

Universiteit van Amsterdam
Interfaculteit der Actuariële
Wetenschappen en Econometrie
A & E Note N 11/82
november 1982

De kosten van een huisvader
G. Renes en J.S. Cramer

Samenvatting

Bij de vaststelling van de kinderbijslag, van alimentatie of van de schade bij het overlijden van een kostwinner rijst de vraag hoe de verschillende gezinsleden in de totale gezinsuitgaven delen. Het antwoord op die vraag is vanouds vervat in verbruikseenheden die aan de leden van het gezin als naar hun leeftijd en/of geslacht worden toegekend, en die aangeven in welke verhouding ieder in de totale uitgaven deelt. Men kan deze eenheden aan de schrijftafel bedenken, maar kan ze ook uit het waargenomen uitgavenpatroon van gezinnen van verschillende samenstelling afleiden. Dit laatste verdient naar onze mening de voorkeur.

Aan de hand van het Budgetonderzoek 1978 van het C.B.S. hebben wij nagegaan wat de kosten van een huisvader zijn. Daarvoor moeten overigens gelijke huishoudens met en zonder vader worden vergeleken. Het budgetonderzoek bevat maar 22 vaderloze gezinnen die voor dit doel kunnen gebruiken, en dat is weinig. Veel betrouwbaarder ramingen zouden kunnen worden verkregen bij een gericht steekproefonderzoek waarin een behoorlijk aantal onvolledige gezinnen worden betrokken.

Met dit voorbehoud kunnen wij uit ons onderzoek concluderen dat een volwassen Man en Vrouw evenveel verbruikseenheden hebben, en een Kind van onder de 17 de helft daarvan, mits men mag aannemen dat het eerste en het tweede kind evenveel verbruikseenheden vertegenwoordigen. Bij een vervolgonderzoek zouden verschillende leeftijdsklassen moeten worden onderscheiden.

Inleiding en probleemstelling

Bij verschillende praktische vraagstukken zoals de vaststelling van de kinderbijslag, van alimentatie of van de schade bij het overlijden van een kostwinner rijst de vraag hoe men de uitgaven van een huishouden aan de afzonderlijke individuele leden kan toerekenen. Nauw verwant hieraan is de vraag hoe de behoeften van gezinnen van verschillende samenstelling zich tot elkaar verhouden: indien men die behoeften gelijk stelt aan de som van de individuele behoeften, worden beide vragen beantwoord door dezelfde 'behoefte-schaal' voor personen van verschillende soort¹⁾. Vanouds stelt men de behoeften van een volwassen man op 1, en kent men vervolgens aan andere personen al naar hun leeftijd en geslacht ieder een (relatief) aantal verbruikseenheden toe. Telt men deze voor alle gezinsleden op dan levert dat het aantal verbruikseenheden per gezin (ve).

Sinds men zich met het sociale vraagstuk bezig houdt, dat wil zeggen vanaf de vorige eeuw, zijn dergelijke schalen van verbruikseenheden ontworpen met als doel behoeften vergelijkbaar te maken. Deze schalen hebben echter ook nog een andere toepassing, en wel bij het vergelijkbaar maken van Engelcurven. Een Engelcurve beschrijft het verband tussen de uitgaven voor een bepaald goed j of x_j , en het inkomen y , zulks voor verschillende gezinnen die op één tijdstip worden waargenomen zoals in een Budgetonderzoek. Behalve het inkomen kan ook de grootte en samenstelling van het gezin variëren, en de vraag rijst hoe hiermee rekening moet worden gehouden. Het antwoord is dat wij uitgaven en inkomen per verbruikseenheid nemen, of

$$x_j/ve = \phi_j(y/ve); \quad (1)$$

- 1) Strikt genomen gaat het hier en ook verderop steeds om huishoudens in de zin van een groep personen die een (consumptieve) economische eenheid vormen, en niet om gezinnen met een familieband. De meeste huishoudens zijn echter gezinnen, en dat geldt in het bijzonder voor de aangeduide problemen en voor de gegevens die wij zullen analyseren. Wij gebruiken de twee termen derhalve door elkaar.

als de verbruikseenheden juist zijn vastgesteld is y/ve een goede maat voor de levensstandaard van uiteenlopende gezinnen, en dus ook voor hun consumptief gedrag.

De meeste verbruikseenheid-schalen zijn in het verleden aan de schrijftafel bedacht. In ons land wordt tot de dag van vandaag gebruik gemaakt van de zg. Amsterdamse schaal, die in 1917 door ambtenaren van de Arbeidsinspectie is opgesteld ([1], p. 68). De volwassen man telt voor 1 verbruikseenheid, en voor andere gezinsleden gelden de volgende getallen:

leeftijd	M	V
≥ 15	1	.90
14	.90	.90
..	telkens .05 minder	
2	.30	.30
1	.20	.20
0	.15	.15

Deze schaal moge aan het gezond verstand zijn ontsproten, zij maakt wel een erg kunstmatige indruk. Vermoedelijk beschrijft zij in gestyleerde vorm de fysiologisch benodigde hoeveelheid voedsel, en daarmee (bij gegeven prijzen) ook de uitgaven voor voeding die voor de verschillende gezinsleden nodig zijn. De andere uitgaven, die een steeds toenemend deel van het gezinsbudget vormen, komen niet aan de orde; als dat de bedoeling zou zijn kon de schaal ook moeilijk van 1917 tot op heden van toepassing zijn. Ook bij de beperking tot voeding blijft de schaal echter onbevredigend, ondanks zijn veelvuldige toepassing, omdat iedere rechtvaardiging voor de gekozen aantallen verbruikseenheden ontbreekt.

Het is beter een stelsel van verbruikseenheden te kiezen dat vervolgens uit het waargenomen gedrag van gezinnen van verschillende samenstelling kan worden geschat. Deze mogelijkheid ontstaat door te eisen dat de verbruikseenheden de Engelcurven van gezinnen van verschillende omvang tot samenvallen brengen.

Hiervoor dient het model dat in 1921 door Sydenstricker en King [7] werd ontwikkeld en dat in 1956 door Houthakker en Prais [6] opnieuw naar voren werd gebracht. Het uitgangspunt is dat er voor ieder van de $j = 1, 2, \dots, J$ verschillende goederen een apart stel verbruikseenheden geldt, evenals voor het inkomen of de totale uitgaven. De Engelcurven voor goed j kunnen dan voor huishoudens van verschillende samenstelling worden geschreven als

$$x_j / \sum_g \mu_{jg} n_g = \phi_j (y / \sum_g \mu_{og} n_g) \quad (2)$$

met

- x_j de uitgaven aan goederencategorie j
- y het inkomen c.q. de totale uitgaven
- μ_{jg} de verbruikseenheid voor goed j van persoon g
- μ_{og} de algemene verbruikseenheid van persoon g
- n_g het aantal personen van type g .

Zoals men zonder erg veel moeite kan afleiden heeft dit model de aantrekkelijke eigenschap dat de algemene verbruikseenheid μ_{og} van persoon g kan worden uitgedrukt als een gewogen gemiddelde van zijn specifieke verbruikseenheden μ_{jg} (zie Houthakker en Prais [6]). De J coëfficiënten μ_{jg} voor persoon g bepalen dus ook zijn algemene verbruikseenheid μ_{og} . Alle coëfficiënten kunnen in beginsel uit een budgetonderzoek worden afgeleid, mits er één voor ieder persoon g gegeven is; zonder dat is het stelsel niet geïdentificeerd (zie Muellbauer [4]). Dit probleem verdwijnt als men de methode van Nicholson volgt, en bijvoorbeeld ter bepaling van de algemene verbruikseenheid van een jong kind de specifieke verbruikseenheid μ_{jg} voor tabak of alcohol gelijk aan nul stelt, omdat een kind deze produkten niet gebruikt (zie Nicholson [5]).

In het onderhavige geval passen wij deze methode toe om de kosten van een huisvader te berekenen. Voor een bepaald goed zoals dameskleding passen wij een Engelcurve zoals (2) aan aan de gegevens van gezinnen die hetzij uit Man, Vrouw en Kind bestaan, hetzij uit alleen Vrouw en Kind.

Wij beschouwen Vrouw en één Kind tezamen als referentiehuishouding, en kennen daar aan voor alle J goederen en voor het algemeen 1 verbruikseenheid toe. Voorts nemen wij aan dat de verbruikseenheid van een Man voor dameskleding nul is. Voor het geval $j = d$ (van dameskleding) wordt (1) dan vereenvoudigd tot

$$x_d = \phi_d(y/(1+\mu_0 D)) \quad (3)$$

met

D een dummy-variabele die 1 is als het gezin een Man bevat en anders 0.

Specificeert men de Engelcurve ϕ_d nader, dan kan uit dit model de algemene verbruikseenheid van een Man μ_0 worden geschat. Deze coëfficiënt geeft aan met welke fractie het inkomen van een referentiehuishouden van Vrouw en Kind moet toenemen om ze als het ware voor de toetreding van een Man te compenseren, in die zin dat ze evenveel dameskleding kopen als vóór zijn verschijnen.

In het onderhavige onderzoek brengen wij deze aanpak in de praktijk. Dit houdt in dat wij uit een budgetonderzoek gegevens bijeen zoeken over de uitgaven aan dameskleding en andere soortgelijke goederen voor volledige en onvolledige gezinnen, dat wij de functie ϕ nader kiezen, en dat wij de parameters van het aldus verkregen stelsel met daaronder μ_0 netjes schatten.

Wij merken nog op dat natuurlijk precies dezelfde methode zou kunnen worden gebruikt om de coëfficiënt van een huisvrouw vast te stellen. In het Budgetonderzoek 1978 zijn er echter te weinig onvolledige gezinnen zonder vrouw om deze analyse mogelijk te maken.

Gegevens

De gegevens voor dit onderzoek zijn de inkomens en uitgaven aan bepaalde goederen van individuele gezinnen uit het Budgetonderzoek 1978 van het C.B.S. De uitgaven hebben betrekking op vier goederencategorieën, die een volwassen man niet gebruikt, zodat men hem er nul specifieke verbruikseenheden voor kan toekennen. Dit zijn

- . dameskleding
- . damesschoeisel
- . kinderkleding
- . kinderschoeisel

Aanvankelijk is ook gebruik gemaakt van de uitgaven voor speelgoed, maar deze voldeden slecht: de aanpassing was heel gebrekkig, en de inkomenselasticiteit was negatief. Wij hebben dit goed vervolgens buiten beschouwing gelaten.

De waarnemingen die wij in het onderzoek opnemen hebben betrekking op huishoudens van een Vrouw met een of twee Kinderen, al dan niet met bovendien een Man. Aangezien Kinderkleding en Kinderschoeisel in de classificatie van het C.B.S. gedefinieerd zijn als kleding en schoeisel voor personen tot 17 jaar hebben wij deze leeftijdsgrens ook voor het kind gehanteerd. Bij toepassing van deze definitie zijn er onder de huishoudens van het Budgetonderzoek maar 22 vaderloze gezinnen van een Vrouw met 1 of 2 Kinderen. Volledige gezinnen met 1 of 2 Kinderen zijn er genoeg, maar het heeft weinig zin die allemaal op te nemen als het aantal onvolledige gezinnen zo gering blijft. Wij hebben ons dan ook beperkt tot een steekproef van ieder vijfde gezin van deze samenstelling, en dat leverde 129 gezinnen op. Alles tezamen zijn er dus 151 waarnemingen in het onderzoek betrokken, naar samenstelling verdeeld zoals in tabel 1 is weergegeven.

De variabelen die wij gebruiken zijn de uitgaven aan de vier genoemde goederencategorieën en het inkomen. In het Budgetonderzoek zijn de uitgaven van de deelnemende gezinnen langs twee wegen verzameld:

Tabel 1. In het onderzoek betrokken huishoudens naar de samenstelling

	zonder Man	met Man	Totaal
Vrouw met een Kind	13	51	64
Vrouw met twee Kinderen	9	78	87
Totaal	22	129	151

afzonderlijke uitgaven van meer dan f 20,00 diende men gedurende het hele jaar te noteren, kleinere uitgaven alleen gedurende de zg. schrijfmaand. Deze laatste werden vervolgens met 11.5¹⁾ vermenigvuldigd om ze op jaarbasis te brengen, en bij de jaarcijfers voor de grote uitgaven geteld. Het is duidelijk dat er voor schoeisel geen kleine uitgaven voorkomen, zodat wij in beginsel een jaarcijfer hebben; voor kleding kunnen echter kleine uitgaven voorkomen, die langs de weg van het maandcijfer in de jaaruitgaven zijn begrepen. Uit deze hoofde zou men voor schoeisel een betere aanpassing mogen verwachten dan voor kleding, waar de toevallige variatie van maand tot maand een rol kan spelen. Het inkomenscijfer dat wij gebruiken heeft betrekking op het beschikbare inkomen; dit is het bruto inkomen verminderd met belasting en sociale premies.

Nadere gegevens over de wijze waarop de gegevens van het Budgetonderzoek zijn verzameld vindt men in [8].

¹⁾ Niet met 12, omdat er gecorrigeerd wordt voor het aantal vakantiedagen.

Specificatie

Wij moeten nu vier Engelcurven zoals (1) aanpassen aan gegevens over huishoudens die alle een Vrouw en een Kind bevatten, soms een Man en soms een tweede Kind. De verbruikseenheden normeren wij op (Vrouw en 1 Kind), zodat wij werken met het stelsel verbruikseenheden dat in tabel 2 is weergegeven.

Tabel 2. Het gebruikte stelsel verbruikseenheden

	Vrouw + 1 Kind	Man	Tweede Kind
Algemeen	1	μ_0	λ_0
Dameskleding, $j = 1$	1	0	0
Damesschoeisel, $j = 2$	1	0	0
Kinderkleding, $j = 3$	1	0	λ_3
Kinderschoeisel, $j = 4$	1	0	λ_4

Voor de functie ϕ gebruiken wij twee verschillende specificaties.¹⁾ De eerste is het model van Working [9], dat gunstige eigenschappen heeft - zie Deaton en Muellbauer([3], p. 75). Het tweede is het semi-logarithmische model, dat het voordeel van de eenvoud heeft. Na de invoering van twee dummy variabelen

D_m 1 als er een Man is, anders 0

D_k 1 als er een tweede Kind is, anders 0

vinden wij dan als Engelcurve in de Working variant

$$(\frac{x_j}{Y}) = \alpha_j (ve_j/ve) + \beta_j (ve_j/ve) \log y^* + \varepsilon_{-j} \quad (4)$$

en de semi-logarithmische vorm

1) Overigens zou het veronderstellen van dezelfde functionele vorm voor alle goederen in een geheel stelsel Engelcurven tot complicaties leiden, maar dat is hier niet aan de orde.

$$\underline{x}_j = \gamma_j \cdot \text{ve}_j + \delta_j \text{ve}_j \log y^* + \underline{\varepsilon}_j \quad (5)$$

met in beide formules

$$\begin{aligned} \text{ve}_j &= 1 + \lambda_j D_k \\ \text{ve} &= 1 + \lambda_o D_k + \mu_o D_m \\ y^* &= y / \text{ve} \end{aligned}$$

Beide specificaties leiden voor $j = 1, 2, 3, 4$ tot een stelsel van vier vergelijkingen met 12 onbekende parameters, te weten 4 intercepten (α_j , resp. γ_j), 4 inkomenscoëfficiënten (β_j , resp. δ_j) en 4 verbruikseenheden, zoals in Tabel 2 vermeld. Hiervan zijn er twee, te weten μ_o en λ_o , die in alle vier vergelijkingen voorkomen. Wij kunnen deze derhalve niet stuk voor stuk schatten, maar moeten steeds het stelsel van vier simultaan aanpassen.

Wij hebben aan (4) en (5) een additieve storingsvector van vier elementen $\underline{\varepsilon}_j$ toegevoegd. Wij nemen aan dat deze vectoren van gezin tot gezin onderling onafhankelijk zijn, en dat ze een simultane normale verdeling volgen met verwachting nul en covariantiematrix

$$\Sigma = E (\underline{\varepsilon} \underline{\varepsilon}') \quad (6)$$

Schatting van de parameters en aanpassing van de vergelijkingen volgens de methode van Maximum Likelihood is nu equivalent met het minimaliseren van

$$\hat{\Sigma} = \frac{1}{n} \sum_i \underline{e}_i \underline{e}_i' \quad (7)$$

waarin $\underline{e}_i$ de vector van vier residuele afwijkingen voorstelt. Dit betekent dat de simultane schattingsmethode ook ruimte laat voor onderlinge correlatie tussen de storingsvectoren van de vergelijkingen net zoals in het geval van lineaire vergelijkingen bij de zg. SURE methode. Achteraf blijkt echter dat deze correlaties gering zijn, zodat uit dezen hoofde geen noodzaak van simultane schatting bestond (zie appendix).

De schattingen, hun (asymptotische) standaardfouten en andere nuttige grootheden zijn berekend door (7) met behulp van de FIML variant van het TSP-standaardprogramma te minimaliseren. Voor de variantie van de geschatte inkomenselasticiteiten gebruikt dit programma een numerieke benadering. Deze bleek in het geval van de Working Engelcurve heel slecht te voldoen. Wij hebben deze varianties op andere wijze herberekend. Wij interesseren ons vooral voor vier grootheden, te weten:

- De waarde van μ_0 , de algemene verbruikseenheid van een Man tegenover (Vrouw + 1 Kind). Dit is de grootheid waar het ons om te doen is.
- De waarden van λ_0 , λ_3 en λ_4 , verbruikseenheden van een (tweede) Kind. Deze worden als bijproduct verkregen en lenen zich voor beoordeling.
- Hetzelfde geldt voor de inkomenselasticiteiten voor de vier goederencategorieën. Deze zijn voor de Engelcurven (4) en (5) respectievelijk gelijk aan

$$e \text{ (Working)} = 1 + \frac{\beta_j}{\alpha_j + \beta_j \log y^*}$$

$$e \text{ (Semilog)} = \frac{\delta_j}{\gamma_j + \delta_j \log y^*}$$

- R^2 , althans een soortgelijke grootheid, per vergelijking gedefinieerd als

$$R^2 = 1 - \frac{\text{residuele kwadratensom}}{\text{kwadratensom afhankelijke variabele in afwijking van zijn gemiddelde}} \quad (8)$$

Men bedenke dat de simultane aanpassing van vier vergelijkingen niet equivalent is met het maximeren van R^2 per vergelijking; dat de twee stelsels verschillende afhankelijke variabelen hebben; dat de vergelijkingen geen intercepten of constante termen bevatten; en dat er tenslotte van een aanpassing van Engelcurven aan gegevens van 151 individuele gezinnen niet veel mag worden verwacht.

Resultaten 1

Tabel 3. Resultaten bij analyse van 151 waarnemingen

	Working	semi-logarithmisch
μ_0 , verbruikseenheid van een volwassen man	1.01 (.27)	.98 (.37)
verbruikseenheden van een Kind:		
λ_0 , algemeen	.79 (.36)	.53 (.30)
λ_3 , kinderkleding	.75 (.39)	.57 (.23)
λ_4 , kinderschoeisel	.71 (.29)	.77 (.39)
inkomenselasticiteit*)		
dameskleding	.89 (.31)	.76 (.23)
damesschoeisel	.71 (.53)	.70 (.37)
kinderkleding	.82 (.38)	.74 (.20)
kinderschoeisel	.20 (.40)	.13 (.26)
R^2 :		
dameskleding	.12	.11
damesschoeisel	.04	.05
kinderkleding	.04	.14
kinderschoeisel	.10	.33

Tussen haakjes staan standaardfouten. *) bij $y^* = \text{fl. 15.000 (1978)}$

Beide modellen kennen aan een Man praktisch 1 ve toe, of evenveel als (Vrouw en Kind). Dit lijkt erg veel, zeker als men kijkt naar de waarden voor een Kind van 0.8 resp. 0.5: er blijft dan voor de Vrouw slechts 0.2 resp. 0.5 over van wat een Man toekomt, als we tenminste veronderstellen dat het eerste kind evenveel verbruikseenheden vertegenwoordigt als het tweede kind. Dit resultaat voor de coëfficiënt van de Vrouw is niet aanvaardbaar.

De coëfficiënten van een Kind verschillen nogal sterk tussen de twee modellen. Dat de specifieke coëfficiënten kleiner dan 1 zijn zou op 'afdragen' van kleren respectievelijk schoeisel kunnen wijzen.

De meeste inkomenselasticiteiten zijn positief en kleiner dan 1, hetgeen is wat men zou verwachten. De verschillen tussen de twee varianten zijn echter groot. Aangezien die verschillen alleen van μ_0 en λ_0 afhangen wekt dit enige twijfel aan de verkregen schattingen.

De aanpassing, zoals gemeten door R^2 , is slecht, maar dat viel te verwachten. Een opvallende uitzondering is kinderschoeisel in het semi-logarithmische model.

Alles tezamen staan wij vrij sceptisch tegenover deze resultaten. Bij een onderzoek van individuele gezinnen, zoals hier, kan echter het afwijkend gedrag of een anomalie in de waarneming bij een enkel gezin al tot rare resultaten leiden. Wij willen daarom nagaan of zich dergelijke 'uitbijters' onder de 151 gezinnen voordoen.

Uitbijters

Waarnemingen met een sterk afwijkend patroon noemen wij uitbijters. Men kan ze in principe herkennen aan de residuele afwijkingen. Onder de gemaakte veronderstellingen zou de grootheid

$$z_{-i} = e_{-i} \hat{\Sigma}^{-1} e_{-i} \quad (9)$$

bij benadering chi-kwadraat verdeeld zijn met 4 vrijheidsgraden. Wij kunnen deze grootheid voor alle waarnemingen berekenen, en nagaan of de grootste groter is dan verwacht mag worden. Bij een overschrijdingskans van α moeten wij dan een kritische waarde z_{α} hanteren volgens de formule

$$P(z_{-i} > z_{\alpha}) = 1 - (1 - \alpha)^{1/n} \quad (10)$$

Bij $\alpha = 0,05$ en $n = 151$ ligt z_{α} tussen 22.1 en 25.6. Deze aanpak kan ons op het spoor brengen van uitbijters, maar helemaal goed is hij niet: eigenlijk zouden wij (8) voor iedere waarneming moeten berekenen ten opzichte van Engelcurves die aangepast zijn aan de overige 150 waarnemingen, en met de daarbij behorende $\hat{\Sigma}$. Dat is echter te bewerkelijk.

Wel kunnen wij nadat wij een of meer vermoedelijke uitbijters hebben opgespoord een wat betere toets toepassen om na te gaan of zij het werkelijk zijn. Dit is een likelihood-ratio-toets die berust op een vergelijking van het maximum van LogL met en zonder zeg m van de n waarnemingen, of

$$\lambda = 2 \{ \log L(n-m) + \log L(n) \} \quad (11)$$

Het weglaten van m waarnemingen is equivalent met het invoeren van m dummy variabelen, die 1 zijn voor die waarnemingen en nul voor alle andere (zie [2]); in onze stelsels van 4 vergelijkingen zijn daar $4m$ extra parameters aan verbonden. Het gebruik van alle n waarnemingen kan

worden opgevat als schatting onder de restrictieve hypothese dat deze 4m parameters nul zijn, en de toetsgrootheid (11) is onder deze hypothese chi-kwadraat verdeeld met 4m vrijheidsgraden.

Als wij de grootheid (9) voor alle 151 waarnemingen in de beide modellen berekenen en toetsen aan de 5% grens van 22 à 25 zijn er bij het model van Working twee uitbijters, en bij het semi-logarithmische model geen een; z komt daar niet hoger dan 20.4. Volgens de tweede toets (11) zijn de twee gezinnen echter in beide modellen uitbijters *, zij het in het semi-log model veel minder duidelijk dan in het model van Working.

Tabel 4. Twee uitbijters

	Working		semi-logarithmisch	
	gezin a	gezin b	gezin a	gezin b
z volgens (9) ¹⁾	37.9	35.9	11.8	13.7
residuele afwijking:				
. dameskleding	-.007	.142	-71.67	2107.01
. damesschoeisel	.033	.005	439.03	79.84
. kinderkleding	-.031	-.034	-435.65	113.26
. kinderschoeisel	-.009	.005	-104.29	86.12
λ volgens (11) ²⁾	93.2		27.8	

1) 5% kritische waarde tussen 22.1 en 25.6

2) 5% kritische waarde is 15.5

* Nadere berekeningen leren dat volgens de likelihood ratio-toets de gezinnen a en b ook in het semi-logarithmische model ieder afzonderlijk een uitbijter zijn.

Resultaten 2

Tabel 5. Resultaten bij analyse van 149 waarnemingen

	Working	semi-logaritmisch
μ_0 , verbruikseenheid van een volwassen man	.67(.20)	.66(.32)
verbruikseenheden van een kind:		
λ_0 algemeen	.39(.18)	.32(.24)
λ_3 kinderkleding	.47(.23)	.47(.22)
λ_4 kinderschoeisel	.64(.28)	.76(.41)
inkomenselasticiteiten:		
dameskleding	1.03(.24)	.92(.33)
damesschoeisel	.88(.54)	.87(.50)
kinderkleding	.89(.34)	.93(.35)
kinderschoeisel	.37(.30)	.15(.30)
R^2 :		
dameskleding	.30	.17
damesschoeisel	.23	.11
kinderkleding	.05	.14
kinderschoeisel	.12	.33

Tussen haakjes staan de standaardfouten.

Vergelijkt men de resultaten na het weglaten van de uitbijters uit Tabel 5 met de eerdere van Tabel 3 dan valt op dat de twee vergelijkingsstelsels nu in grote trekken tot dezelfde schattingen leiden, en dat een aantal anomalieën is verdwenen. Men kan ook opmerken dat de resultaten wel heel sterk veranderen door het weglaten van maar twee waarnemingen; maar het betreft dan ook onvolledige gezinnen, waar we er maar 22 van hebben, en het waren tenslotte uitbijters.

Voor een Man vinden wij nu .66 verbruikseenheden bij (Vrouw + Kind) = 1, en voor een Kind zeg .35, mits wij het eerste en tweede kind over een kam mogen scheren. Dit betekent dat Man en Vrouw praktisch een gelijke verbruikseenheid hebben. De specifieke coëfficiënten van Kinderen wijzen er op dat Kinderkleding veel en Kinderschoenen wat minder worden afgedragen.

De inkomenselasticiteiten van dameskleding, damesschoenen en kinderkleding zijn 0,9 à 1 die van kinderschoenen is praktisch nul, hetgeen zou inhouden dat zij niet van het inkomen afhangen. Dit zou wellicht voor kleine kinderen met groeiende voeten kunnen gelden, maar wij betwijfelen het vooralsnog.

Bijna alle geschatte coëfficiënten - vooral de verbruikseenheden - hebben behoorlijke standaardfouten, maar het aantal waarnemingen is ook gering.

Conclusies

Wij concluderen dat het mogelijk is langs de aangegeven weg verbruikseenheden te schatten, en dat het er niet zo heel veel toe doet welke vorm van de Engelcurve men kiest. Voorshands stellen wij de algemene verbruikseenheden van een volwassen Man en Vrouw aan elkaar gelijk, en van een Kind onder de 17 op zeg .5 van een volwassene.

Direct moet hieraan worden toegevoegd dat deze schattingen nog een voorlopig karakter dragen, en dat zij aanzienlijk verbeterd zullen worden als men meer waarnemingen verzamelt, en wel in het bijzonder bij onvolledige gezinnen. Het zou de moeite waard zijn eens een gericht steekproef-onderzoek in te stellen naar het uitgavenpatroon van onvolledige gezinnen en daarbij vooral aandacht te besteden aan uitgavencategorieën die zoals hier een nulcoëfficiënt voor de Man of voor de Vrouw hebben.

Bij die gelegenheid zou tevens een nader onderscheid moeten worden gemaakt tussen kinderen van verschillende leeftijd en geslacht, want het is wel duidelijk dat deze groep te heterogeen is om ze als één geheel te behandelen.

Appendix. Correlatiematrices van de residuen.

Nummering van de vergelijkingen:

1. dameskleding
2. damesschoeisel
3. kinderkleding
4. kinderschoeisel

Aanpassing aan 151 waarnemingen

Working				semi-logarithmisch			
1	2	3	4	1	2	3	4
1	.26	.23	.26	1	.26	.30	.18
2		.07	.08	2		.14	.13
3			.08	3			.15
4				4			

Aanpassing aan 149 waarnemingen

Working				semi-logarithmisch			
1	2	3	4	1	2	3	4
1	.29	.27	.26	1	.27	.31	.18
2		.14	.14	2		.17	.15
3			.07	3			.15
4				4			

Literatuur

- [1] Arbeidersbudgets gedurende de crisis (1917),
Leiden, Boek- en Steendrukkerij Eduard Ijdo.
- [2] Belsley, D.A., E. Kuh en R.E. Welsch (1980),
"Regression Diagnostics: Identification of influential
data and sources of collinearity", Wiley series in
probability and mathematical statistics.
- [3] Deaton, A. en J. Muellbauer (1980), "Economics and
consumer behaviour", Cambridge University Press.
- [4] Muellbauer, J. (1975), "Identification and consumer unit
scales", *Econometrica*, Vol. 43.
- [5] Nicholson, J.L. (1949), "Variations in working class
family expenditures", *Journal of the Royal Statistical
Society series A*, Vol. 112, pp. 359-411.
- [6] Prais, S.J. en H.S. Houthakker (1971), "The analysis
of family budgets", Cambridge University Press,
2nd edition overbridged.
- [7] Sydenstricker, E. en W.I. King (1921), "The
measurement of the relative economic status
of families", *Quarterly Publication of the
American Statistical Association*, Vol. 17, pp. 842-57.
- [8] Werknemersbudgetonderzoek 1974/75, deel V/Va,
Methodologie, C.B.S.
- [9] Working, H. (1943), "Statistical laws of family
expenditure", *Journal of the American Statistical
Association*, Vol. 38.