

In de "margin" DEFX bestaat er een drie-variabelen-interactie tussen beroep vader, geslacht en intelligentie. Dat betekent dat het effect van beroep vader op intelligentie voor jongens anders is dan voor meisje. Voor de totale groep geldt dat kinderen uit hogere milieus het hoogst scoren op de variabele intelligentie (λ=-.498), kinderen waarvan de vader tot de restgroep behoord scoren gemiddeld (λ=.013), en kinderen van zelfstandigen en arbeiderskinderen zijn het minst intelligent (λ=.214 en λ=.271). Bij de jongens zijn deze effecten nog groter: wanneer hun vader tot de hogere beroepsgroep behoort is het effect van beroep vader op intelligentie -.702 (=(-.498)+(-.204)), als hun vader zelfstandige of boer is .449, wanneer hun vader arbeiders is .660 en als hun vader tot de restgroep behoort -.312. Voor meisjes is het effect van beroep vader op intelligentie juist kleiner als voor de totale groep. Deze drie-variabelen-interactie is waarschijnlijk het gevolg van een niet geheel toevallige uitval. Meisjes zijn bovendien iets intelligenter dan jongens met constanthouding van de andere variabelen (λ=.109). Verder is er een groot direct effect van het opleidingsniveau van de vader op de intelligentie van de respondent (λ=.412).

Het opleidingsniveau van de respondent wordt direct beinvloed door de opleiding van de vader (λ=.185) en door het beroep van de vader, respondenten uit hogere milieus hebben een hogere opleiding (λ=.-.399) en respondenten uit een arbeidersmilieus een lagere opleiding (λ=.389) dan de rest. Bovendien bestaat er een directe samenhang tussen de intelligentie en het opleidingsniveau van de respondenten (λ=.372).

Schema 5.1: <u>Het best passende model.</u>



$\lambda^{DE}_{11}=-.845$	$\lambda^{DE}_{21}=.359$	$\lambda^{DE}_{31}=.498$	$\lambda^{DE}_{41}=-.012$
$\lambda^{DF}_{11}=.000$	$\lambda^{DF}_{21}=.143$	$\lambda^{DF}_{31}=.086$	$\lambda^{DF}_{41}=-.230$
$\lambda^{DFX}_{31}=-.204$	$\lambda^{DFX}_{31}=.235$	$\lambda^{DFX}_{311}=.294$	$\lambda^{DFX}_{31}=-.325$
$\lambda^{DX}_{11}=-.498$	$\lambda^{DX}_{21}=.214$	$\lambda^{DX}_{31}=.271$	$\lambda^{DX}_{41}=.013$
$\lambda^{DG}_{11}=-.399$	$\lambda^{DG}_{21}=-.058$	$\lambda^{DG}_{31}=.389$	$\lambda^{DG}_{41}=.067$
$\lambda^{DH}_{11}=-.185$	$\lambda^{DH}_{21}=-.104$	$\lambda^{DH}_{31}=.160$	$\lambda^{DH}_{41}=.129$

Het beroepsniveau van de onderzoekspersonen (H) is afhankelijk van het beroep van de vader: respondenten met een vader met de hoger beroep of met een vader die boer of zelfstandige is, scoren hoger op de beroepsschaal dan arbeiderskinderen of de restcategorie. Mannen hebben gemiddeld hogere beroepen dan vrouwen (λ=-.339), intelligente respondenten hebben hogere beroepen dan niet intelligente (λ=.132) en mensen met een hoge opleiding hebben hogere beroepen dan mensen met een lage opleiding (λ=.543), steeds met contstanthouding van alle andere variabelen.

De L-waarde die men krijgt voor het totale model (la) is onderaan in tabel 5.2 te vinden (L=451.9). Deze zou gelijk moeten zijn aan de waarde die verkregen wordt door optelling van de L-waarden van de diverse deelmodellen, die de stappen in de "modified path analyses" vormen, maar door afrondingsfouten klopt dat niet helemaal (L=449.4). De L-waarden voor de tabellen DEF, DEFX, DEFGX en DEFGHX zijn verkregen door conditionele toetsing van de modellen voor deze tabellen ten opzichte van model 1.

Wat opvalt is dat het model zeer goed past bij de data (p(x)=.729). Wel moet enige voorzichtheid in acht worden genomen in verband met het grote aantal cellen zonder of met te weinig waarnemingen. Bovendien wijken de waarden van de Pearson-X^2 nogal af van de L-waarden (zie bijlage IV). Een ander aspect dat niet uit het oog mag worden verloren is het exploratieve karakter van de hier uitgevoerde analyse. Het door ons gevonden model zou opnieuw moeten worden getoetst met behulp van een nieuwe onafhankelijke steekproef.

5.3. Correctie voor nonresponse op door de Fuchs voorgestelde wijze.

In deze subparagraaf zullen de correctieprocedure van Fuchs en de problemen die zich daarbij voordoen worden geillustreerd. Daartoe moeten de onderzoekspersonen ingedeeld worden in subgroepen op grond van de variabelen waarop ze hebben gerespondeerd. Hier hebben we te maken met vier subgroepen, namelijk: 1] ABCDEFGH, 2] ABCDFGH, 3] ABCDEF en 4] ABCDF. In de eerste subgroep zitten degenen waarvan alles bekend is, in de tweede degenen waarbij de opleiding van de vader niet is gemeten (niet deelgenomen aan de tweede golf), in de derde degenen waarvoor de gegevens over opleiding en beroep ontbreken (niet deelgenomen aan de derde golf) en tot de vierde groep behoren de respondenten waarvan noch de opleiding van de vader noch de eigen opleiding en beroep bekend zijn. De frequentietabellen die horen bij deze subgroepen zijn in vinden in bijlage II, tabel 1 t/m 4. Uit de totalen van deze tabellen blijkt dat het de moeite waard is om ook degenen waarvoor gegevens ontbreken in de analyse te betrekken. In subgroep ABCDEFGH, de groep waarop de vorige analyse betrekking had, zitten immers maar 1279 respondenten terwijl de oorspronkelijke onderzoeksgroep uit meer dan 5000 personen bestaat. Door de correctiemethode van Fuchs toe te passen worden 3208 onderzoekspersonen extra bij de analyse betrokken.

Voordat we de eigenlijke analyse uitvoeren zullen we aan de hand van de relatieve verdelingen van de variabelen bekijken waarin de subgroepen van elkaar verschillen. Deze relatieve frequenties zijn te vinden in bijlage III. Het is bekend dat bij de tweede golf in eerste instantie alleen intelligente kinderen en later ook kinderen uit boeren- en zelfstandigengezinnen en uit arbeidersmilieus naar de opleiding van de ouders is gevraagd. Dit komt ook tot uitdrukking in de relatieve frequenties van de subgroepen 1(ABCDEFGH) en 2(ABCDFGH). In subgroep 1 zitten relatief meer respondent in de categorie 2(=hoog) van de intelligentievariabelen A, B en C, respectievelijk 65,0%, 61,8% en 63,9%, dan in subgroep 2, waar 24,7%, 43,1% en 31,2% hoog scoort op deze variabelen. Wat betreft het beroep van de vader verschillen deze subgroepen niet zoveel. Alleen het aantal respondenten waarvan de vader in de restcategorie(4) valt is een stuk hoger in subgroep 2 dan in subgroep 1. Dit komt omdat deze respondenten een kleine kans hadden om ondervraagt te worden als ze niet tot de intelligente groep hoorden. Verder valt op dat relatief meer mannen dan vrouwen hebben deelgenomen aan de tweede golf. In subgroep 1 is het percentage mannen 69,0%, terwijl dat in subgroep 2 maar 49,8% is. Dit verschil is niet te verklaren op grond van de informatie die we hebben over de uitval in de tweede golf. Ook zitten relatief meer respondenten met een lage opleiding en een lager beroep in subgroep 2.

Subgroep 3 verschilt niet erg veel van subgroep 1 wat betreft de relatieve verdelingen van de afzonderlijke variabelen. Alleen het percentage mannen is kleiner in subgroep 3 (50,5%) dan in subgroep 1. Dit is niet vreemd omdat de onderzoekers in de derde golf bij mannen meer moeite hebben gedaan om de gewenste informatie te krijgen dan bij vrouwen: alleen mannelijke nonrespondenten werden een tweede maal benaderd (Hartog,1986). Subgroep 4 lijkt sterk op subgroep 2, alleen zijn daar de afwijkingen ten aanzien van subgroep 1 nog extremer. Het aantal intelligente respondenten is nog kleiner (18,9%, 35,3% en 26,1%) en hetzelfde geldt voor het relatieve aantal mannen (27,7%).

Wanneer we de analyse uit de vorige paragraaf nogmaals uitvoeren, maar dan mét gebruikmaking van de gegevens van de respondenten waarvoor een deel van de gegevens ontbreekt, dan kunnen zowel de parameterschattingen als de toetsingsresultaten veranderen.
Om de effecten van de uitval bij de tweede en de derde golf op deze schattingen en toetsingsuitkomsten uit elkaar te kunnen houden, zijn er drie aparte analyses uitgevoerd. De eerste met behulp van de subgroepen 1 en 2, om het effect van de uitval in de tweede golf te bekijken, de tweede met de subgroepen 1 en 3, om het effect van de uitval in de derde golf te kunnen zien, en de laatste met alle vier de subgroepen tegelijk. Steeds zijn er twee modellen opgesteld, namelijk het model dat gevonden werd in de vorige paragraaf en het model dat we in de vorige paragraaf model 1 noemden: het model waarin het bestaan van de latente variabele intelligentie getoets werd en waarin verder het verza-

digde model gold.

In alle drie de analyses moesten steeds beide modellen worden verworpen (voor de toetsingresultaten in bijlage V). Dit wijst er op dat de nonresponse niet "missing completely at random" is, maar enkel "missing at random". Dit komt overeen met onze verwachting omtrent de aard van de nonresponse. We verwachtten immers op grond van de relatieve frequenties dat de uitval bij variabele opleiding vader zou samenhangen met het beroep van de vader en met de intellingentie en het van de respondent, en dat de uitval bij de opleiding en het beroep van de respondent zou afhangen van het geslacht van de respondenten.

Weliswaar konden de modellen door het niet geheel toevallig zijn van de uitval niet worden getoetst op hun houdbaarheid, het was natuurlijk wel mogelijk om de omvang van de effecten te schatten. Dit is gedaan voor het in de vorige paragraaf gevonden model. Bij de analyse met behulp van de subgroepen 1 en 2 valt op dat de directe verbanden tussen beroep vader en geslacht (λ=-.015, λ=.032, λ=.002 en λ=-.019) en tussen geslacht en intelligentie (λ=.030) vrijwel verdwijnen en dat de drie-variabelen-interactie DFX (λ=-.156, λ=.042, λ=.077 en λ=.036) veel kleiner wordt. Bij de subgroepen 1 en 3 wordt het verband tussen de opleiding van de vader en het geslacht van de respondent (λ=.035) erg klein. Wanneer we alle subgroepen samen nemen verdwijnen de vreemde interacties DF (λ=.023, λ=-.014, λ=-.008 en λ=-.001) en EF (λ=-.007) helemaal, verder wordt de interactie FX (λ=.050) vrijwel 0 en wordt de drie-variabelen-interactie EFX (λ=-.045, λ=-.022, λ=-.046 en λ=.113) erg klein. Alle drie de modellen met de daarbijhorende λ parameters zijn te vinden in bijlage V.

Dit alles wijst er op dat de analyse uit de vorige paragraaf, waarbij geen correctie werd uitgevoerd voor nonresponse, tot vertekende resultaten heeft geleid. Alle directe effecten, waarvan we vermoedden dat ze te maken hadden met een niet geheel toevallige uitval, verdwijnen wanneer de subgroepen met ontbrekende gegevens bij de analyse betrokken worden. De veranderingen hebben vooral betrekking op het eerste deel van het causaal model, namelijk op de marginalen DEF en DEFX. De effecten op de opleiding en het beroep van de respondent veranderen vrijwel niet. Wanneer we echter ook de houdbaarheid van een aangepast model willen aantonen, dan zal zo'n model eerst moeten worden getoetst. Omdat dat in het geval van een "missing at random" nonresponse niet kan met behulp van de door Fuchs aangedragen toetsingprocedure, zal in de volgende subparagraaf de door Fay voorgestelde methode worden toegepast.

5.4. Correctie met behulp van response-variabelen.

In deze subparagraaf zal allereerst worden getracht om responsemodellen op te stellen voor de uitval in de tweede en derde golf. In eerste instantie zullen we ons daarbij beperken tot modellen waarin het response-mechanisme "ignorable" is. Later zullen voor de nonresponse bij de variabelen opleiding en beroep ook "nonignorable" response-modellen worden opgesteld. De bedoeling was om te beginnen met twee aparte response-modellen voor de nonresponse in de tweede (opleiding vader) en derde golf (opleiding en beroep respondent), om deze daarna gecombineren tot één verklaringsmodel voor beide vormen van uitval. De computerfaciliteiten die tot onze beschikking stonden lieten dit laatste jammer genoeg niet toe. Het was niet mogelijk om twee response-variabelen tegelijkertijd op te nemen in het causale model met de onderzoeksvariabelen. Daarom moeten we volstaan me de twee afzonderlijke response-modellen.

Behalve het bepalen van de response-mechanisme moet er natuurlijk ook een verklaringsmodel voor de relaties tussen de onderzoeksvariabelen worden gezocht. Als hypothetisch model wordt daarvoor steeds het model genomen, dat met behulp van de procedure van Fuchs niet getoetst kon worden, maar dat op grond van de omvang van de parameters en op inhoudelijke gronden erg plausibel lijkt. Bij de analyse met de subgroepen 1 en 2 is dat het model uit paragraaf 5.2 maar dan zonder de interacties DF, EF, DFX en FX. Voor de analyse met de subgroepen 1 en 3 is dat bij een "ignorable" response-mechanisme ook het model uit paragraaf 5.2 maar dan alleen zonder de interactie EF. Wanneer het response-mechanisme "nonignorable" is, is het op grond van de informatie die we tot nu toe hebben niet mogelijk om een hypothetisch model voor de relaties tussen de onderzoeksvariabelen op te stellen. Daarom zal er naar een model moeten worden gezocht dat bij de data past. Daarbij is het interessant om te zien of de relaties tussen de onderzoeksvariabelen in het "ignorable" en in het "nonignorable" response-model erg verschillen.

5.4.1. Response-modellen.

Om de door Fay ontwikkelde techniek te kunnen toepassen moet allereerst de nonresponse die zich in deze dataset voordoet vertaald worden in termen van response-variabelen. Het al dan niet responderen op de variabele opleiding vader (tweede golf) wordt de response-variabele R en het al dan niet responderen op opleiding en beroep (derde golf) S, waarbij de score 1 betekent wel responderen en de score 2 niet responderen.

De effecten op R werden onderzocht aan de hand van de data van de subgroepen ABCDEFGH en ABCDFGH. In bijlage VI.1 zijn de resultaten van de uitgevoerde toetsen te vinden. Het response-model dat het beste past bij de data is het model met de vier variabelen interactie DFRX (model 3). Er is echter de voorkeur gegeven aan het spaarzamere model 6b, waarin alleen de drie variabelen interacties DFR en FRX zijn opgegenomen. Daarbij heeft de conditionele toets tussen dit model en model 2 (p(x)=.50) de doorslag gegeven. Zoals we hadden verwacht is het al dan niet aanwezig zijn van de gegevens over de opleiding van de vader direct afhankelijk van het beroep van de vader en van de intelligentie en het geslacht van de respondent. Respondenten met een vader met een hoog of middelbaar beroep responderen bij constanthouding van de andere variabelen gemiddeld (λ=-.001), degenen met een vader die zelfstandige of arbeider is vaker (λ=203 en λ=.353) en de restgroep minder vaak (λ=-.556). Het directe effect van intelligentie op R is zeer groot (λ=-.719): niet intelligente kinderen zullen veel vaker niet reponderen dan intelligente kinderen. Het directe effect van geslacht op R is .161, hetgeen betekent dat voor mannen vaker de opleiding van de vader bekend is dan voor vrouwen. Het drie-variabelen-effect FRX heeft een omvang van .176. Dat houdt in dat voor mannen het directe effect van intelligentie op response kleiner is dan voor de vrouwen. De interactie DFR geeft aan dat het effect van beroep vader op response sterker is voor de mannen dan voor de vrouwen. In bijlage VII.1.2 zijn de λ-parameters, die de omvang van de directe effecten op R weergeven, te vinden.

De verklaring van response-variabele S met behulp van data in de subgroepen ABCDEFGH en ABCDEF leverde meer problemen op omdat model 2 (zie bijlage VI.2), het model waarin de latente variabele X wordt gedefiniëerd en waarin alle mogelijke relaties naar S zijn gelegd uitgaande van de "ignorable" nonresponse assumptie, maar een overschijdingkans van 1% heeft. Dat dit "bijna-verzadigde" model eigenlijk moet worden verworpen heeft twee mogelijke oorzaken. Het kan zijn dat X^2 benadering van de toetsingsgrootheid L niet opgaat vanwege het grote aantal cellen zonder of met een te gering aantal waarnemingen. Het kan er echter ook op wijzen dat bij de analyse met behulp van deze twee subgroèpen de relaties tussen de intelligentie-indicatoren A, B en C niet kunnen worden verklaard vanuit X zoals dat tot nu toe steeds het geval was. Dit is echter erg onwaarschijnlijk. Omdat elk getoetst model moest worden verworpen is er door middel van van conditionele toetsing naar een passend model gezocht. Van deze toetsen is immers bekent dat ze robuster zijn ten aanzien van de schending van de voorwaarden waaronder ze mogen toegepast.

In bijlage VI.2 zijn de toetsingsresulaten te vinden. Bij de bepaling van het beste verklaringsmodel voor S is zowel op de uitkomsten van de conditionele toetsen als op de spaarzaamheid van het model gelet. Op grond van deze twee criteria bleek model 4c het best passende model te zijn. S (het het dan niet deelnemen aan de derde golf) wordt daarin verklaard vanuit F (geslacht), X (intelligentie) en de drie-variabelen-interactie FSX. Het directe effect van F op S is .241, dit betekent dat mannen relatief vaker responderen dan vrouwen (voor de λ-parameters zie bijlage VII.2). Daarnaast is er een klein direct effect van X op S (λ=-.085): intelligente mensen responderen vaker dan niet intelligente. Het effect FSX (λ=.097) betekent dat bij mannen het effect van X op S kleiner is dan bij vrouwen. Bij de mannen is dat .012 (=-.085+ .097) en bij de vrouwen -.182 (=-.085-.097).

Het directe effect tussen geslacht en het al dan niet responderen op de variabelen opleiding en beroep was te verwachten op grond van de informatie die we hadden over deze uitval: bij mannen is meer moeite gedaan om de gewenste informatie te verkrijgen dan bij vrouwen. Er blijkt bij de vrouwen echter ook een direct verband te bestaan tussen intelligentie en de response-variabele S: intelligente vrouwen responderen vaker dan niet intelligente.

Hierboven hebben we ons beperk tot de "ignorable" response-modellen. Het is echter ook mogelijk om deze assumptie los te laten. Voor de uitval bij de variabelen opleiding en beroep (S) zou dat betekenen dat deze direct mag samenhangen met de opleiding en het beroep van de respondenten. Modellen 7a en 7b zijn "nonignorable" response-modellen (zie bijlage VI.2). In model 7a is behalve de interactie FSX ook de drie-variabelen-interactie GHS opgenomen, terwijl in model 7b wordt volstaan met de effecten GS en HS. Model 7b wordt verkozen omdat uit de toetsingsresultaten blijkt dat de interactie GHS niet significant is. Bovendien blijkt uit de conditionele toets dat dit model beter bij de data past dan het "ignorable" response-model 4c. Het is echter niet mogelijk om op grond van de data te bepalen of dit "nonignorable" response-model zinnig is. Dit type response-modellen heeft alleen maar zin als de onderzoeker vermoedt dat de uitval op een bepaalde onderzoeksvariabele direct samenhangt met de score op deze variabele.

In het "nonignorable" model is het directe effect FS iets kleiner (λ=.170) en zijn de effecten SX (λ=-.146) en FSX (λ=.119) iets groter dan in het "ignorable" response-model. Het directe effect van opleiding op S is groot (λ=.457). Mensen met een lage opleiding zullen vaker responderen dat mensen met een hoge opleiding. Het directe effect van beroep op de response-variabele S is precies andersom: λ= -.536. Bij onderzoekspersonen die tot de lage beroepscategorie behoren is het uitval veel groter dan bij degenen met een hoog beroep.

5.4.2. De relaties tussen de onderzoeksvariabelen.

Bij de analyse die is uitgevoerd met de frequentietabellen van de subgroepen ABCDEFGH en ABCDFGH blijkt, dat de relaties tussen de onderzoeksvariabelen waarvan werd verwacht dat ze niet significant zouden zijn, inderdaad kunnen worden weggelaten. Model 2 in bijlage VII.1.1 is het model dat werd gevonden in paragraaf 5.2 (zie schema 5.1). Dat model kan ook nu worden aanvaard. Model 3 is bijna hetzelfde model, alleen zijn hier de interacties DF, EF, DFX en FX weggelaten. Conditionele toetsing wijst uit dat deze interacties niet significant zijn (L=14.0, df=8, p(x)=.1), hetgeen overeenkomt met onze verwachtingen. De andere directe effecten veranderen weinig ten opzichte van de eerdere gevonden resultaten, alleen het effect van beroep vader op intelligentie is anders (zie bijlage VII.1.2). Eerder vonden we dat kinderen met een vader met een hoger of middelbaar beroep hoger scoorden op de intelligentievariabele. Dat is nu ook het geval (λ=-.436). Bij kinderen van boeren of zelfstandigen en bij arbeiderskinderen is het negatieve effect van beroep vader op intelligentie echter bijna geheel verdwenen (λ=.029 en λ=.100), terwijl nu juist de restgroep relatief slechter scoort op de intelligentievariabele (λ=.307).

Uit de analyse met behulp van de subgroepen 1 (ABCDEFGH) en 3 (ABCDEF) blijkt dat de interactie EF niet significant is. In bijlage VII.2.1 zijn de toetsingresultaten te vinden voor het model dat we in paragraaf 5.2 vonden (model 2) en het model zonder de interactie EF (model 3). Beide modellen kunnen worden aanvaard (na conditionele toetsing t.o.v. model 1), maar de conditionele toets (L=1.7, df=1, p(x)=.2) geeft aan dat model 3 te verkiezen valt boven model 2. De andere effecten zijn vrijwel gelijk aan de effecten die we vonden bij de analyse uit paragraaf 5.2. Alleen de drie-variabelen-interactie DFX is iets kleiner.

In het geval van het "nonignorable" response-mechanisme zijn de directe effecten tussen D,E,F en X vrijwel identiek aan het "ignorable" model (zie bijlage VII.3.2). Bij de verklaring van het beroep van de respondent (H) blijkt het nu echter nodig te zijn om de interactie FHX (λ=.138) op te nemen in het model. Uit de conditionele toets tussen model 2 en 3 in bijlage VII.3.1 blijkt dat dit effect significant is ($L=13.1$, $df=1$, $p(x)<.001$). Dat betekent dat het effect van intellingentie op beroep bij mannen een stuk sterker wordt ($\lambda=.183+.138=.321$) en dat bij de vrouwen dat effect bijna geheel verdwijnt ($\lambda=.183-.138=.045$). Bovendien is het directe effect van opleiding op beroep een stuk kleiner dan tot nu toe steeds het geval was ($\lambda=.362$).

5.5. Conclusies.

Met behulp van Fay's techniek hebben we een duidelijk beeld gekregen van de nonresponse in deze dataset. De uitval in de tweede golf is zoals we vermoedden direct afhankelijk van de intelligentie van de respondent en van het beroep van de vader. Daarnaast is er een verschil tussen mannen en vrouwen wat betreft de nonresponse bij de variabele opleiding vader. De uitval in de derde golf is te verklaren vanuit het geslacht van de respondenten en bij de vrouwen is er ook een effect van intelligentie op responsekans. In het "nonignorable" response-model is er bovendien een effect van de variabelen beroep en opleiding op de nonresponse bij de deze variabelen.

Deze correctiemethode voor nonresponse heeft een grote invloed op de relaties tussen de onderzoeksvariabelen. Ondanks het feit dat het niet mogelijk was om de analyse uit te voeren met alle vier de subgroep tegelijk, kunnen we met op grond van de omvang van de parameterschattingen (die kennen we immers wel) en op grond van de aparte analyses voor de twee response-modellen, met een vrij grote zekerheid stellen dat de interacties DF, EF, FX en DFX verdwijnen en dat het effect van het beroep van de vader op intelligentie verandert. Bovendien treden er nog enkele andere veranderingen op wanneer een "nonignorable" response-mechanisme wordt verondersteld.

6. CONCLUSIES.

Het doel van dit onderzoek was om een oplossing te vinden voor het probleem, dat bij de correctie voor nonresponse door het schatten van de "sufficient statistics", de toetsingsprocedure uitgaat van de "missing completely at random" assumptie. Door de nonresponse te modelleren met behulp van response-variabelen, die in het causale model met de onderzoeksvariabelen worden opgenomen, blijkt het ook in gevallen waarin de nonresponse niet geheel toevallig is mogelijk om het opgesteld causaal model te toetsen. Het is nu zelfs mogelijk om zowel bij de schattings- als toetsingsprocedure uit te gaan van een "nonignorable" response-mechanisme.

De response-modellen hebben behalve dat ze een oplossing bieden voor het hierboven aangehaalde toetsingsprobleem ook een eigen waarde. Ze vormen een belangrijk instrument bij het achterhalen van de oorzaken van de uitval bij bepaalde onderzoeksvariabelen.

Wij hebben hier bovendien iets nieuws toegevoegd aan Fay's techniek, namelijk de uitbreiding naar modellen met latente variabelen. In paragraaf 4 werden de formules die daarbij van belang zijn gegeven en in paragraaf 5 is een toepassing van deze correctiemethode voor nonresponse in loglineaire modellen met latente variabelen besproken.

Twee zaken zouden mijns inziens nader moeten worden onderzocht aan de hand van andere bestaande datasets: ten eerste de mogelijkheden die "nonignorable" response-modellen bieden bij de verklaring van nonresponse en ten tweede de toepassingsmogelijkheden van de hier beschreven techniek in modellen met latente variabelen.

GEBRUIKTE LITERATUUR.

Chen T. (1979),
Log-linear models for categorized data with misclassification and double sampling, Journal of the American Statistical Association, jrg. 74, pag. 481-488.

Chen T. en S.E. Fienberh (1974),
Two-dimensional contingency tables with both completely and partially cross-classified data, Biometrics, jrg. 30, pag. 629-642.

Dronkers J. (1987),
Ouders, liefde en geld: De relaties tussen ouderlijk milieu, schoolloopbaan, beroep, huwelijk en gezinsinkomen bij vrouwen, Katholieke Universiteit Brabant Subfaculteit SCW Working paper Series nr. 14.

Dronkers J. en B. Bakker (1986),
Milieu, onderwijs, huwelijk, beroep en gezinsinkomen: een replicatie en uitbreiding, SISWO, Amsterdam.

Dronkers J. en B.F.M. Bakker (1986),
Van een dubbeltje en het kwartje: de relatie tussen milieu, onderwijs, partnerkeuze, beroep en gezinsinkomen, Economisch Statistische Berichten, jrg. 71, nr 3584.

Fay R.E. (1986),
Causal models for patterns of nonresponse, Journal of the American Statistical Association, jrg. 81, pag. 354-365.

Fienberg S.E. (1977),
The analysis of cross-classified categorical data, MIT Press, Cambridge.

Fuchs C. (1982),
Maximum likelihood estimation and model selection in contingency tables with missing data, Journal of the American Statistical Association, jrg. 77, pag. 270-278.

Jöreskog K.G. en D. Sörbom (1984),
Lisrel VI: analysis of linear structural relationschips by maximum likelihood, instrumental variables and least squares methodes, University of Uppsala, Sweden.

Hagenaars J.A.P. (1985),
Loglineaire analyse van herhaalde surveys; panel- trend- en cohortonderzoek, Dissertatie, Tilburg.

Hagenaars J.A.P. (1986),
Ontbrekende gegevens bij loglineaire analyses met en zonder latente variabelen, in: J.H.G. Segers en E.J. Bijnen (red.), Onderzoeken: reflecteren en meten, pag. 57-77, Tilburg University Press, Tilburg.

Hagenaars J.A.P. en A.G.J.J. Heinen (1980),
Analyse van contingentietabellen met behulp van het loglineaire model, in: Sociologische onderzoeksmethoden deel II: technieken van causale analyse, hfdst.6, pag 185-258, Van Gorcum, Assen.

Hagenaars J.A.P en R. Luijkx (1987),
LCAG: latent class models and other models with latent variables, Tilburg University Department of Sociology Working Paper Series # 17, Tilburg.

Hanushek E.A. en J.E. Jackson (1977),
Statistical methods for social scientists, Academic Press, New York/San Francisco/Londen.

Hays W.L. (1981),
Statistics, CBS College Publishing, New York.

Hartog J. (1986),
Survey non-response in relation to ability and family background: structure and effects on estimated earning functions, Research Memorandum no. 8620, Universiteit van Amsterdam.

Hocking R.R. en H.H. Oxsprings (1971),
Maximum likelihood estimation with incomplete multinominal data, Journal of the American Statistical Association, jrg. 66, pag. 65-70.

Hocking R.R. en H.H. Oxsprings (1974),
The analysis of partially categorized contingenty data, Biometrics, jrg. 30, pag. 469-483.

Korteweg P.J. en J. van Weesep (red.) (1983),
Ruimtelijk onderzoek: leidraad voor de opzet, uitvoering en verwerking, Romen, Bussum.

Little R.J.H. (1982),
Models for nonresponse in sample surveys, Journal of the American Statistical Association, jrg. 77, pag. 237-250.

Little R.J. en D.B. Rubin (1987)
Statistical analysis with missing data, Wiley, New York.

Marini M.M., A.R. Olsen en D.B. Rubin (1979),
Maximum-likelihood estimation in panel studies with missing data, in: K. Schuessler (red.), Sociological Methodology 1980, pag. 314-357, Jossey Bass, San Francisco.

Muthen B., D. Kaplan en Hollis M. (1987),
On structural equation modeling with dat that are not missing completely at random, Psychometrica, jrg. 52, pag. 431-462.

Rubin D.B.(1984)
Characterizing the estimation of parameters in incomplete-data problems, Journal of the American Statistical Association, jrg. 69, pag. 467-474.

Scott Long J. (1983),
Confirmatory factor analysis, Sage Publications, Beverly Hills.

Scott Long J. (1983),
Covariance structure models: an introduction to LISREL, Sage Publications, Beverly Hills.

Segers J.H.G. (1977),
Sociologische onderzoeksmethoden: inleiding tot de structuur van het onderzoeksproces en tot de methoden van dataverzameling, Van Gorcum, Assen/Amsterdam.

Segers J.H.G. en J.A.P. Hagenaars (red.) (1980)
Sociologische onderzoeksmethoden deel II: technieken van causale analyse, Van Gorcum, Assen.

BIJLAGE I: Codering van de in paragraaf 5 gebruikte variabelen uit de uitgebreide Matthijssen-Sonnemans cohort.

<u>Variabele A</u>: Intelligentiecode LO-IV = een intelligentietest bestaande uit drie verbale en drie nonverbale test.

<u>Variabele B</u>: Intelligentiecode P.M.

<u>Variabele C</u>: Intelligentiecode W.S = een woordenschattest.

Variabelen A t/m C zijn variabelen van intervalniveau. De schaal loopt van 67 t/m 148.

Hercodering van A, B en C: 1] = 67 t/m 100
2] = 101 t/m 148

<u>Variabele D</u>: Beroep vader 1952.

BEROEP	CODE	%
Onderwijskrachten	11	0.4
Leidinggevenden	12	3.1
Onderwijzers	21	1.3
Middelbaar personeel	22	8.0
Zelfstandige agrariërs	23	17.0
Zelfstandige ambachtslieden	24	14.9
Lager administratief personeel	31	2.1
Landarbeiders	32	2.3
Arbeiders in de industrie	33	35.4
Overige werknemers	34	5.4
Werkend s.g. onbekend	35	0.4
Tijdelijk niet werkend	36	0.1
Pensioen, rentenier	37	0.3
Invalide	40	0.2
Voornaamste bezigheid onbekend	42	9.0
Niet van toepassing	43	0.2

Hercodering: 1] = 11, 12, 21 en 22
2] = 23 en 24
3] = 31, 32, 33 en 34
4] = 35, 36, 37, 40, 42 en 43

Variabele E: Genoten dagonderwijs vader.

ONDERWIJSNIVEAU	CODE	%
Voorafgaand aan het eerste le	1	0.0
le niveau	2	41.4
2e niveau, le trap	3	9.0
2e niveau, 2e trap	4	3.7
3e niveau, le trap	5	1.7
3e niveau, 2e trap	6	0.9
Niet ingevuld	0	4.0
Onbekend	9	39.4

Hercodering: 1] = 1 en 2
2] = 3 t/m 6

Variabele F: Geslacht

Codering: 1] = Jongen
2] = Meisje

Variabele G: Hoogst genoten opleiding

OPLEIDINGSNIVEAU	CODE
Lager onderwijs	1
VGLO, LAVO, MULO, MAVO	2
Lager beroepsonderwijs	3
Middelbaar beroepsonderwijs	4
HBS, MMS, HAVO, VWO	5
Hoger beroepsonderwijs	6
Wetenschappelijk onderwijs	7

Hercodering: 1] = 1 t/m 3
2] = 4 t/m 7

Variabele H: Beroepsniveau

BEROEPSNIVEAU	CODE
Zeer eenvoudige arbeid	1
Eenvoudige arbeid	2
Enigszins ingewikkelde arbeid	3
Vrij ingewikkelde arbeid	4
Ingewikkelde arbeid	5
Zeer ingewikkelde arbeid	6
Wetenschappelijke arbeid	7

Hercodering: 1] = 1 t/m 4
2] = 5 t/m 7

BIJLAGE II: Frequentietabellen voor de diverse subgroepen.

TABEL II.1: Subgroep ABCDEFGH

				D	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	
				E	1	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2	
				H	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	
G	F	A	B	C																	
1	1	1	1	1	0	33	60	1	0	3	2	0	0	25	28	1	0	3	0	0	156
2	1	1	1	1	0	2	1	0	0	1	0	0	2	8	3	1	0	1	0	0	19
1	2	1	1	1	0	3	13	2	0	0	0	0	0	0	1	0	0	1	0	0	20
2	2	1	1	1	0	1	0	0	0	0	0	0	0	2	0	0	0	0	0	0	3
1	1	2	1	1	1	5	13	0	0	1	0	0	1	6	4	0	0	0	0	0	31
2	1	2	1	1	0	0	0	0	0	0	0	0	0	3	4	0	1	0	0	0	8
1	2	2	1	1	0	5	4	1	0	0	1	0	0	1	0	0	0	0	0	0	12
2	2	2	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	1	2	1	0	6	23	1	0	0	4	0	0	12	7	0	0	1	1	0	55
2	1	1	2	1	0	1	0	0	0	0	0	0	0	4	3	0	0	1	3	0	12
1	2	1	2	1	1	3	7	1	0	0	1	0	0	2	0	0	0	0	0	0	15
2	2	1	2	1	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	3
1	1	2	2	1	0	11	14	3	0	1	2	0	1	19	15	0	2	1	1	0	70
2	1	2	2	1	0	1	2	0	1	2	0	1	2	11	3	0	2	2	0	0	27
1	2	2	2	1	0	1	6	0	1	1	1	0	1	1	2	0	0	0	1	1	16
2	2	2	2	1	1	1	1	0	1	1	1	0	0	6	1	0	1	0	0	0	14
1	1	1	1	2	0	4	26	0	0	1	0	0	0	4	12	0	1	0	2	0	50
2	1	1	1	2	1	1	3	0	0	0	0	0	1	3	3	0	2	1	1	0	16
1	2	1	1	2	3	2	9	0	1	0	0	1	0	1	1	0	0	0	0	0	18
2	2	1	1	2	0	0	0	0	1	0	0	0	0	1	1	1	0	0	1	0	5
1	1	2	1	2	0	5	13	1	0	1	1	0	2	9	17	1	0	3	1	0	54
2	1	2	1	2	0	3	2	1	0	0	0	0	0	9	3	1	11	4	7	0	41
1	2	2	1	2	1	9	14	0	1	2	2	0	0	1	2	0	0	1	1	0	34
2	2	2	1	2	1	0	2	0	0	0	1	1	2	3	3	0	4	1	2	1	21
1	1	1	2	2	1	4	18	0	1	0	2	0	1	3	7	2	0	0	0	0	39
2	1	1	2	2	1	1	1	0	1	0	0	0	0	1	8	0	5	0	1	0	19
1	2	1	2	2	0	3	4	1	1	1	0	1	0	1	2	0	0	0	0	0	14
2	2	1	2	2	0	0	0	0	0	0	0	0	0	2	1	0	1	0	0	0	4
1	1	2	2	2	3	19	33	3	1	2	9	0	4	13	28	3	7	3	4	1	133
2	1	2	2	2	2	2	2	1	2	1	2	1	11	32	21	3	41	17	12	2	152
1	2	2	2	2	3	16	46	3	7	7	14	1	1	2	5	0	1	1	2	1	110
2	2	2	2	2	1	4	5	1	4	3	1	0	4	20	20	2	28	6	3	5	107
					20	147	323	20	23	28	44	6	33	206	205	15	107	48	43	11	1279

TABEL II.2: Subgroep ABCDFGH

F A B C																	
D	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	
H	1	1	1	1	2	2	2	2	1	1	1	1	1	1	1	1	
G	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	
1 1 1 1	2	17	21	25	1	9	8	4	1	0	0	1	2	4	1	1	97
2 1 1 1	6	23	45	5	2	5	4	0	0	1	2	0	1	3	1	0	98
1 2 1 1	0	0	1	3	0	1	1	0	0	0	0	0	0	0	0	1	7
2 2 1 1	0	2	6	1	0	1	2	0	0	0	0	0	0	1	0	0	13
1 1 2 1	0	5	15	2	2	3	3	3	0	1	0	1	2	5	2	0	44
2 1 2 1	1	13	21	8	0	5	1	0	0	1	1	0	0	3	0	0	54
1 2 2 1	0	2	4	0	0	2	3	1	0	1	0	0	0	1	1	4	19
2 2 2 1	1	5	4	1	0	0	1	0	0	1	0	0	0	2	1	0	16
1 1 1 2	2	3	10	2	1	2	7	2	0	0	0	0	0	0	0	0	29
2 1 1 2	1	3	9	4	2	0	0	0	0	0	1	1	0	1	0	2	24
1 2 1 2	0	1	1	1	0	2	0	1	0	0	0	0	2	0	1	1	10
2 2 1 2	1	2	1	1	1	0	0	0	0	1	1	0	1	1	0	0	10
1 1 2 2	1	2	6	2	0	2	1	1	0	0	0	0	1	1	0	1	18
2 1 2 2	0	2	9	0	0	1	2	0	0	1	0	1	0	0	1	0	17
1 2 2 2	1	3	4	0	1	2	5	1	1	0	0	0	5	3	1	1	28
2 2 2 2	3	4	2	2	0	1	1	0	0	1	1	0	1	3	0	3	22
	19	87	159	57	10	36	39	13	2	8	6	4	15	28	9	14	506

TABEL II.3: Subgroep ABCDEF

A B C																	
D	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	
E	1	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2	
F	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	
1 1 1	1	47	65	0	1	1	7	0	0	10	35	1	3	2	3	0	176
2 1 1	1	10	11	0	1	0	1	0	0	8	7	0	1	1	0	0	41
1 2 1	2	19	27	0	1	1	4	1	3	14	15	2	0	3	3	0	95
2 2 1	4	7	27	1	3	3	1	0	5	16	28	1	7	6	2	2	113
1 1 2	1	17	32	7	7	6	6	2	5	18	44	6	3	3	3	0	160
2 1 2	3	12	31	0	12	6	5	2	4	25	31	7	8	4	6	0	156
1 2 2	2	6	17	3	2	5	3	1	2	13	16	4	7	4	2	1	88
2 2 2	21	47	84	4	42	18	21	7	23	56	80	10	57	23	16	6	515
	35	165	294	15	69	40	48	13	42	160	256	31	86	46	35	9	1344

TABEL II.4: Subgroep ABCDF

	D	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	
	F	1	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2	
	A	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	
B	C																	
1	1	11	53	57	49	12	164	253	49	3	5	2	4	2	14	13	2	693
2	1	0	16	20	13	7	71	84	26	1	7	8	4	6	18	19	10	310
1	2	5	7	20	15	10	29	49	18	5	1	4	5	1	7	8	2	186
2	2	2	1	9	3	3	11	22	12	9	13	8	16	10	12	17	21	169
		18	77	106	80	32	275	408	105	18	26	22	29	19	51	57	35	1358

BIJLAGE III: Relatieve frequenties per variabele in de diverse subgroepen.

Variabele	Categorie	Subgroepen ABCDEFGH	ABCDFGH	ABCDEF	ABCDF
A	1	.350	.753	.386	.811
	2	.650	.247	.614	.189
B	1	.382	.569	.397	.647
	2	.618	.431	.603	.353
C	1	.361	.688	.316	.739
	2	.639	.312	.684	.261
D	1	.143	.091	.173	.064
	2	.335	.314	.306	.316
	3	.481	.421	.471	.437
	4	.041	.174	.051	.183
E	1	.758	---	.743	---
	2	.242	---	.257	---
F	1	.690	.498	.505	.277
	2	.310	.502	.495	.723
G	1	.647	.830	---	---
	2	.353	.170	---	---
H	1	.478	.676	---	---
	2	.522	.324	---	---

BIJLAGE IV: de toetsingsresultaten van de modellen op grond van de "volledige" tabel (tabel 1).

IV.1 Modellen en toetsingsresultaten

MODEL	L	df	p(x)	X^2	p(x)
1] {AX,BX,CX}{DEFGHX}	363.6	378	.694	370.3	.602
1a]{AX,BX,CX}{DE,EF,DF}{DEF,DFX,EX} {DEFX,DG,EG,XG}{DEFGX,DH,FH,HX,GH}	451.9	471	.729	536.8	.019
niet opnemen					
2a]{AX,BX,CX}{DE}{DEFGHX}	387.0	385	.461	418.9	.113
2b]{AX,BX,CX}{DE,EF,DF}{DEFGHX}	367.4	381	.683	380.1	.503
2c]{AX,BX,CX}{DE,DF}{DEFGHX}	374.2	382	.602	391.4	.358
3a]{AX,BX,CX}{DEF,DX,EX,FX}{DEFGHX}	380.8	388	.593	418.7	.136
3b]{AX,BX,CX}{DEF,DEX,FX}{DEFGHX}	378.0	385	.591	413.1	.155
3c]{AX,BX,CX}{DEF,DEX,DFX}{DEFGHX}	364.3	382	.734	370.7	.651
3d]{AX,BX,CX}{DEF,DFX,EX}{DEFGHX}	366.6	385	.742	370.9	.688
3e]{AX,BX,CX}{DEF,DFX}{DEFGHX}	414.1	386	.155	415.3	.146
4a]{AX,BX,CX}{DEFX,DG,EG,GX}{DEFGH}	388.5	404	.702	404.1	.490
4b]{AX,BX,CX}{DEFX,DG,EG,FG,GX}{DEFGH}	386.6	403	.713	404.1	.475
4c]{AX,BX,CX}{DEFX,DEG,FG,GX}{DEFGH}	384.4	400	.704	407.2	.391
4d]{AX,BX,CX}{DEFX,DGX,EG,FG}{DEFGH}	386.2	400	.681	402.2	.459
4e]{AX,BX,CX}{DEFX,DFG,EG,GX}{DEFGH}	385.5	400	.689	401.6	.468
4f]{AX,BX,CX}{DEFX,DG,GX}{DEFGH}	407.2	405	.459	446.1	.077
4g]{AX,BX,CX}{DEFX,DG,EG}{DEFGH}	451.0	405	.057	461.5	.027
4h]{AX,BX,CX}{DEFX,EG,FG}{DEFGH}	456.6	405	.045	491.6	.003
5a]{AX,BX,CX}{DEFGX,DH,FH,HX,GH}	417.7	435	.716	483.8	.053
5b]{AX,BX,CX}{DEFGX,DH,EH,FH,HX,GH}	417.5	434	.707	482.5	.053
5c]{AX,BX,CX}{DEFGX,DGH,FH,HX}	416.6	432	.694	476.4	.069
5d]{AX,BX,CX}{DEFGX,DFH,HX,GH}	415.9	432	.702	474.5	.077
5e]{AX,BX,CX}{DEFGX,DHX,FH,GH}	411.4	432	.755	465.6	.128
5f]{AX,BX,CX}{DEFGX,DH,FH,GH}	425.5	436	.631	497.1	.023
5g]{AX,BX,CX}{DEFGX,FH,HX,GH}	435.1	438	.530	509.7	.010
5h]{AX,BX,CX}{DEFGX,DH,HX,GH}	492.9	436	.031	552.6	.000
5i]{AX,BX,CX}{DEFGX,DH,FH,HX}	605.9	436	.000	605.9	.000
5j]{AX,BX,CX}{DEFGX,DFH,DHX,GH}	418.5	429	.755	453.5	.200

IV.2 Conditionele toetsen

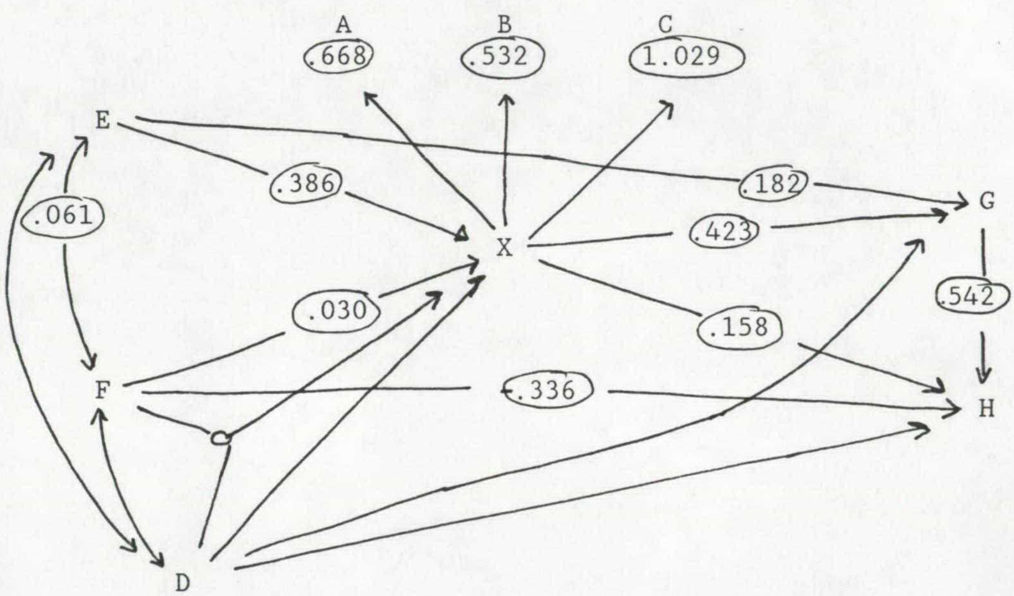
MODELLEN	$L_{b/u}$	$df_{b/u}$	p(x)
1a-1	88.3	93	.6
2a-1	23.4	7	.002
2b-1	3.8	3	.25
2c-2b	6.8	1	.01
3a-1	17.2	10	.1
3a-3b	2.8	3	.5
3b-3c	13.7	3	.01
3a-3d	14.2	3	.005
3d-1	3.0	7	.9
4a-1	24.9	26	.5
5a-1	54.1	57	.3
5a-5b	0.2	1	.6
5a-5c	1.1	3	.75
5a-5d	1.8	3	.6
5a-5e	6.3	3	.1
5e-1	47.8	54	.75

BIJLAGE V: Toetsingresultaten en parameterschattingen van de analyses met de diverse subgroepen (methode Fuchs).

V.1.1. Toetsingsresultaten voor subgroepen ABCDEFGH en ABCDFGH.

MODEL	L	df	p(x)	X^2	p(x)
1] {AX,BX,CX}{DEFGHX}	1114.2	633	.000	1202.2	.000
1a]{AX,BX,CX}{DE,EF,DF}{DEF,DFX,EX} {DEFX,DG,EG,XG}{DEFGX,DH,FH,HX,GH}	1211.2	726	.000	1378.8	.000

V.1.2. Parameterschattingen voor model 1a.



$\lambda^{DE}_{11}=-.863$ $\lambda^{DE}_{21}=.324$ $\lambda^{DE}_{31}=.466$ $\lambda^{DE}_{41}=.072$

$\lambda^{DF}_{11}=-.015$ $\lambda^{DF}_{21}=.032$ $\lambda^{DF}_{31}=.002$ $\lambda^{DF}_{41}=-.019$

$\lambda^{DFX}_{111}=-.156$ $\lambda^{DFX}_{211}=.042$ $\lambda^{DFX}_{311}=.077$ $\lambda^{DFX}_{411}=.036$

$\lambda^{DX}_{11}=-.416$ $\lambda^{DX}_{21}=.015$ $\lambda^{DX}_{31}=.083$ $\lambda^{DX}_{41}=.318$

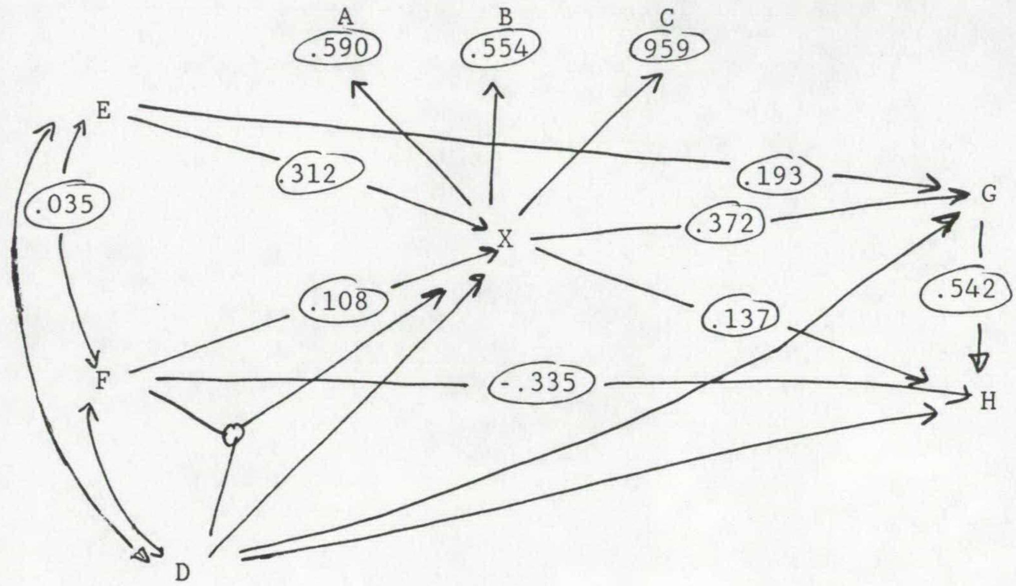
$\lambda^{DG}_{11}=-.362$ $\lambda^{DG}_{21}=-.072$ $\lambda^{DG}_{31}=.403$ $\lambda^{DG}_{41}=.030$

$\lambda^{DH}_{11}=-.213$ $\lambda^{DH}_{21}=-.127$ $\lambda^{DH}_{31}=.144$ $\lambda^{DH}_{41}=.196$

V.2.1. Toetsingsresultaten voor subgroepen ABCDEFGH en ABCDEF.

MODEL	L	df	p(x)	X^2	p(x)
1] {AX,BX,CX}{DEFGHX}	701.5	505	.000	710.2	.000
1a]{AX,BX,CX}{DE,EF,DF}{DEF,DFX,EX} {DEFX,DG,EG,XG}{DEFGX,DH,FH,HX,GH}	786.3	598	.000	847.4	.000

V.2.2. Parameterschattingen voor model 1a.



$\lambda^{DE}_{11}=-.794$ $\lambda^{DE}_{21}=.309$ $\lambda^{DE}_{31}=.516$ $\lambda^{DE}_{41}=-.031$

$\lambda^{DF}_{11}=-.047$ $\lambda^{DF}_{21}=.108$ $\lambda^{DF}_{31}=.119$ $\lambda^{DF}_{41}=-.180$

$\lambda^{DFX}_{111}=-.181$ $\lambda^{DFX}_{211}=.158$ $\lambda^{DFX}_{311}=.101$ $\lambda^{DFX}_{411}=-.078$

$\lambda^{DX}_{11}=-.551$ $\lambda^{DX}_{21}=.206$ $\lambda^{DX}_{31}=.303$ $\lambda^{DX}_{41}=.042$

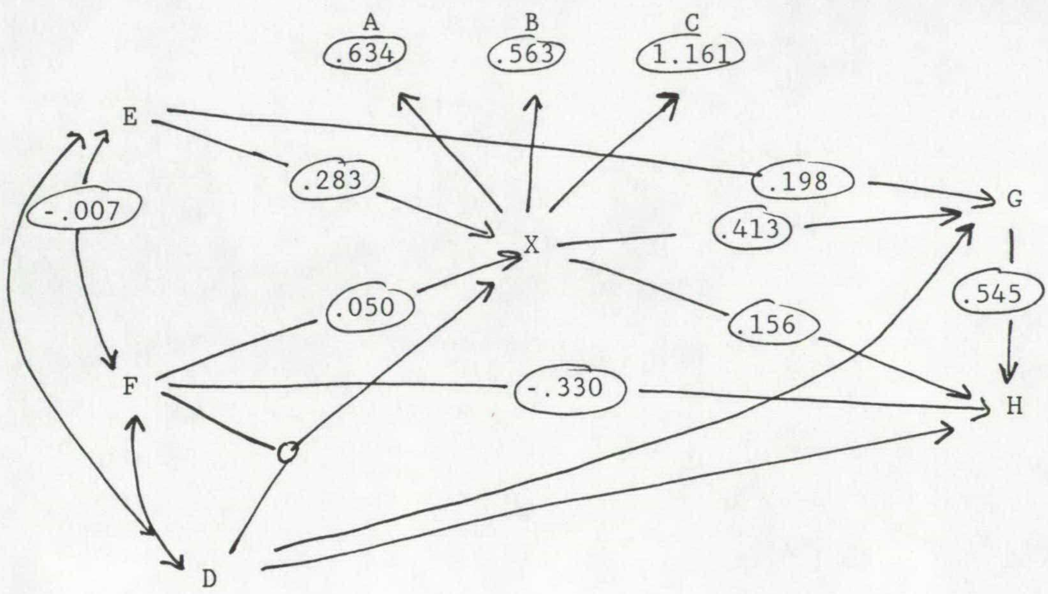
$\lambda^{DG}_{11}=-.399$ $\lambda^{DG}_{21}=-.052$ $\lambda^{DG}_{31}=.393$ $\lambda^{DG}_{41}=.058$

$\lambda^{DH}_{11}=-.184$ $\lambda^{DH}_{21}=-.103$ $\lambda^{DH}_{31}=.162$ $\lambda^{DH}_{41}=.125$

V.3.1. Toetsingsresultaten voor subgroepen ABCDEFGH, ABCDFGH, ABCDEF en ABCDF.

MODEL	L	df	p(x)	X^2	p(x)
1] {AX,BX,CX}{DEFGHX}	2858.2	823	.000	2777.0	.000
1a]{AX,BX,CX}{DE,EF,DF}{DEF,DFX,EX} {DEFX,DG,EG,XG}{DEFGX,DH,FH,HX,GH}	2950.8	916	.000	2945.9	.000

V.3.2. Parameterschattingen voor model 1a.



$\lambda^{DE}_{11}=-.830$	$\lambda^{DE}_{21}= .300$	$\lambda^{DE}_{31}= .505$	$\lambda^{DE}_{41}= .025$
$\lambda^{DF}_{11}= .023$	$\lambda^{DF}_{21}=-.014$	$\lambda^{DF}_{31}=-.008$	$\lambda^{DF}_{41}= .001$
$\lambda^{DFX}_{111}=-.045$	$\lambda^{DFX}_{211}=-.022$	$\lambda^{DFX}_{311}=-.046$	$\lambda^{DFX}_{411}= .113$
$\lambda^{DX}_{11}=-.543$	$\lambda^{DX}_{21}= .095$	$\lambda^{DX}_{31}= .138$	$\lambda^{DX}_{41}= .380$
$\lambda^{DG}_{11}=-.362$	$\lambda^{DG}_{21}=-.074$	$\lambda^{DG}_{31}= .398$	$\lambda^{DG}_{41}= .038$
$\lambda^{DH}_{11}=-.214$	$\lambda^{DH}_{21}=-.126$	$\lambda^{DH}_{31}= .144$	$\lambda^{DH}_{41}= .196$

BIJLAGE VI: Responsemodellen.

VI.1 Responsemodellen voor de subgroepen ABCDEFGH EN ABCDFGH.

Toetsingsresultaten.

MODEL	L	df	p(x)	X^2	p(x)
1]{AX,BX,CX,DEFGHX,R}	1114.3	633	.000	1202.3	.000
2]{AX,BX,CX,DEFGHX,DFGHRX}	556.2	570	.65	549.9	.724
3]{AX,BX,CX,DEFGHX,DFRX}	597.2	618	.72	623.1	.447
4]{AX,BX,CX,DEFGHX,DFR,DRX,FRX}	607.0	621	.65	634.6	.35
5]{AX,BX,CX,DEFGHX,DR,FR,RX}	651.3	628	.355	709.6	.001
6a]{AX,BX,CX,DEFGHX,DFR,DRX}	617.3	622	.548	646.7	.242
6b]{AX,BX,CX,DEFGHX,DFR,FRX}	**612.4**	**624**	**.629**	**628.2**	**.453**
6c]{AX,BX,CX,DEFGHX,DRX,FRX}	615.4	624	.595	686.8	.038
6d]{AX,BX,CX,DEFGHX,FRX,DR}	625,5	627	.52		

Conditionele toetsen.

MODEL	L	df	p(x)
3-2	48.0	41	.75
4-3	9.8	3	.03
5-4	44.3	7	<.001
6a-4	10.7	1	.001
6b-4	5.4	3	.25
6c-4	8.4	3	.025<p(x)<.05
6d-4	18.5	6	.005
6d-2	69.3	57	.035
6b-2	56.2	54	.5

VI.2. Responsemodellen voor de subgroepen ABCDEFGH EN ABCDEF.

Toetsingsresultaten.

MODEL	L	df	p(x)	X^2	p(x)
1]{AX,BX,CX,DEFGHX,S}	701.5	505	.000	710.1	.000
2]{AX,BX,CX,DEFGHX,DEFSX}	545.9	474	.010	539.3	.017
3a]{AX,BX,CX,DEFGHX,DEFS,DESX,DFSX,EFSX}	548.0	477	.010	542.6	.017
3b]{AX,BX,CX,DEFGHX,DESX,FX}	583.0	489	.001	609.3	.000
3c]{AX,BX,CX,DEFGHX,DFSX}	574.0	490	.004	561.8	
4a]{AX,BX,CX,DEFGHX,DES,DFS,EFS,DSX,ESX,FSX}	573.0	487	.003	592.3	.000
4b]{AX,BX,CX,DEFGHX,FSX,DS}	586.3	499	.003	590.2	.002
4c]{AX,BX,CX,DEFGHX,FSX}	**591.9**	**502**	**.003**	**595.6**	**.001**
5a]{AX,BX,CX,DEFGHX,DS,ES,FS,SX}	597.0	499	.001	592.8	.001
5b]{AX,BX,CX,DEFGHX,FS,SX}	604.4	503	.000		
7a]{AX,BX,CX,DEFGHX,FSX,GHS}	565.0	499	.02	581.6	.005
7b]{AX,BX,CX,DEFGHX,FSX,GS,HS}	**565.8**	**500**	**.02**	**581.9**	**.005**

Conditionele toetsen.

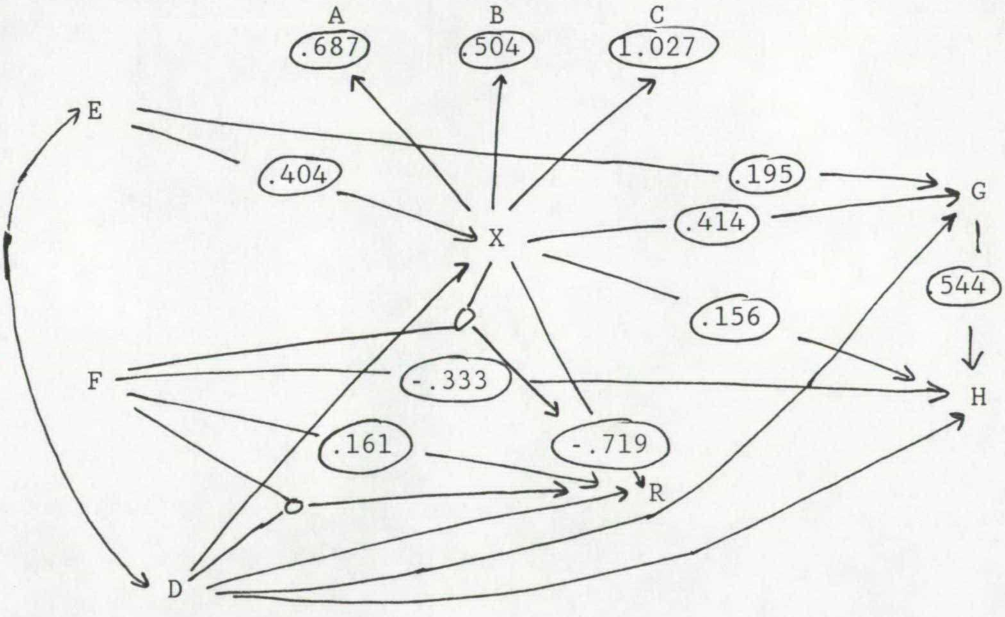
MODEL	L	df	p(x)
3a-2	2.1	3	.50
4a-3a	25.0	10	.005
5a-4a	24.0	12	.02
3b-3a	35.0	12	<.001
3c-3a	26.0	13	.015
3c-2	28.1	16	.03
4b-3c	12.3	9	.20
4c-4b	5.6	3	.15
4c-3c	17.9	12	.12
4c-2	46.0	28	.02
5b-4c	7.5	1	.005
7b-7a	.8	1	.3
4c-7b	20.5	3	.000

BIJLAGE VII: Toetsingresultaten en parameterschattingen van de analyses met de diverse subgroepen (methode Fay).

VII.1.1. Toetsingsresultaten voor subgroepen ABCDEFGH en ABCDFGH.

MODEL	L	df	p(x)	X^2	p(x)
1]{AX,BX,CX,DEFGHX,DFR,FRX}	612.4	624	.629	628.2	.453
2]{AX,BX,CX}{DE,EF,DF}{DEF,DFX,EX} {DEFX,DFR,FRX}{DEFRX,DG,EG,XG} {DEFGRX,DH,FH,HX,GH}	707.2	717	.60	794.6	.02
3]{AX,BX,CX}{DE,F}{DEF,DX,EX} {DEFX,DRX,FRX}{DEFRX,DG,EG,XG} {DEFGRX,DH,FH,HX,GH}	721.2	725	.504	816.1	.008

VII.1.2. Parameterschattingen voor model 3.



$\lambda^{DE}_{11}=-.871$ $\lambda^{DE}_{21}=.321$ $\lambda^{DE}_{31}=.458$ $\lambda^{DE}_{41}=.092$

$\lambda^{DX}_{11}=-.436$ $\lambda^{DX}_{21}=.029$ $\lambda^{DX}_{31}=.100$ $\lambda^{DX}_{41}=.307$

$\lambda^{DR}_{11}=-.001$ $\lambda^{DR}_{21}=.203$ $\lambda^{DR}_{31}=.353$ $\lambda^{DR}_{41}=-.556$

$\lambda^{DFR}_{111}=-.062$ $\lambda^{DFR}_{211}=.206$ $\lambda^{DFR}_{311}=.128$ $\lambda^{DFR}_{411}=-.273$

$\lambda_{111}^{FRX} = .176$

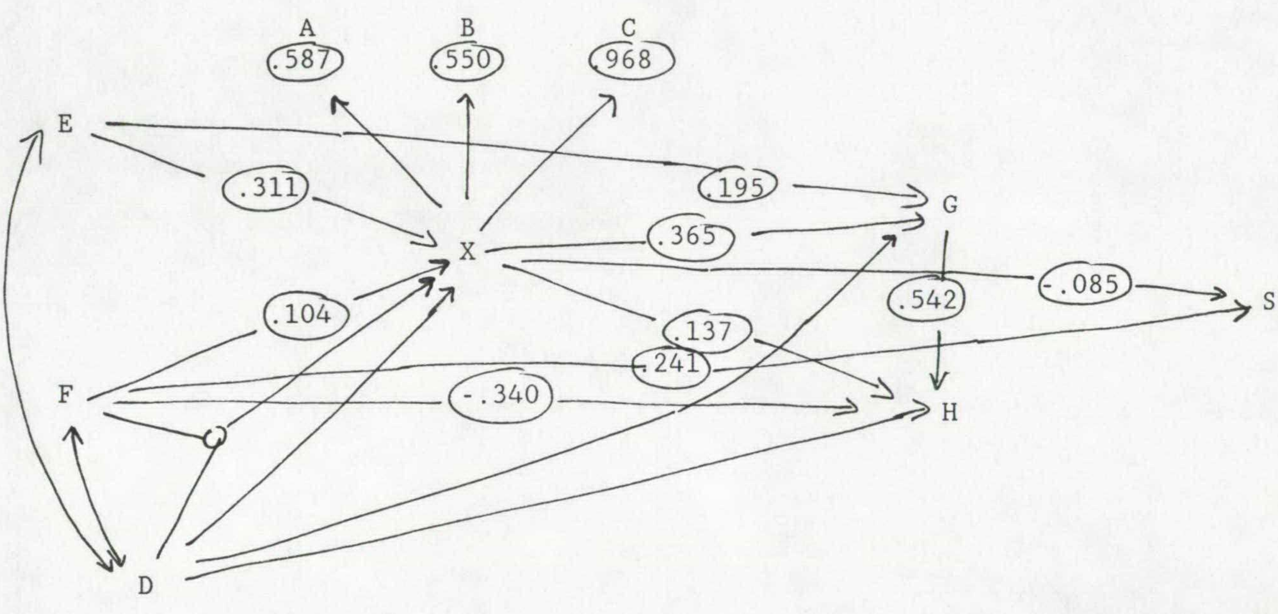
$\lambda_{11}^{DG} = -.359$ $\lambda_{21}^{DG} = -.074$ $\lambda_{31}^{DG} = .402$ $\lambda_{41}^{DG} = .031$

$\lambda_{11}^{DH} = -.213$ $\lambda_{21}^{DH} = -.128$ $\lambda_{31}^{DH} = .143$ $\lambda_{41}^{DH} = .198$

VII.2.1. Toetsingsresultaten voor subgroepen ABCDEFGH en ABCDEF bij een "ignorable" response-mechanisme.

MODEL	L	df	p(x)	X^2	p(x)
1]{AX,BX,CX,DEFGHX,FSX}	591.9	502	.003	595.6	.001
2]{AX,BX,CX}{DE,EF,DF}{DEF,DFX,EX} {DEFX,DG,EG,XG}{DEFGX,DH,FH,HX,GH} {DEFGHX,FSX}	677.0	595	.009	735.3	.000
3]{AX,BX,CX}{DE,DF}{DEF,DFX,EX} {DEFX,DG,EG,XG}{DEFGX,DH,FH,HX,GH} {DEFGHX,FSX}	678.7	596	.009	738.2	.000

VII.2.2. Parameterschattingen voor model 3.



$\lambda^{DE}_{11}=-.796$ $\lambda^{DE}_{21}= .313$ $\lambda^{DE}_{31}= .520$ $\lambda^{DE}_{41}=-.038$

$\lambda^{DF}_{11}=-.071$ $\lambda^{DF}_{21}= .118$ $\lambda^{DF}_{31}= .133$ $\lambda^{DF}_{41}=-.179$

$\lambda^{DFX}_{111}=-.180$ $\lambda^{DFX}_{211}= .155$ $\lambda^{DFX}_{311} = .103$ $\lambda^{DFX}_{411}=-.078$

$\lambda^{DX}_{11}=-.538$ $\lambda^{DX}_{21}= .201$ $\lambda^{DX}_{31}= .296$ $\lambda^{DX}_{41}= .042$

$\lambda^{DG}_{11}=-.401$ $\lambda^{DG}_{21}=-.050$ $\lambda^{DG}_{31}= .392$ $\lambda^{DG}_{41}= .060$

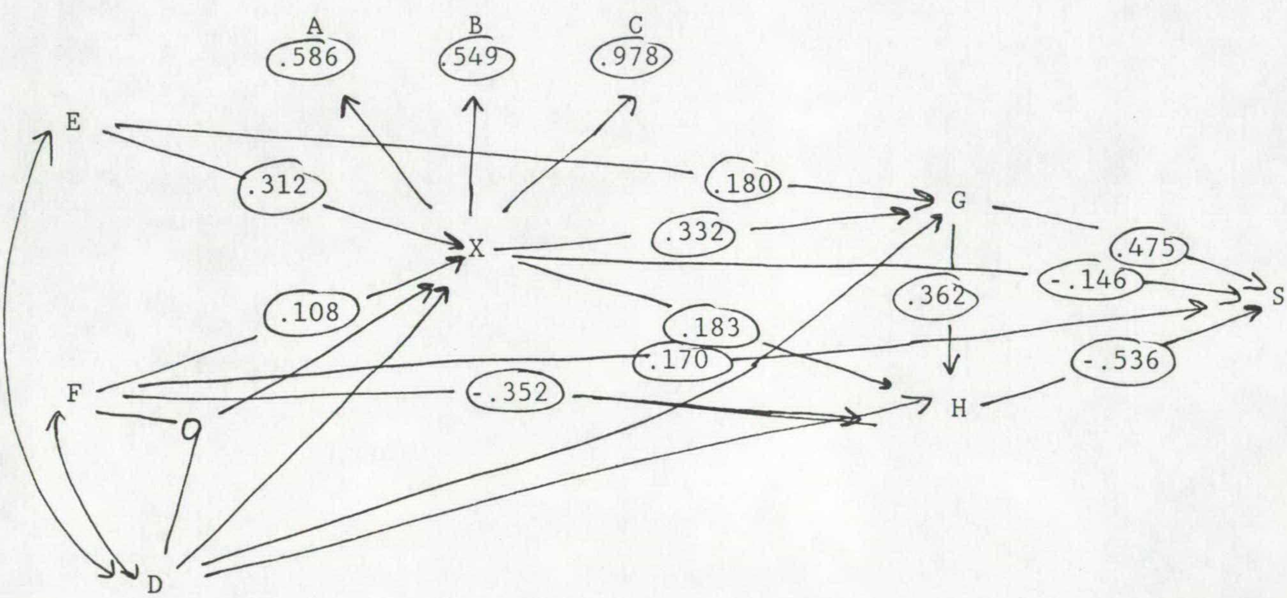
$\lambda^{DH}_{11}=-.184$ $\lambda^{DH}_{21}=-.103$ $\lambda^{DH}_{31}= .162$ $\lambda^{DH}_{41}= .125$

$\lambda^{FRX}_{111}= .097$

VII.3.1. Toetsingsresultaten voor subgroepen ABCDEFGH en ABCDEF bij een "nonignorable" response-mechanisme.

MODEL	L	df	p(x)	X^2	p(x)
1]{AX,BX,CX,DEFGHX,FSX,GS,HS}	565.8	500	.02	581.9	.005
2]{AX,BX,CX}{DE,DF}{DEF,DFX,EX} {DEFX,DG,EG,XG}{DEFGX,DH,FH,HX,GH} {DEFGHX,FSX,GS,HS}	674.6	594	.01	737.1	.000
3]{AX,BX,CX}{DE,DF}{DEF,DFX,EX} {DEFX,DG,EG,XG}{DEFGX,DH,FHX,GH} {DEFGHX,FSX,GS,HS}	661.5	593	.023	721.1	.000

VII.3.2. Parameterschattingen voor model 3.



$\lambda_{11}^{DE}=-.796$ $\lambda_{21}^{DE}=.313$ $\lambda_{31}^{DE}=.520$ $\lambda_{41}^{DE}=-.038$

$\lambda_{11}^{DF}=-.071$ $\lambda_{21}^{DF}=.118$ $\lambda_{31}^{DF}=.133$ $\lambda_{41}^{DF}=-.179$

$\lambda_{111}^{DFX}=-.174$ $\lambda_{211}^{DFX}=.151$ $\lambda_{311}^{DFX}=.099$ $\lambda_{411}^{DFX}=-.076$

$\lambda_{11}^{DX}=-.536$ $\lambda_{21}^{DX}=.199$ $\lambda_{31}^{DX}=.295$ $\lambda_{41}^{DX}=.041$

$\lambda_{11}^{DG}=-.400$ $\lambda_{21}^{DG}=-.003$ $\lambda_{31}^{DG}=.369$ $\lambda_{41}^{DG}=.034$

$\lambda_{11}^{DH}=-.138$ $\lambda_{21}^{DH}=-.138$ $\lambda_{31}^{DH}=.153$ $\lambda_{41}^{DH}=.123$

$\lambda_{111}^{FHX}=.138$ $\lambda_{111}^{FSX}=.116$