

SIMULTANE LOGITANALYSEEen uiteenzetting en toepassing van de methode van Mallar

H.H.U. Konijnendijk*, J.J. Siegers**, C.A. van den Wijngaart***
Economisch Instituut, Rijksuniversiteit Utrecht

1. Inleiding

Herhaaldelijk doet zich in de sociale wetenschappen de combinatie voor van:

- a. tweewaardigheid van een of meer te verklaren variabelen en
- b. simultaniteit.

Wanneer een te verklaren variabele slechts twee waarden kan aannemen, wordt het onderzoek naar de kwantitatieve invloed die uitgaat van de verklarende variabelen bemoeilijkt, doordat de handzame en veel gebruikte analysetechniek van de volgens de gewone kleinste-kwadratenmethode uitgevoerde lineaire regressierekening in een dergelijke situatie op problemen stuit. Logitanalyse vormt dan een geeignende onderzoeksmethode. Twee voorbeelden van tweewaardige (operationaliseringen van) variabelen zijn arbeidsmarktparticipatie (iemand participeert hetzij wel, hetzij niet op de arbeidsmarkt) en vruchtbaarheid (er zijn hetzij wel, hetzij geen kinderen in het gezin). Met betrekking tot de term vruchtbaarheid zij opgemerkt dat hiermee in de demografische literatuur wordt bedoeld op het aantal (levend)geborenen; de biologische potentie kinderen te kunnen krijgen -waarnaar de term vruchtbaarheid op grond van het dagelijkse spraakgebruik lijkt te verwijzen- wordt aangeduid met de term fecunditeit.

Van simultaniteit is sprake wanneer twee of meer variabelen el-

kaar over en weer causaal beïnvloeden. Zo laten de arbeidsmarkt-participatie door gehuwde vrouwen en de vruchtbaarheid van gehuwde vrouwen een sterke onderlinge correlatie zien zonder dat op voorhand de richting van de causaliteit kan worden bepaald. Een dergelijk causaliteitsprobleem valt uiteen in vier vragen:

- loopt de causaliteit van arbeidsmarktparticipatie naar vruchtbaarheid?
- loopt de causaliteit van vruchtbaarheid naar arbeidsmarktparticipatie?
- is er sprake van een wederzijdse causale beïnvloeding en, zo ja, welke van de twee invloeden is dan het sterkste?
- is er sprake van een of meer andere variabelen die zowel arbeidsmarktparticipatie als vruchtbaarheid causaal beïnvloeden, zonder dat beide laatste variabelen een causale samenhang vertonen?

Om deze vragen te kunnen beantwoorden dient een simultane schattingsmethode te worden uitgevoerd. Dat is bovendien vereist, omdat niet-simultane methoden tot foutieve schatters van de regressiecoëfficiënten kunnen leiden (Bronfenbrenner, 1953; Koutsoyiannis, 1973, blz. 322-325). Wanneer er sprake is van niet-lineaire specificaties, zoals bij de toepassing van logitanalyse, kan evenwel niet zonder meer gebruik worden gemaakt van de in geval van een stelsel van lineaire vergelijkingen gewoonlijk gehanteerde methode der kleinste kwadraten in twee of drie ronden gezien het feit dat in het algemeen geldt $Ef(x) \neq f(Ex)$ (Goldfeld en Quandt, 1972, blz. 236). Ook de simultane maximum-likelihoodmethoden bieden nauwelijks uitkomst omdat zij hoewel theoretisch adequaat in de praktijk tot vrijwel onoplosbare problemen leiden (zie Mallar, 1975). Door een iets andere specificatie van het model te kiezen

kon door Mallar een handzame methode worden ontwikkeld om de parameters van een simultaan stelsel van logistische vergelijkingen te kunnen schatten (Mallar, 1975, 1976 en 1977).

In dit artikel¹ wordt de methode van Mallar uiteengezet en nader toegelicht aan de hand van een simultane logitanalyse van de arbeidsmarktparticipatie door gehuwde vrouwen en de aanwezigheid van jonge kinderen in het gezin. Alvorens daartoe over te gaan wordt in paragraaf 2 de wenselijkheid besproken om in geval van tweewaardigheid van een te verklaren variabele in plaats van lineaire regressie logitanalyse toe te passen. Paragraaf 3 bevat een bespreking van Mallar's methode voor het uitvoeren van simultane logitanalyse. Een toepassing van deze methode volgt in paragraaf 4.

2. Lineaire regressie en logitanalyse bij tweewaardige afhankelijke variabelen

Wanneer de te verklaren variabele tweewaardig is, dan geeft de toepassing van lineaire regressie volgens de gewone kleinste-kwadratenmethode onzuivere en niet-consistente schatters voor de varianties van de geschatte regressiecoëfficiënten (Koutsoyiannis, 1973, blz. 178). Dit is het gevolg van het feit, dat de tweewaardigheid van de afhankelijke variabelen er toe leidt, dat er sprake is van heteroscedasticiteit (Ashenfelter, 1969, blz. 644-648; Kooyman, 1971, blz. 39-46), dat wil zeggen dat de variantie van de storingsterm van de te verklaren reeks niet constant is (Goris, 1972, blz. 208). Hoewel deze heteroscedasticiteit in beginsel kan worden vermeden door in plaats van de gewone kleinste-kwadraten-

methode een variant van de gegeneraliseerde kleinste-kwadratenmethode te hanteren, blijkt een dergelijke oplossing in de praktijk niet zonder problemen, doordat de voor deze methode benodigde gewichten extreem groot worden wanneer de geschatte waarden van de te verklaren variabele dicht bij nul of één liggen. Bovendien rijzen moeilijkheden wanneer de geschatte waarden van de te verklaren variabele op een van de grenzen van het interval $[0,1]$ liggen of daarbuiten (zie hieronder) vallen (Siegers, 1977a, blz. 216-218; zie voor de specificatiegevoeligheid van deze oplossing Pindyck en Rubinfeld, 1976, blz. 249).

De te verklaren tweewaardige variabele kan worden geïnterpreteerd als een kansvariabele die de waarde één aanneemt indien een bepaalde gebeurtenis optreedt en de waarde nul indien de betreffende gebeurtenis niet optreedt. Waarden die liggen buiten het interval $[0,1]$ zijn dan op logische gronden onmogelijk. Wanneer in een dergelijk geval een lineair schattingsmodel wordt gehanteerd, doet zich evenwel het probleem voor, dat er geen garantie is, dat de geschatte waarden van de te verklaren variabelen binnen het interval $[0,1]$ vallen (Gunderson, 1974, blz. 90; Siegers, 1977a, blz. 217). Een specificatie die aan dit bezwaar tegemoet komt, wordt gevormd door de logistische curve. De parameters van een dergelijke curve kunnen onder meer met behulp van met name door Goodman en Theil in de sociale wetenschappen ingevoerde methoden op een betrekkelijk eenvoudige wijze worden geschat op basis van in een kruistabel gegroepeerde gegevens (Goodman, 1971-1972 en 1972; Theil, 1969, 1970, 1971, blz. 632-635 en 1972, blz. 166 e.v.; zie ook Davis, 1969; Gunderson, 1977; Heckman, 1976; Nerlove en Press, 1973; Siegers, 1979a; Tacq, 1979; Bishop

c.s., 1977). In geval van continue variabelen luidt het werken met gegroepeerde gegevens evenwel tot verlies van informatie, doordat de variantie tussen de waarnemingen binnen elke afzonderlijke cel wordt verwaarloosd.

In geval van een lineaire specificatie is de kans op het optreden van een bepaalde gebeurtenis gelijk aan:

$$P(Y=1 | X) = \beta'X$$

waarbij:

$P(Y=1 | X)$ = de kans op het optreden van de te verklaren gebeurtenis, gegeven X

Y = tweewaardige dummyvariabele die gelijk is aan één indien de te verklaren gebeurtenis optreedt en gelijk is aan nul indien de gebeurtenis niet optreedt

X = vector van verklarende variabelen

β = vector van regressiecoëfficiënten (β' geeft de getransponeerde van de vector β aan).

(Vectoren worden steeds aangegeven met een vette letter.)

Wanneer een logistische specificatie wordt gehanteerd², dan luidt de kans:

$$P(Y=1 | X) = \frac{1}{1 + e^{-\beta'X}}$$

waarbij $P(Y=1 | X)$ van nul naar één loopt, wanneer $\beta'X$ loopt van $-\infty$ naar $+\infty$. Een methode waarmee de parameters van de logistische curve op basis van individuele waarnemingen kunnen worden geschat, is de maximum-likelihoodmethode (zie met name Warner, 1962, blz. 19-20; de Donnea, 1971, blz. 220-222 en 124-125).

3. De methode van Mallar

Uitgangspunt vormt een model dat G afhankelijke binaire variabelen Y_g bevat. Aangenomen wordt dat de waarschijnlijkheid P_g dat de g^e afhankelijke variabele Y_g gelijk is aan één, niet alleen wordt beïnvloed door gepredetermineerde variabelen, maar eveneens een directe invloed ondervindt van de waarschijnlijkheden op het gelijk aan één zijn van de respectieve andere afhankelijke variabelen in het model. Het model ziet er dan als volgt uit:

$$P_{gi} = F(X_{gi}, \beta_g, P_{gi}^*, Y_g)$$

waarbij $g = 1..G$ en de index $i = 1..N$ het rangnummer van de waarneming aangeeft. De waarschijnlijkheid P_{gi} dat $Y_{gi} = 1$, wordt behalve door de vector X_{gi} van de gepredetermineerde onafhankelijke variabelen met bijbehorende parametervector β_g , ook bepaald door de niet-waarneembare waarschijnlijkheden in de andere vergelijkingen, P_{gi}^* :

$$P_{gi}^* = (P_{1,i}, \dots, P_{g-1,i}, P_{g+1,i}, \dots, P_{G,i})'$$

met bijbehorende parametervector Y_g . De waarschijnlijkheden zijn conditioneel met betrekking tot de verklarende variabelen met vaste coëfficiënten (β_g, Y_g) . In beginsel ligt het voor de hand om als specificatie aan te nemen, dat er een indexfunctie I_{gi} bestaat, die lineair is in $X_{gi}, \beta_g, P_{gi}^*$ en Y_g , zodat de bovenstaande vergelijking kan worden herschreven als:

$$P_{gi} = F(I_{gi})$$

met:

$$I_{gi} = \beta'_g X_{gi} + \gamma'_g P_{gi}$$

In feite komt elke kansverdelingsfunctie in aanmerking voor de functie F . De logistische functie brengt evenwel naar verwachting de relatief geringste hoeveelheid rekenkundige bewerkingen met zich mee zonder dat de resultaten noemenswaard afwijken van de resultaten, die worden verkregen via op andere kansverdelingsfuncties berustende specificaties. Wanneer het logitmodel wordt gehanteerd, geldt:

$$P_{gi} = \left[1 + \exp(-I_{gi}) \right]^{-1}$$

ofwel:

$$I_{gi} = \ln \left[\frac{P_{gi}}{1-P_{gi}} \right] \equiv \text{logit}_{P_{gi}}$$

De cruciale verandering van de specificatie die door Mallar is voorgesteld, is dat wordt aangenomen, dat de indexfunctie I_{gi} niet lineair is in P_{gi}^* maar in de indexfuncties I_{gi}^* van het model, zodat voor de indexfunctie I_{gi} kan worden geschreven:

$$I_{gi} = \beta'_g X_{gi} + \gamma'_g I_{gi}^*$$

met:

$$I_{gi}^* = (I_{1,i}, \dots, I_{g-1,i}, I_{g+1,i}, \dots, I_{G,i})'$$

Deze structuurvergelijking kan eenvoudig worden herschreven in de bijbehorende gereduceerde vorm, waarbij de indexfunctie beschouwd wordt als functie van alle exogene variabelen Z_i in het model:

$$I_{gi} = \pi_g' Z_i$$

waarbij π_g de parametervector is die bij de variabelen Z_i behoort. Op basis van de gereduceerde-vormvergelijkingen kunnen voor alle gepredetermineerde variabelen de totale (multiplier)effecten worden berekend (Siegers, 1979b):

$$\frac{\partial P_g}{\partial X_{gk}} = \pi_{gk} f(I_g)$$

waarbij f de kansdichtheidsfunctie is, die behoort bij de kansverdelingsfunctie F . Omdat de kansdichtheidfunctie niet-negatief is, is het teken van elk van de totale (multiplier)effecten gedetermineerd door het teken van elk van de gereduceerde-vormcoëfficiënten π_{gk} .

Consistente schattingen van de gereduceerde-vormparameters π_g kunnen worden verkregen door de likelihoodfunctie L_g van iedere afhankelijke variabele Y_g afzonderlijk te maximaliseren. De log-likelihoodfunctie van de g^e vergelijking kan worden geschreven als:

$$L_g^* = \ln L_g = \sum_{i=1}^N \left[Y_{gi} \ln P_{gi} + (1 - Y_{gi}) \ln (1 - P_{gi}) \right]$$

met:

$$P_{gi} = F(\pi' z_i)$$

De hier gekozen methode om L_g^* te maximaliseren verloopt als volgt. De vergelijkingen voor de te schatten parameters π_g volgen uit de hierboven gegeven vergelijking voor L_g^* door de afgeleiden van L_g^* naar de gereduceerde-vormparameters nul te stellen. Als er K parameters zijn, dan is er sprake van K niet-lineaire vergelijkingen met K onbekenden, die simultaan moeten worden opgelost. Dit is equivalent aan het berekenen van de waarden die de elementen van de vector π_g moeten aannemen, opdat de volgende functie S_L haar minimum bereikt (waarbij bekend is dat de minimumwaarde van de functie gelijk is aan nul):

$$S_L = \sum_{i=1}^K \left(\frac{\partial L_g^*}{\partial \pi_{gi}} \right)^2$$

De structuurparameters β_g en γ_g kunnen niet direct worden geschat, omdat de indexfuncties I_{gi} niet waarneembaar zijn. Wel kunnen deze parameters worden bepaald met behulp van een door Mallar ontwikkelde methode (Mallar, 1975, 1977), die overeenkomsten vertoont met de kleinste-kwadratenmethode in twee ronden:

- (1) In de eerste ronde worden de gereduceerde-vormparameters π_g geschat door S_L te minimaliseren. Op basis van de aldus geschatte parameters worden met behulp van de gereduceerde-vorm-

vergelijkingen voor elk van de G functies de geschatte waarden van de indexfuncties, $\hat{I}_{gi}^*$, berekend.

- (2) In de tweede ronde worden de parameters van de structuurvergelijkingen geschat, dat wil zeggen van:

$$I_{gi} = \beta_g' x_{gi} + \gamma_g' \hat{I}_{gi}^*$$

waarbij in elk van de G vergelijkingen voor elk van de N waarnemingen voor I_{gi}^* , de in de eerste ronde geschatte waarden, $\hat{I}_{gi}^*$, wordt gesubstitueerd. Evenals de schatting van de gereduceerde-vormparameters in de eerste ronde, kan ook die van de parameters van de structuurvergelijkingen in de tweede ronde worden verkregen door een aan S_L analoge functie te minimaliseren, nu niet met betrekking tot π_g maar met betrekking tot β_g en γ_g .

4. Een simultane analyse van de arbeidsmarktparticipatie door gehuwde vrouwen en de aanwezigheid van jonge kinderen in het gezin

Op grond van theoretische overwegingen lijkt het aannemelijk dat de aanwezigheid van jonge kinderen in het gezin een belangrijke determinant vormt van de arbeidsmarktparticipatie door gehuwde vrouwen (zie bijvoorbeeld Siegers, 1976, en de aldaar vermelde literatuur). Niet-simultane empirische analyses hebben niet tot weerlegging van deze hypothese geleid (zie bijvoorbeeld Siegers, 1976). Evenzo ondersteunen niet-simultane empirische analyses de op grond van theoretische overwegingen geformuleerde

hypothese dat de arbeidsmarktparticipatie door de vrouw een belangrijke determinant vormt van de aanwezigheid van jonge kinderen in het gezin (Siegers, 1977b en 1978b). Overeenkomstig hetgeen in de inleiding is opgemerkt, lijkt een simultane schattingsprocedure vereist. De simultane logitanalyse, waarvan hier de resultaten worden gepresenteerd, heeft betrekking op de gegevens van 350 dertigjarige, gehuwde vrouwen. De betreffende data zijn afkomstig van het Nationaal Onderzoek Vruchtbaarheid en Ouderschapsmotivatie (NOVOM)³. Ten behoeve van het NOVOM zijn in het voorjaar van 1975 ruim 4500 Nederlandse gehuwde vrouwen uit de huwelijkscohorten van 1963 t/m 1973 geënquêteerd (Moors, 1975; Moors e.a., 1976). De arbeidsmarktparticipatie door de vrouw is gelijk aan één gesteld wanneer zij liet weten ten tijde van het interview te werken en gelijk aan nul wanneer zij liet weten dat niet te doen. De aanwezigheid van jonge kinderen in het gezin is geoperationaliseerd door de hierop betrekking hebbende variabele gelijk te stellen aan één wanneer in het gezin kinderen jonger dan zes jaar aanwezig waren en gelijk aan nul wanneer dat niet het geval was. In de vergelijking ter verklaring van de arbeidsmarktparticipatie zijn tevens opgenomen een dummyvariabele die betrekking heeft op het feit of de echtgenoot wel of niet werkzaam was en een continue variabele die het opleidingsniveau van de vrouw representeert. Laatstgenoemde variabele is uitgedrukt in effectieve scholingsjaren door een lagere opleiding gelijk te stellen aan zes effectieve scholingsjaren, een voortgezette lagere opleiding aan negen, een middelbare opleiding aan twaalf, een hogere beroepsopleiding aan vijftien en een opleiding op het hoogste opleidingsniveau aan zeventien (vergelijk Bartels en Bertens, 1975, blz. 139; Bartels, 1977,

blz. 220). De vergelijking ter verklaring van de aanwezigheid van jonge kinderen in het gezin bevat naast de arbeidsmarktparticipatie door de vrouw de huwelijksduur, zowel lineair als kwadratisch, en een viertal dummyvariabelen die betrekking hebben op de kerkelijke gezindte van de vrouw. De voor elk van laatstgenoemde variabelen gevonden regressiecoëfficiënt geeft het effect aan van het behoren tot de betreffende kerkelijke gezindte ten opzichte van het effect van het niet behoren tot een kerkelijke gezindte (Siegers, 1976, blz. 211).

Zoals in paragraaf 3 is uiteengezet, dient voor de schatting van de parameters van het model een aantal malen een functie van meerdere variabelen te worden geminimaliseerd. Omdat er in de praktijk betrekkelijk weinig bekend is omtrent de convergentie van de voorhanden zijnde minimaliseringsmethoden, werd een drietal methoden gehanteerd.

1. De methode van Fletcher en Powell (Fletcher en Powell, 1963).

Dit is een iteratieve methode die in elke stap de schatting van de parametervector verbetert met behulp van alle afgeleiden van de te minimaliseren functie (hier S_L).

2. De methode van Marquardt (Marquardt, 1963). Ook deze methode maakt gebruik van de afgeleiden van de te minimaliseren functie. Ze kan worden beschouwd als een interpolatie tussen de methode van Gauss en de gradiënt-methoden (zoals de "steepest-descent"-methode).

3. De "direct search"-methode van Hooke en Jeeves die in tegenstelling tot de beide voorgaande procedures geen gebruik maakt van de afgeleiden van de te minimaliseren functie (Hooke en Jeeves, 1961). In elke iteratiestap wordt de functie onderzocht in de omgeving van de in de vorige stap bere-

kende parametervector. Door alleen de waarden van de parametervector te accepteren, die een kleinere functiewaarde opleveren, wordt getracht het minimum te vinden.

De resultaten van de uitgevoerde logitanalyse zijn vermeld in tabel 1. De gebruikte symbolen hebben de volgende betekenis:

- E = opleidingsniveau van de vrouw
- Lh = arbeidsmarktparticipatie door de echtgenoot
- D = huwelijksduur
- DD = huwelijksduur, kwadraat
- R1 = kerkelijke gezindte: gereformeerd
- R2 = kerkelijke gezindte: Rooms-katholiek
- R3 = kerkelijke gezindte: Nederlands hervormd
- R4 = kerkelijke gezindte: overige gezindten
- I_{CH} = indexfunctie (zie blz. 6) met betrekking tot de aanwezigheid van jonge kinderen in het gezin
- I_L = indexfunctie (zie blz. 6) met betrekking tot de arbeidsmarktparticipatie door de vrouw

Op grond van het feit dat zowel de bij I_{CH} als de bij I_L behorende regressiecoëfficiënten significant van nul verschillen, kan worden geconcludeerd dat tussen de arbeidsmarktparticipatie door gehuwde vrouwen van dertig jaar en de aanwezigheid van jonge kinderen in hun gezin een wederzijdse causaliteit lijkt te bestaan.

De drie gehanteerde minimaliseringmethoden gaven uiteindelijk alle hetzelfde in tabel 1 vermelde resultaat. Bij de methode van Fletcher en Powell kwam het herhaaldelijk voor, dat tijdens het minimaliseringsproces met grote stappen over het

Tabel 1. Schatting van de parameters met behulp van de maximum-likelihoodmethode*

Variabele	Herleide-vormvergelijkingen						Structuurvergelijkingen					
	Arbeidsmarkt-participatie			Aanwezigheid in het gezin van kinderen jonger dan zes jaar			Arbeidsmarkt-participatie			Aanwezigheid in het gezin van kinderen jonger dan zes jaar		
	L			CH			L			CH		
	b	s	t	b	s	t	b	s	t	b	s	t
E	0.279	0.062	4.52	-0.118	0.067	1.76	0.223	0.060	3.71	- - -		
Lh	-1.529	0.598	2.56	1.576	0.606	2.60	-1.135	0.635	1.79	- - -		
D	-0.650	0.306	2.13	1.794	0.346	5.18	- - -			1.394	0.376	3.70
DD	0.047	0.020	2.34	-0.117	0.023	5.20	- - -			-0.090	0.025	3.65
R1	-0.633	0.479	1.32	1.816	0.783	2.32	- - -			1.550	0.787	1.97
R2	-0.366	0.310	1.18	0.783	0.360	2.18	- - -			0.566	0.372	1.52
R3	-0.440	0.369	1.19	0.183	0.392	0.47	- - -			0.004	0.403	0.01
R4	0.168	0.781	0.22	-1.345	0.741	1.82	- - -			-1.280	0.729	1.76
I _L	- - -			- - -			- - -			-0.582	0.206	2.83
I _{CH}	- - -			- - -			-0.337	0.130	2.59	- - -		
Constante	-0.001	1.470	0.001	-5.552	1.624	3.42	-1.557	0.784	1.90	-4.328	1.325	3.27

*) b = regressiecoëfficiënt; s = asymptotische standaardfout; t = t waarde.

minimum heen werd gesprongen. De methode van Hooke en Jeeves bleek steeds langzamer te convergeren naarmate het minimum werd genaderd. Overeenkomstig hetgeen op grond van deze eigenschappen van beide methoden kan worden verwacht, bleek het efficiënt om het gebied waar het minimum ligt te zoeken met behulp van de methode van Hooke en Jeeves om vervolgens over te gaan op de methode van Fletcher en Powell. Steeds bleek echter de minimaliseringmethode van Marquardt probleemloos en het snelst naar het minimum te leiden. De methoden van Fletcher en Powell en Hooke en Jeeves, al dan niet in combinatie toegepast, benodigden 150 tot 300 berekeningen van de gradiënt van S_L om het minimum te bereiken, de methode van Marquardt daarentegen slechts 30 tot 100.

Noten

- *) Ten tijde van het onderzoek werkzaam bij het Economisch Instituut van de RUU op een post die financieel werd mogelijk gemaakt door een subsidie van het Nationaal Programma Demografisch Onderzoek.
- xx) Wetenschappelijk medewerker bij de sectie macro-economie van het Economisch Instituut van de RUU.
- xxx) Ten tijde van het onderzoek werkzaam bij het Economisch Instituut van de RUU op een post uit de universitaire researchpool.

De schrijvers danken drs. G.P. van der Vorst en drs. A.A. de Jong voor hun commentaar op een eerdere versie van dit artikel.

- 1) De paragrafen 2 en 3 zijn gebaseerd op Siegers (1978c); paragraaf 4 is grotendeels gebaseerd op Siegers en van den Wijngaart (1979).
- 2) Zie voor andere mogelijke transformaties Cox (1970).
- 3) Het NOVOM werd uitgevoerd door het Nederlands Interuniversitair Demografisch Instituut in samenwerking met het CBS en de Werkgroep sociologie van bevolkingsvraagstukken van de Rijksuniversiteit Utrecht. De projectleiding berustte bij het NIDI (Dr. H.G. Moors). Het NOVOM werd mogelijk gemaakt door een subsidie van ZWO. De schrijvers zijn de stuurgroep en de projectleiding van het NOVOM erkentelijk voor het ter beschikking stellen van de enquête-uitkomsten.

Literatuur

- Ashenfelter, O., Some statistical difficulties in using dummy dependent variables, in: W. Bowen en T.A. Finegan, The economics of labor force participation, Princetown, N.J. 1969.
- Ashton, W.D., The logit transformation, Londen, 1957.
- Barteld, C.P.A. en A.M. Bertens, The size distribution of income in Dutch municipalities and regions. Some critical remarks, De Economist, 1975.
- Bartels, C.P.A., Economic aspects of regional welfare, Leiden, 1977.
- Bishop, Y.M.M., S.E. Fienberg en P.W. Holland, Discrete multivariate analysis, Cambridge, Mass., derde druk, 1977.
- Bronfenbrenner, J., Sources and size of least-squares bias in a two-equation model, in: W.C. Hood en T.C. Koopmans (red.), Studies in econometric method, New York, 1953.
- Cox, D.R., The analysis of binary data, Londen, 1970.
- Cramer, J.S. en A.F. de Vos, Simultane Vergelijkingen, Instituut voor Actuarial en Econometrie, Universiteit van Amsterdam, Amsterdam, 1972.
- Davis, N.H.W., Some methods of analysing cross-classified census data. The case of labour force participation rates, Dominion Bureau of Statistics, Ottawa, 1969.
- Domencich, Th.A. en D. McFadden, Urban travel demand. A behavioral analysis, Amsterdam, 1975.
- Donnea, F.X. de, The determinants of transport mode choice in Dutch cities, Rotterdam, 1971.
- Fletcher, R. en M.J.D. Powell, A rapidly convergent descent method for minimization, The Computer Journal, 1963.

- Goldberger, A.S., Econometric theory, New York, 1964.
- Goldfeld, S.M. en R.E. Quandt, Nonlinear methods in econometrics, Amsterdam, 1972.
- Goodman, L.A., A general model for the analysis of surveys, American Journal of Sociology, 1971-1972.
- Goodman, L.A., The modified multiple regression approach to the analysis of dichotomous variables, American Sociological Review, 1972.
- Goris, H., Inleiding in de econometrie, Amsterdam, 1972.
- Gunderson, M., Retention of trainees, A study with dichotomous dependent variables, Journal of Econometrics, 1974.
- Gunderson, M., Logit estimates of labour force participation based on census cross-tabulations, Canadian Journal of Economics, 1977.
- Heckman, J.J., The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models, Annals of Economic and Social Measurement, 1976.
- Hooke, R. en T.A. Jeeves, Direct search solution of numerical and statistical problems, Journal of the Association of Computing Machinery, 1961.
- Johnston, J., Econometric methods, New York, tweede druk, 1972.
- Kooyman, M.A., Onechte variabelen in de econometrie, Assen, 1971.
- Koutsoyiannis, A., Theory of econometrics. An introductory exposition of econometric methods, Londen, 1973.
- Mallar, C.D., The effects of income maintenance on the productive activities of young adults, ongepubliceerde dissertatie, Universiteit van Wisconsin-Madison, 1975.

- Mallar, C.D., Estimating a simultaneous probability model: the school and labor force decisions of youths, Proceedings of the Business and Economic Statistics Section, American Statistical Association, 1976.
- Mallar, C.D., The estimation of simultaneous probability models, Econometrica, 1977.
- Marquardt, D., An algorithm for least-squares estimation of non-linear parameter, Journal of the Society of Industrial and Applied Mathematics, 1963.
- Moors, H.G., Vruchtbaarheidsonderzoek in Nederland: Nationaal onderzoek vruchtbaarheid en ouderschapsmotivatie (NOVOM, 1975), Bevolking en Gezin, 1975
- Moors, H.G., P.J.W.J. van Dongen, M. Niphuis-Nell en H. de Vries, Nationaal Onderzoek Vruchtbaarheid en Ouderschapsmotivatie. Eerste resultaten, Nederlands Interuniversitair Demografisch Instituut, Voorburg, 1976.
- Nerlove, M. en S.J. Press, Univariate and multivariate log-linear and logistic models, Sancta Monica, 1973.
- Pindyck, R.S. en D.L. Rubinfeld, Econometric models and economic forecasts, New York, 1976.
- Siegers, J.J., Participatie op de arbeidsmarkt door gehuwde vrouwen jonger dan 35 jaar in Nederland, Bevolking en Gezin, 1976; ook verschenen als: Labour force participation of married women under 35 years of age in the Netherlands, in: H. G. Moors e.a. (red.), Population and family in the Low Countries II, Leiden, 1978.
- Siegers, J.J., Onderzoek van participatie op de arbeidsmarkt door middel van regressie-analyse waarbij de afhankelijke variabele tweewaardig is, Bevolking en Gezin, 1977a.

- Siegers, J.J., Een regressie-analyse ter verklaring van het aantal kinderen jonger dan zes jaar in gezinnen waarvan de vrouw jonger is dan 35 jaar, Economisch Instituut, Utrecht, 1977b.
- Siegers, J.J., Derksen's study of labour force participation of married women in the Netherlands: some comment and a further analysis, De Economist, 1978a.
- Siegers, J.J., De invloed van huwelijksduur en huwelijksleeftijd op het aantal kinderen jonger dan zes jaar in gezinnen waarvan de vrouw jonger is dan 35 jaar, Economisch Instituut, Utrecht, 1978b.
- Siegers, J.J., Simultane logitanalyse volgens de methode van Mallar, Economisch Instituut, 1978c.
- Siegers, J.J., Logitanalyse op basis van kruistabellen: arbeidsmarktparticipatie door jonge gehuwde vrouwen in Nederland, in: J.J. Klant e.a. (red.), Samenleving en onderzoek, Leiden, 1979a.
- Siegers, J.J., Logitanalyse in geval van niet-lineaire indexfuncties, Economisch Instituut, Rijksuniversiteit Utrecht, 1979b.
- Siegers, J.J. en C.A. van den Wijngaart, Een computerprogramma voor simultane logitanalyse, Economisch Instituut, Utrecht, 1979.
- Tacq, J., Multiplicatieve causale analyse van meerdimensionale contingentietabellen: een toepassing uit de medische sociologie, Sociale Wetenschappen, 1979.
- Theil, H., Logit specifications in the multivariate analysis of qualitative data, University of Chicago, Chicago, 1969.
- Theil, H., On the estimation of relationships involving qualita-

- tive variables, American Journal of Sociology, 1970.
- Theil, H., Principles of econometrics, Amsterdam, 1971.
- Theil, H., Statistical decomposition analysis, Amsterdam, 1972.
- Warner, S.L., Stochastic choice of mode in urban travel: a study in binary choice, Evanston, 1962.