Tom Van Ourti *

Armoedemaatstaven, IGL-curven en de bootstrap: een empirische studie naar de evolutie van het armoedeniveau in België tussen 1985 en 1997

In toegepast onderzoek naar de evolutie van het armoedeniveau gebruikt men meestal de headcount, die het proportioneel aantal armen meet. A. Sen merkte reeds in de jaren zeventig op dat andere aspecten van armoede even belangrijk kunnen zijn, met name het gemiddeld inkomen van de armen en de inkomensongelijkheid onder de armen. In deze studie ga ik na in hoeverre armoedemeting aan de hand van deze additionele aspecten een ander beeld oplevert dan armoedemeting aan de hand van de headcount. Bovendien ga ik na of de veranderingen in het Belgische armoedeniveau tussen 1985 en 1997 statistisch significant zijn.

Inleiding

Armoede is een belangrijk fenomeen in onze samenleving. Niet alleen heeft de overheid allerhande voorzieningen uitgebouwd om armoede te bestrijden (denken we aan bestaansminima, enz.), ook de media besteden regelmatig aandacht aan armoede. Toch krijg ik de indruk dat de visie op het heersende armoedeniveau al te vaak beïnvloed wordt door subjectieve opinies. Dat kan leiden tot allerhande misvattingen. Daarom moeten de beleidsmakers, die mee het armoedebestrijdingsbeleid formuleren, 'objectieve' informatie verkrijgen over het heersende armoedeniveau. Deze zal hen in staat stellen de beperkte middelen voor armoedebestrijding ‘optimaal’ te alloceren.

Armoedemaatstaven zijn instrumenten die tegemoetkomen aan de eis van objectieve informatie. Ze drukken immers het heersende armoedeniveau uit in één getal en worden berekend via vooraf vastgestelde procedures. In een eerste stap stelt men een

* UFSIA (Universiteit Antwerpen)
Deze paper werd mogelijk gemaakt door de medewerking van enkele personen. Dankzij M. Hubert kreeg ik inzicht in de bootstrap methode. K. Van den Bosch suggereerde enkele aanpassingen van de bootstrap procedure. G. Erreygers, D. De Graeve, participanten aan een UFSIA workshop en een anonieme referee leverden kritische commentaren.

armoedegrens op die de populatie opdeelt in armen en niet-armen. Nadien drukt men aan de hand van een armoedemaatstaf het heersende armoedeniveau uit in één getal. De meest gebruikte armoedemaatstaf is de headcount of armoedegraad. Deze geeft het proportioneel aantal armen weer. Met andere woorden, een headcount van 10 percent betekent dat 10 percent van de bevolking arm is. Onder andere het Centrum voor Sociaal Beleid (CSB) gebruikt deze maatstaf. Uit hun onderzoek komt naar voren dat het armoedeniveau quasi constant bleef tussen 1985 en 1992, maar in beperkte mate toenam tussen 1992 en 1997 (Cantillon, De Lathouwer, Marx, Van Dam & Van den Bosch, 1999, blz. 15). Natuurlijk is de headcount niet de enige armoedemaatstaf. Sterker nog, het merendeel van de armoedemaatstaven werd geformuleerd nadat in de economische literatuur een belangrijke kritiek werd geformuleerd op de headcount.

Sen was één van de eersten die wees op de beperkingen van de headcount (Sen, 1976, blz. 219). Hij stelde dat een armoedemaatstaf rekening moet houden met drie aspecten, die volgens hem een wezenlijk deel uitmaken van het begrip armoede. Ten eerste moet men rekening houden met het (proportioneel) aantal armen. Ten tweede moet men het gemiddeld inkomen van de armen betrekken in de analyse. Tenslotte stipt Sen de inkomensongelijkheid onder de armen aan als een belangrijk aspect van armoede.

Deze drie aspecten van armoede zijn belangrijk voor de evaluatie van het beleid. Indien men het (proportioneel) aantal armen meet, kan men enkel het armoedebestrijdingsbeleid evalueren dat het (proportioneel) aantal armen wil reduceren. Met andere woorden, een beleid dat een reductie van de inkomensongelijkheid onder de armen beoogt en/of het gemiddeld inkomen van de armen wil optrekken, beïnvloedt de headcount niet (voor zover het aantal armen ongewijzigd blijft). Een theoretisch voorbeeld: nemen we aan dat de armoedegrens 18.000 BEF bedraagt. In de eerste fase heeft 10 percent van de bevolking een inkomen van 0 BEF en het inkomen van de resterende 90 percent ligt ruim boven de armoedegrens. De overheid beslist dan om een bestaansminimum van 14.000 BEF uit te keren. Deze uitkering wordt op zulke wijze gefinancierd dat geen van de resterende 90 percent arm wordt. Indien de beleidsmakers het armoedebestrijdingsbeleid (enkel)

evalueren aan de hand van de headcount, heeft het beleid geen invloed op het heersende armoedeniveau.
Gaat men hiermee akkoord, dan stelt er zich geen probleem bij het gebruik van de headcount. Vindt men daarentegen dat zulk een beleidsmaatregel de armoedemaatstaf moet beïnvloeden, dan is het noodzakelijk dat men meer dan alleen het (proportioneel) aantal armen in rekening brengt.

Om aan de kritiek op de headcount tegemoet te komen, ontwikkelde Sen een maatstaf die rekening houdt met het aantal armen, hun gemiddeld inkomen en de inkomensongelijkheid onder hen. Nadien werden nog andere maatstaven ontwikkeld, die ook de drie aspecten van armoede incorporeren. Ondanks de consensus over de drie aspecten, verschillen deze maatstaven onderling. Deze verschillen kunnen leiden tot tegengestelde conclusies bij een vergelijking van meerdere inkomensverdelingen. Zo vindt Zheng een verschillende evolutie van het armoedeniveau in de VS naargelang de gebruikte maatstaf (Zheng, 1993, blz. 123). Met andere woorden, een armoedemaatstaf zal steeds in staat zijn meerdere inkomensverdelingen te rangschikken naar armoedeniveau (complete ordening), maar de rangorde hoeft niet dezelfde te zijn voor verschillende armoedemaatstaven. Daarom is er gedurende de laatste 15 jaar veel aandacht uitgegaan naar dominantiecriteria.[1] In deze literatuur is het de bedoeling ondubbelzinnige conclusies te trekken met betrekking tot (de evolutie van) het heersende armoedeniveau voor een groep van armoedemaatstaven.[2] Bij een vergelijking van meerdere inkomensverdelingen zal men niet langer één armoedemaatstaf berekenen, maar eerder dominantiecriteria toepassen die aangeven of alle armoedemaatstaven (van een welbepaalde groep) de inkomensverdelingen op dezelfde wijze rangschikken naar armoedeniveau. In deze paper komen invers gegeneraliseerde Lorenz curven (IGL-curven) aan bod die toelaten ondubbelzinnige conclusies te trekken met betrekking tot het armoedeniveau voor een welbepaalde groep van armoedemaatstaven.

[1] Atkinson (1987), Foster & Shorrocks (1988), Jenkins & Lambert (1997, 1998) en Zheng (1993).
[2] Dominantiecriteria worden eveneens gebruikt om ondubbelzinnige conclusies te trekken met betrekking tot het heersende armoedeniveau voor een groep van armoedegrenzen en voor een groep van equivalentieschalen. Voor een overzicht, zie Zheng (1993).

De eerste onderzoeksvraag in dit artikel luidt als volgt: "Nuanceren de additionele aspecten van armoede het verkregen beeld van armoede of bevestigen ze veeleer de resultaten van de headcount?" Ik zal trachten deze vraag te beantwoorden door na te gaan wat de resultaten van de verschillende maatstaven en de IGL-curven zijn voor de evolutie van het armoedeniveau in België tussen 1985 en 1997.
De tweede onderzoeksvraag focust op statistische significantie. Mills en Zandvakili geven aan dat in het merendeel van de studies rond wijzigingen in inkomensongelijkheid niet wordt nagegaan of de wijzigingen ook statistisch significant zijn (Mills & Zandvakili, 1997, blz. 133). Dezelfde tekortkoming kan men terugvinden bij empirische meting van het armoedeniveau. Een bijkomend probleem betreft het gebrek aan klassieke statistische methoden om na te gaan of veranderingen in het armoedeniveau significant zijn. Daarom zal ik aan de hand van de bootstrap methode nagaan of eventuele verschuivingen in het armoedeniveau statistisch significant zijn. In de appendix wordt een uitvoerige beschrijving van deze methode gegeven.
Merk op dat geen van beide onderzoeksvragen ingaat op gevolgen en/of oorzaken van armoede.

In de volgende paragraaf bespreek ik de verschillende armoedemaatstaven en de IGL-curven. In paragraaf 2 worden de gegevens besproken. De derde paragraaf behandelt de resultaten.

1. De meting van armoede: maatstaven en dominantiecriteria

In deze paragraaf beschrijf ik de IGL-curven en enkele armoedemaatstaven. Ik beoog hierbij geenszins volledigheid, maar beperk me tot zes armoedemaatstaven die later in het empirische luik aan bod komen.[3]

A. Assumpties

[3] Voor een goed overzicht, zie Seidl (1986) en Zheng (1997).

Vooraleer de IGL-curven en maatstaven zelf te bespreken, komen drie belangrijke kanttekeningen aan bod. Alle zijn essentieel bij armoedemeting, maar worden in dit artikel niet behandeld aangezien ik focus op de twee onderzoeksvragen. Verder wordt een kort overzicht van symbolen gegeven.

Ten eerste neem ik bij de beschrijving van de armoedemaatstaven aan dat reeds uitgemaakt is wie tot de armen behoort. Met andere woorden, een armoedegrens werd geconstrueerd die de bevolking opdeelt in armen en niet-armen. In het empirisch gedeelte kies ik voor één welbepaalde armoedegrens. Ik zal niet nagaan wat de implicaties zijn van andere armoedegrenzen. Nochtans is dit een belangrijke onderzoeksvraag op zich omdat de hoogte van de armoedegrens de uiteindelijke resultaten aanzienlijk kan beïnvloeden.[4]

Ten tweede wil ik wijzen op de verschillende waardeoordelen die impliciet of expliciet gepaard gaan met het meten van armoede. Enkele voorbeelden: moeten we rekening houden met de inkomens van de niet-armen, met hun aantal of zijn geen van beiden van belang? Daalt het armoedeniveau sterker indien we de 'armste' der armen of de 'rijkste' der armen uit de armoede lichten? Op welke manier moet een armoedemaatstaf reageren op inkomenstransfers tussen de armen? enz. In de theoretische literatuur werden axioma's ontwikkeld. Elk axioma komt overeen met een welbepaald waardeoordeel. De axioma's waaraan een armoedemaatstaf voldoet, geven uitdrukking aan de waardeoordelen waarmee een maatstaf rekening houdt.[5]

Ten derde houd ik enkel rekening met het inkomen van de individuen. Met andere worden, er wordt abstractie gemaakt van het multi-dimensionale karakter van armoede. Dat dit een zeer sterke assumptie is, behoeft geen betoog. In het empirisch gedeelte corrigeer ik enkel voor verschillen tussen individuen uit één- en meerpersoonsgezinnen.

[4] Voor een uitgebreid overzicht, zie Van den Bosch (1999).

[5] De axiomatische benadering is dominant. Andere benaderingen zijn gebaseerd op sociale welvaartsfuncties of zijn totaal ad hoc.

Ten vierde, ga ik uit van een geordende (gerangschikt naar inkomen) inkomensverdeling met n leden waarbij y_i staat voor het inkomen van het i^{de} individu. Verder wordt een armoedegrens (z) gedefinieerd die de populatie opdeelt in q armen en $n-q$ niet-armen waarbij $y_1 \leq y_2 \leq ... \leq y_q < z$.

B. Armoedemaatstaven en IGL-curven

Grosso modo zijn armoedemaatstaven onder te verdelen in twee klassen, namelijk de axiomatische armoedemaatstaven, die de drie aspecten van armoede incorporeren en de ad hoc armoedemaatstaven die met enkel één aspect rekening houden. Alle armoedemaatstaven zijn zodanig genormaliseerd dat ze een waarde tussen nul (minimale armoede) en één (maximale armoede) aannemen. Tenslotte sta ik ook stil bij IGL-curven.

In de onderstaande uiteenzetting gebruik ik armoedekloven als leidraad. De armoedekloof van het i^{de} individu wordt gedefinieerd als de inkomensafstand tussen de armoedegrens en het inkomen van het i^{de} individu, met name $z-y_i$. De armoedemaatstaven verschillen enkel in de wijze waarop ze armoedekloven optellen.

1. De ad hoc armoedemaatstaven

Hieronder komen de headcount, de inkomenskloof ratio en de Gini coëfficiënt aan bod. Deze maatstaven zijn relevant in zoverre men enkel het onderliggende aspect wil meten.

De headcount houdt enkel rekening met het proportioneel aantal armen. Formeel kan deze voorgesteld worden als een (gewogen) som van de armoedekloven verheven tot de nulde macht over de armen:

$$H = \frac{1}{n}\sum_{i=1}^{q}(z-y_i)^0 = \frac{q}{n} \tag{1}$$

Uit de formule wordt meteen duidelijk dat het gemiddeld inkomen van de armen en de verdeling van armoede buiten beschouwing worden gelaten.

De inkomenskloof ratio daarentegen benadrukt het gemiddeld inkomen van de armen. De maatstaf wordt gedefinieerd als het procentuele verschil tussen de armoedegrens en het gemiddeld inkomen van de armen. Algebraïsch komt men tot een (gewogen) som van de armoedekloven over de armen:

$$I = \frac{1}{qz}\sum_{i=1}^{q}(z - y_i) = 1 - \frac{\mu_q}{z} \tag{2}$$

met μ_q het gemiddeld inkomen van de armen. Uit de algebraïsche vorm blijkt dat weinig belang gehecht wordt aan het proportioneel aantal armen. Ook de verdeling speelt geen rol. Immers, een inkomenstransfer tussen twee armen beïnvloedt het gemiddeld inkomen van de armen (μ_q) niet.

De laatste ad hoc armoedemaatstaf is de Gini coëfficiënt voor de inkomensverdeling van de armen. Formeel komt men tot de volgende voorstelling:

$$G_q = \frac{1}{\mu_q q^2}\sum_{i=1}^{q}\sum_{j=1}^{q}\left|\left(z - y_j\right) - \left(z - y_i\right)\right| = 1 + \frac{1}{q} - \frac{2}{\mu_q q^2}\sum_{i=1}^{q}(q + 1 - i)y_i \tag{3}$$

In feite is dit geen armoedemaatstaf, maar een ongelijkheidsmaatstaf. Ik gebruik deze om het ongelijkheidaspect van armoede te becijferen.

2. De axiomatische armoedemaatstaven

In deze paragraaf komen de Sen (1976), de Foster, Greer & Thorbecke (1984) en de Watts (1968) maatstaf aan bod. Alhoewel deze drie maatstaven onderling sterk verschillen, houden ze alle rekening met de drie aspecten van armoede.

Zoals vermeld in de inleiding, is de Sen maatstaf de eerste axiomatische armoedemaatstaf.[6] Algebraïsch wordt deze voorgesteld als volgt:

[6] De Sen maatstaf is de eerste axiomatische armoedemaatstaf die theoretisch onderbouwd werd. Watts formuleerde vóór Sen een axiomatische maatstaf, maar deze is niet theoretisch onderbouwd.

$$S = \frac{2}{(q+1)nz}\sum_{i=1}^{q}(z-y_i)(q+1-i) = H\left[I + (1-I)G_q\left(\frac{q}{q+1}\right)\right] \qquad (4)$$

De tweede vergelijking verduidelijkt dat de Sen maatstaf rekening houdt met de drie aspecten van armoede. De maatstaf is immers te herleiden tot een functie van de headcount, inkomenskloof ratio en de Gini coëfficiënt voor de inkomensverdeling van de armen. De implementatie van de verdelingsaspecten kan beter afgeleid worden uit de eerste vergelijking. De verdelingsaspecten worden weergegeven door middel van de rangorde van de armen ($q+1-i$). Met andere woorden, diegene met een lage rang (geassocieerd met een laag inkomen) krijgen een zwaarder gewicht bij de (gewogen) sommatie van de armoedekloven dan diegene met een hoge rang, namelijk de armste arme krijgt een gewicht q, de volgende $q-1$ en de rijkste arme wordt het gewicht 1 toegekend.

De belangrijkste kritiek op deze maatstaf betreft juist het gebruik van de rangorde voor het weergeven van de verdelingsaspecten. Zo stelt Zheng dat geen aandacht besteed wordt aan de grootte van de inkomensafstand tussen twee opeenvolgende armen: "In poverty, where each poor person knows exactly his income distance from the 'reference point' - the poverty line (or maximum income of the population) - it is less convincing to use a person's relative ranking, which is rarely known" (Zheng, 1997, blz. 145). Met andere woorden, de hoogte van het inkomen van een arme heeft een beperkte invloed op de rangorde.
Een tweede kritiek betreft het verband met de Gini index. De Gini index (en dus ook de Sen maatstaf) hecht des te meer belang aan inkomenstransfers naarmate zich meer individuen tussen beide 'transfererende' partijen bevinden. "If, as is typically the case, most of the poor are bunched at or near the poverty line, then it is there, on the Sen index, that transfers would be most effective in alleviating poverty" (Borooah, 1991, blz.31).[7]

[7] Een andere kritiek betreft vergelijkingen tussen inkomensverdelingen met een verschillend aantal leden. De Sen maatstaf zal een verschillend armoedeniveau meten voor een inkomensverdeling X en een inkomensverdeling Y indien Y bestaat uit r replicaties van X. Werkt men echter met voldoende grote steekproeven, dan is dit verschil verwaarloosbaar.

De Foster, Greer & Thorbecke maatstaf heeft de meeste aandacht gekregen in de economische literatuur. Dit omdat Foster, Greer & Thorbecke de gewichten waarmee de genormaliseerde armoedekloven $(1-{y_i}/{z})$ worden opgeteld op een zeer eenvoudige en zeer intuïtieve wijze weergeven. De gewichten zijn namelijk niets meer dan de genormaliseerde armoedekloven zelf, verheven tot de macht $\alpha - 1$. Algebraïsch wordt deze maatstaf als volgt voorgesteld:

$$FGT(\alpha) = \frac{1}{n}\sum_{i=1}^{q}\left(1-\frac{y_i}{z}\right)\left(1-\frac{y_i}{z}\right)^{\alpha-1} \quad met\ \alpha > 1 \tag{5}$$

α kan geïnterpreteerd worden als een armoede-aversieparameter. Elke keuze van α komt overeen met een welbepaalde visie op armoede. Een hogere α benadrukt de laagste inkomens. De parameter heeft als voordeel dat de onderzoeker de maatstaf voor verschillende waarden α kan berekenen en zodoende kan nagaan of de armoedeordening tussen twee inkomensverdelingen beïnvloed wordt door de visie op armoede.

Voor α=1 bekomen we het product van de headcount en de inkomenskloof ratio. De gewichten zijn dan gelijk aan 1. Indien α>1 verkrijgen we axiomatische armoedemaatstaven.

De derde axiomatische armoedemaatstaf die ik zal gebruiken, is de Watts maatstaf. Algebraïsch wordt deze voorgesteld als volgt:

$$W = \frac{1}{n}\sum_{i=1}^{q}(\ln z - \ln y_i) \tag{6}$$

In vergelijking met de Foster, Greer & Thorbecke maatstaf bezit deze maatstaf geen parameter. Zodoende kan men niet nagaan wat de implicaties zijn van verschillende visies op armoede.

Alhoewel de axiomatische armoedemaatstaven rekening houden met de drie aspecten van armoede, verschillen ze onderling. Deze verschillen kunnen leiden tot tegengestelde conclusies met betrekking tot (de evolutie van) het armoedeniveau. Als daarentegen alle armoedemaatstaven behorend tot een bepaalde groep dezelfde tendens weergeven, kan men (meer) ondubbelzinnige uitspraken doen over (de evolutie van) het armoedeniveau. De benadering van dominantiecriteria komt hieraan tegemoet. In de volgende paragraaf worden IGL-curven besproken.[8]

3. *Invers generaliseerde Lorenz curven (IGL-curven)*

Jenkins & Lambert ontwikkelden de IGL-curven. Deze curven laten toe dominantiecriteria na te gaan voor een groep van *axiomatische* armoedemaatstaven, met name de 'gegeneraliseerde armoedekloof' armoedemaatstaven (GAA). Met andere woorden, indien men op basis van de IGL-curven besluit dat het armoedeniveau van een inkomensverdeling X hoger is dan dat van een inkomensverdeling Y, dan geldt deze conclusie niet voor één armoedemaatstaf, maar voor alle armoedemaatstaven die behoren tot de GAA. Bovendien geven de IGL-curven de drie aspecten van armoede op een visuele wijze weer (Jenkins & Lambert, 1997, blz. 317-327). Een IGL-curve wordt als volgt gedefinieerd:

$$IGL(g,\frac{k}{n}) = \sum_{i=1}^{k} \frac{g_i}{n};\ g \in R^n;\ g_1 \geq g_2 \geq \ldots \geq g_n;\ g_i = max(z - y_i, 0) \qquad (7)$$

Men berekent dan n punten door k te laten variëren van 1 tot n. Aangezien k een discreet getal is, wordt de continue curve bekomen door middel van interpolatie (Jenkins & Lambert, 1998, blz. 41).

Indien men inkomensverdelingen met een verschillende armoedegrens wenst te vergelijken, kan men de gewone IGL-curven niet gebruiken. Men maakt dan beter gebruik van de genormaliseerde IGL-curven. Deze curven zijn dusdanig gedefinieerd dat er zich geen problemen voordoen met correcties voor inflatie en koopkrachtpariteit. Aangezien ik de evolutie van het Belgische armoedeniveau tussen 1985 en 1997 bestudeer, gebruik ik de genormaliseerde curven. Zij worden als volgt gedefinieerd:

$$IGL(r,\frac{k}{n}) = \sum_{i=1}^{k} \frac{r_i}{n};\ r \in R^n;\ r_1 \geq r_2 \geq \ldots \geq r_n;\ r_i = max(\frac{z - y_i}{z}, 0) \qquad (8)$$

[8] Andere dominantiecriteria zijn gebaseerd op gegeneraliseerde Lorenz curven (Foster & Shorrocks (1988)) en 'poverty deficit curves' (Atkinson (1987)).

Figuur 1 beeldt een genormaliseerde IGL-curve af. Deze curve vat de drie aspecten van armoede samen. Het proportioneel aantal armen vinden we terug op de horizontale as, met name dat punt op de horizontale as (q/n) waar de genormaliseerde IGL-curve horizontaal wordt. De inkomenskloof ratio vinden we terug in de helling van de rechte die de oorsprong verbindt met het punt ($q/n, \sum_{i=1}^{q} r_i/n$). De verdeling van armoede wordt weergegeven door de concaviteit van de genormaliseerde IGL-curve (Jenkins & Lambert, 1997, blz. 319).

[figuur 1]

Belangrijker dan de visuele weergave van de drie aspecten van armoede zijn de dominantiecriteria verbonden met de genormaliseerde IGL-curven. Indien de genormaliseerde IGL-curve van een inkomensverdeling X volledig boven de genormaliseerde IGL-curve van een inkomensverdeling Y ligt (of X IGL-domineert Y), kan men besluiten dat X gekenmerkt wordt door een hoger armoedeniveau dan Y. Deze conclusie geldt voor alle armoedemaatstaven die behoren tot de GAA (Jenkins & Lambert, 1998, blz. 44), met name de maatstaven ontwikkeld door Chakravarty (1983), Clark, Hemming & Ulph (type 2) (1981), Foster, Greer & Thorbecke (1984), Hagenaars (1987), Johnson (1988), Pyatt (1987), Shorrocks (1995), Thon (1983) en Watts (1968). Indien daarentegen de genormaliseerde IGL-curven van de twee inkomensverdelingen snijden, kan men niet dezelfde uitspraak doen voor alle maatstaven die behoren tot de GAA. Met andere woorden, de IGL-curven geven slechts een partiële armoedeordening van de inkomensverdelingen, aangezien enkel in het geval van IGL-dominantie gegarandeerd wordt dat alle GAA maatstaven de inkomensverdelingen op dezelfde wijze rangschikken.

Jenkins en Lambert ontwikkelden daarnaast een methodologie die toelaat na te gaan hoever de armoedegrens van X mag dalen (terwijl die van Y constant blijft) zonder de IGL-dominantie aan te tasten (Jenkins & Lambert, 1998, blz. 49). Op deze wijze kan men nagaan in hoeverre de IGL-dominantie gevoelig is voor de hoogte van de armoedegrens van X.[9] Een voorbeeld: stel dat (1) de armoedegrens van X 18.000 BEF en die van Y 16.000 BEF bedraagt en dat (2) X gekenmerkt wordt door een

[9] Voor een uitgebreide discussie, zie theorema 4 en 5 (Jenkins & Lambert, 1998, blz. 49-51).

hoger armoedeniveau dan Y (of X IGL-domineert Y). Gegeven deze uitgangssituatie zou men kunnen vermoeden dat het hogere armoedeniveau van X te wijten is aan de hogere armoedegrens van X. Maar indien blijkt dat de armoedegrens van X mag dalen tot bijvoorbeeld 14.000 BEF, zonder de IGL-dominantie aan te tasten, kan men besluiten dat zelfs voor deze lagere armoedegrens het armoedeniveau van X hoger is dan dat van Y. Dit wijst erop dat niet (enkel) de armoedegrens verantwoordelijk is voor het hogere armoedeniveau. Indien in de uitgangssituatie de armoedegrens van X reeds lager is dan die van Y is het meteen duidelijk dat de armoedegrens van X niet verantwoordelijk is voor de IGL-dominantie.

Samengevat kan men stellen dat er drie voordelen verbonden zijn aan het gebruik van (genormaliseerde) IGL-curven. Ten eerste, IGL-dominantie garandeert dat de armoedeordening niet voor één, maar voor alle GAA armoedemaatstaven geldt. Ten tweede, de IGL-curven geven de drie aspecten van armoede visueel weer en men hoeft maar één curve te berekenen in plaats van alle GAA maatstaven. Ten derde, kan men de robuustheid van de IGL-dominantie nagaan met betrekking tot de hoogte van de armoedegrens. Dit kan de discussie omtrent het belang van de hoogte van de armoedegrens vergemakkelijken.

2. De gegevens

Bij het berekenen van de verschillende armoedemaatstaven maak ik gebruik van panelgegevens verzameld door het Centrum voor Sociaal Beleid (CSB). Het CSB voerde in 1985, 1988, 1992 en 1997 enquêtes uit bij een representatieve steekproef van Belgische huishoudens. Wel moet opgemerkt worden dat de steekproef louter bestaat uit private huishoudens. Hieruit volgt dat ze niet representatief is voor daklozen, personen in een instelling, enz. Verder wordt de steekproef in alle jaren gewogen. De weging is doorgevoerd om te corrigeren voor "differentiële selectiekansen ten gevolge van de panel opvolgingsregels en eventuele differentiële non-respons" (Cantillon, Marx, Proost & Van Dam, 1993, blz. 8). Zodoende blijft de steekproef representatief voor de Belgische huishoudens op het moment van de enquêtes.

Er werden 6471 huishoudens ondervraagd in 1985, 3779 in 1988, 3821 in 1992 en 4616 in 1997. "Het inkomen wordt in de CSB-enquêtes bevraagd op maandbasis. Het totaal beschikbaar inkomen van het huishouden omvat voor alle leden de nettolonen uit hoofd- en bijberoep, de netto-bedrijfsinkomsten voor zelfstandigen, de sociale zekerheidsuitkeringen en diverse inkomens. Het vermogen en inkomsten uit roerende goederen blijven in deze analyse buiten beschouwing" (ibidem, blz. 9).

Aangezien ik de evolutie van armoede beschouw, gebruik ik het individu (en niet het huishouden) als analyse-eenheid. Een huishouden is immers onderhevig aan veranderingen (echtscheidingen, kinderen die het ouderlijk huis verlaten, etc.). Bovendien impliceert het gebruik van het huishouden als analyse-eenheid dat huishoudens van een verschillende grootte hetzelfde effect hebben op het armoedeniveau. “There seems to be no good reason why the poverty measure would increase more when two poor single persons enter poverty, than when a couple with two children does so” (Van den Bosch et al., 1993, blz. 247).

Zodoende komt men tot 17715, 10383, 10004 en 11322 individuen in respectievelijk 1985, 1988, 1992 en 1997.[10] Om vergelijkbare inkomens voor de individuen te bekomen, worden equivalentieschalen toegepast. Meer bepaald gebruik ik de EG equivalentieschaal.[11] Deze kent een gewicht 1 toe aan de eerste volwassene, 0.7 aan bijkomende volwassenen en 0.5 aan kinderen. Het *individueel* equivalent inkomen wordt dan gedefinieerd als het huishoudinkomen gedeeld door de som van de equivalentiefactoren voor dat huishouden. Dit equivalent inkomen wordt toegekend aan elk lid van het huishouden.

3. De resultaten

Deze paragraaf is opgedeeld in twee delen. In het eerste deel komen de resultaten voor de verschillende armoedemaatstaven aan bod. In het tweede gedeelte presenteer ik de resultaten voor de IGL-curven.

[10] Het betreft de gewogen aantallen.
[11] Het CSB gebruikt dezelfde schaal. Dit maakt deze studie vergelijkbaar met de resultaten van het CSB.

A. De armoedegrenzen en de armoedemaatstaven

In deze studie heb ik gekozen voor de EG-grens. De EG-grens is relatief van aard en wordt gedefinieerd als de helft van het gemiddeld *individueel* equivalent inkomen. De definitie impliceert dat de armoedegrens in dezelfde richting evolueert als het algemene welvaartspeil. Ik kies voor de EG-grens omwille van de vergelijkbaarheid met het CSB. In tabel 1 worden de armoedegrenzen weergegeven.

[tabel 1]

Op basis van de linkerkolom is een sterke stijging van de armoedegrens waar te nemen tussen 1985 en 1997. Zelfs indien we corrigeren voor inflatie (rechterkolom) blijft de stijging aanzienlijk. Dit is ook hetgeen men zou verwachten aangezien de grens relatief van aard is. Een hoger gemiddeld inkomensniveau impliceert een hogere armoedegrens. In de appendix worden teststatistieken weergegeven (appendix C, tabel A.1). Daaruit kunnen we afleiden dat de stijging van de armoedegrens statistisch significant is. In paragraaf B wordt met behulp van IGL-curven nagegaan in hoeverre het verschil tussen de armoedegrenzen determinerend is voor het uiteindelijke resultaat.

[tabel 2]

In tabel 2 zijn de resultaten van de armoedemaatstaven per jaar opgenomen. Merk op dat in deze tabel nog abstractie gemaakt wordt van statistische significantie.
Indien men de ordening van de armoedemaatstaven bestudeert, merkt men dat alle axiomatische armoedemaatstaven (laatste drie rijen) dezelfde ordening opleveren. Met name, 1992 wordt gekenmerkt door het hoogste armoedeniveau en 1988 door het laagste. 1985 en 1997 worden steeds als respectievelijk twee en drie geordend. Vergelijken we deze rangorde met die van de headcount, dan komt men tot de conclusie dat er enkel voor 1988 overeenstemming is. Dit geeft, mijn inziens, reeds een eerste indicatie dat de axiomatische maatstaven een alternatief zijn voor de headcount.

De resultaten van de headcount verschillen wezenlijk van de resultaten van de inkomenskloof ratio en de Gini coëfficiënt. De ordeningen van deze laatste twee maatstaven vertonen daarentegen meer overeenkomsten met de armoedeordening van de axiomatische armoedemaatstaven. Hieruit kan men besluiten dat, tussen 1985 en 1997, het gemiddeld inkomen van de armen en de inkomensongelijkheid onder hen primeert boven het proportioneel aantal armen. Beperken we bijvoorbeeld de aandacht tot 1992 en 1997. Hoewel 1997 gekenmerkt wordt door een hoger proportioneel aantal armen, zijn de inkomenskloof ratio en de Gini coëfficiënt van zulk een waarde in 1992, dat alle axiomatische armoedemaatstaven 1992 als armer bestempelen dan 1997.

In het voorgaande heb ik abstractie gemaakt van statistische significantie. De vraag stelt zich of de veranderingen van het armoedeniveau tussen 1985 en 1997 ook statistisch significant zijn. Figuur 2 geeft een synthese van de resultaten weer. In appendix A, B en C wordt de methodologie beschreven. De originele teststatistieken kan men terugvinden in appendix C (tabel A.2).

[figuur 2]

Figuur 2 geeft een Hasse diagram per armoedemaatstaf weer. Deze diagrammen illustreren de respectievelijke armoedeordeningen, waarbij rekening wordt gehouden met statistische significantie. Hoger geplaatste jaartallen duiden op een hoger armoedeniveau en de lijnen geven de statistisch significante relaties weer. Indien er zich geen lijn bevindt tussen twee jaartallen, wijst dit op een niet statistisch significant verschil in armoedeniveau.

Bekijken we eerst het Hasse diagram voor de headcount. Desondanks het feit dat het proportioneel aantal armen in 1997 ten minste 1.2 percent punt hoger is dan in de voorgaande jaren, blijkt dat deze verschillen niet statistisch significant zijn.
Voor zowel de inkomenskloof ratio als de Gini coëfficiënt bekomen we een verschillend beeld. Alhoewel het proportioneel aantal armen niet significant wijzigt, blijkt dat de procentuele afstand tussen de armoedegrens en het gemiddeld inkomen

van de armen enerzijds en de inkomensongelijkheid onder de armen anderzijds, hoger was in 1992 dan in 1985, 1988 (en 1997).

Tenslotte, welke evolutie vinden we terug voor de axiomatische armoedemaatstaven, die de drie aspecten van armoede incorporeren? Alle geven aan dat het armoedeniveau in 1992 hoger was dan in 1988. Met betrekking tot 1985 en 1997 vinden we minder éénduidigheid. In geen van de drie gevallen kan men een onderscheid maken tussen 1992 en 1997 en tussen 1985 en 1988.

Wat kan men nu besluiten uit het bovenstaande? Ten eerste, op basis van het proportioneel aantal armen komt men tot de conclusie dat het armoedeniveau constant bleef tussen 1985 en 1997.

Ten tweede, het procentueel verschil tussen de armoedegrens en het gemiddeld inkomen van de armen nam toe tussen 1988 en 1992. De inkomensongelijkheid onder hen nam eveneens toe tussen 1988 en 1992, maar daalde na 1992. Indien men deze aspecten belangrijk vindt, zijn de inkomenskloof ratio en de Gini coëfficiënt een nuttige aanvulling op de headcount.

Ten derde, aan de hand van de axiomatische armoedemaatstaven kan men enkel concluderen dat 1988 het laagste armoedeniveau kende en 1992 het hoogste. De andere relaties zijn niet in alle gevallen significant en lenen zich daarom niet tot éénduidige gevolgtrekkingen. Toch lijkt het merendeel van de resultaten aan te duiden dat 1985 en 1988 jaren van een lager en 1992 en 1997 jaren van hoger armoedeniveau zijn.

B. Genormaliseerde invers gegeneraliseerde Lorenz curven

In figuur 3 zijn de resultaten voor de *genormaliseerde* IGL-curven opgenomen. Bemerk dat ik opteerde voor de genormaliseerde curven aangezien elk jaar een andere armoedegrens kende. Bovendien ga ik in deze figuur nog niet in op statistische significantie.

[figuur 3]

De drie voordelen van de curven worden geïllustreerd in figuur 3. Ten eerste kruist geen van de curven. Met andere woorden, het armoedeniveau (aan de hand van alle GAA armoedemaatstaven) in 1992 was hoger dan dat in 1997 en in dalende lijn volgen 1985 en 1988. Dit stemt overeen met de armoedeordening uit tabel 2. Ten tweede kunnen we de drie aspecten van armoede afleiden uit de figuur. De headcount lezen we af op de horizontale as, de inkomenskloof ratio op de helling die de oorspong verbindt met het punt ($q/n, \sum_{i=1}^{q} r_i/n$) en de inkomensongelijkheid onder de armen komt overeen met de concaviteit van de IGL-curve. Men zou dus kunnen stellen dat de curven een synthese weergeven van de werkwijze in de vorige paragraaf. Zowel ad hoc als axiomatische maatstaven kunnen immers afgeleid worden uit de curven. Ten derde kunnen we nagaan hoeveel de armoedegrens van het dominerende jaar mag dalen zonder dat de IGL-dominantie wordt aangetast. Men bekomt als dusdanig het bereik van de armoedegrens waarvoor IGL-dominantie geldt.[12] Tabel 3 geeft de resultaten van deze methodologie.

[tabel 3]

In de eerste kolom worden de dominerende (hoogste armoedeniveau) en de gedomineerde (laagste armoedeniveau) jaren weergegeven. De armoedegrenzen van de gedomineerde jaren worden vermeld in de tweede kolom. Deze zijn steeds uitgedrukt in prijzen van het dominerend jaar omdat ze vergeleken dienen te worden met het bereik van de armoedegrens van het dominerend jaar. Dit bereik wordt afgebeeld in de laatste kolom en geeft weer hoeveel de armoedegrens van het dominerend jaar mag dalen zonder dat de IGL-dominantie wordt aangetast. Een vergelijking van de tweede en derde kolom geeft dan een indicatie van het belang van de hoogte van de armoedegrenzen.

Indien de grens in de tweede kolom lager is dan de ondergrens van het bereik in de derde kolom, kan men besluiten dat de bevinding van IGL-dominantie niet langer opgaat indien de armoedegrens van het dominerend jaar gelijk is aan diegene van het gedomineerd jaar. Met andere woorden, de hogere armoedegrens van het dominerend

[12] Voor deze methodologie wordt geen statistische significantie nagegaan. Dat laat ik voor verder onderzoek.

jaar is een *noodzakelijk* voorwaarde voor de bevinding van IGL-dominantie. Bekijken we bijvoorbeeld 1992 versus 1985. Op basis van de IGL-curven uit figuur 3 kan men afleiden dat 1992 gekenmerkt wordt door een hoger armoedeniveau. De vraag stelt zich nu of dit niet te wijten is aan het verschil in de hoogte van de armoedegrenzen. In 1985 bedroeg de armoedegrens 15039 BEF (prijzen van 92) en 17217 BEF in 1992. De grens van 1992 mag dalen tot 15825 BEF zonder de IGL-dominantie aan te tasten. Met andere woorden, mochten we de armoedegrens van 1985 (in prijzen van 1992), met name 15039 BEF, toepassen op de gegevens van 1992, dan zouden de IGL-curven snijden en zou men geen éénduidige uitspraak kunnen doen over de evolutie van het armoedeniveau. Dezelfde redenering kan toegepast worden op de vergelijkingen tussen 1997 en 1988, 1997 en 1985.

Indien de grens in de tweede kolom behoort tot of hoger is dan de bovengrens van het bereik in de derde kolom, kan men besluiten dat de hoogte van de armoedegrens van het dominerend jaar minder determinerend is voor de bevinding van IGL-dominantie. De vergelijking tussen 1988 en 1985 is hiervan een voorbeeld. De armoedegrens van 1988 is immers hoger dan de armoedegrens van 1985. Dezelfde redenering kan toegepast worden voor 1992 versus 1988, en 1997 versus 1992.

Aan de hand van tabel 3 trachtte ik na te gaan in hoeverre IGL-dominantie te wijten is aan verschillen in de armoedegrenzen. Het bleek dat voor alle vergelijkingen, behalve voor 1988 versus 1985, 1992 versus 1988, en 1997 versus 1992, het verschil in armoedegrens determinerend is voor de bevinding van IGL-dominantie. Met andere woorden, het toepassen van welvaartsvaste armoedegrenzen (behalve voor de drie bovenvernoemde vergelijkingen) zou resulteren in een snijding van de IGL-curven.

Tot slot zou ik willen ingaan op het belang van statistische significantie. Wederom gebruik ik de bootstrap methode (zie appendix A, B & C) om betrouwbaarheidsregio's voor de IGL-curven te construeren. In de appendix zijn de originele testen opgenomen (figuur A.1 – A.6). Figuur 4 geeft een Hasse diagram weer, waarbij rekening is gehouden met statistische significantie.

[figuur 4]

Het blijkt dat de armoedeordening op basis van figuur 3 onvoldoende is om robuuste conclusies te trekken. 1992 wordt nog steeds gekenmerkt door een statistisch significant hoger armoedeniveau dan 1988, maar is niet langer verschillend van 1985 en 1997. Ook over de relatie tussen 1985 en 1997 kunnen de IGL-curven geen uitspraak doen.
Een vergelijking van dit Hasse diagram met de Hasse diagrammen voor de armoedemaatstaven doet besluiten dat de belangrijkste conclusies onveranderd blijven. De armoedeordening van de headcount verschilt wezenlijk van de ordening van de IGL-curven. Bovendien wordt nog steeds 1992 gekenmerkt door het hoogste armoedeniveau en 1988 door het laagste.

De conclusie van deze paragraaf kan kort samengevat worden. Ten eerste, geven de IGL-curven dezelfde tendens weer als de armoedemaatstaven, hoewel de IGL-ordening inherent partieel is. Ten tweede verduidelijkt tabel 3 dat de keuze voor een welbepaalde armoedegrens de resultaten aanzienlijk kan beïnvloeden. Alhoewel Sen benadrukte dat armoedemeting bestaat uit twee gescheiden denkoefeningen[13], blijken beide een belangrijke invloed uit te oefenen op de uiteindelijke resultaten.

Besluit

In dit artikel trachtte ik een antwoord te formuleren op twee onderzoeksvragen. Ten eerste, nuanceren de IGL-curven en de axiomatische armoedemaatstaven het verkregen beeld van armoede of bevestigen ze veeleer de resultaten van de headcount? Ten tweede focuste ik op het belang van statistische significantie. Ik gebruikte daarvoor de bootstrap methode.

De gebruikte armoedemaatstaven zijn de headcount, de inkomenskloof ratio, de Gini coëfficiënt voor de armen, de Sen, de Foster, Greer Thorbecke en de Watts maatstaf. De eerste drie catalogiseerde ik onder de noemer ad hoc. Geen van de drie houdt immers tegelijkertijd rekening met het aantal armen, hun gemiddeld inkomen en de

[13] "(1) *identification* of the poor, and (2) *aggregation* of the statistics regarding the identified poor to derive an overall index of poverty" (Sen, 1992, blz. 102).

verdeling van armoede. Ze zijn nuttig voor zover men enkel één van de drie aspecten wenst te becijferen. Wil men daarentegen de drie aspecten combineren, dan moet men gebruik maken van de laatste drie maatstaven, namelijk de axiomatische.
IGL-curven zijn een andere manier om het armoedeniveau te bestuderen. Zij vatten het armoedeniveau niet samen in één getal, maar beschouwen de gehele 'arme' inkomensverdeling. Drie voordelen zijn inherent verbonden aan deze methodologie. Ten eerste kan men voor een uitgebreide groep van axiomatische armoedemaatstaven een ondubbelzinnige uitspraak doen in een vergelijkende studie naar armoede. Ten tweede, geven de curven een visuele voorstelling van armoede. Ten derde kan men nagaan in hoeverre de hoogte van de armoedegrens determinerend is voor de resultaten.

In het empirisch onderzoek gebruikte ik Belgische microdata voor 1985, 1988, 1992 en 1997, verzameld door het CSB. De conclusies kunnen als volgt samengevat worden.
Ten eerste, het proportioneel aantal armen bleef nagenoeg constant tussen 1985 en 1997. Bijna alle andere maatstaven (ook ad hoc) daarentegen wezen op een significante stijging van het armoedeniveau tussen 1985 en 1992. De resultaten zijn echter minder éénduidig omtrent de evolutie na 1992. Toch komt duidelijk naar voren dat de headcount een andere armoedeordening aangeeft dan de andere armoedemaatstaven. Daarom lijkt het aangewezen om in toekomstig toegepast onderzoek naar armoede verschillende armoedemaatstaven te gebruiken. Op die wijze krijgen de beleidsmakers weliswaar een grotere hoeveelheid informatie, maar zullen ze wel de keuze hebben om hun beleid te evalueren aan de hand van verschillende aspecten van armoede.
Ten tweede, de IGL-curven zijn niet alleen theoretisch, maar ook empirisch een zinvol alternatief voor de verschillende armoedemaatstaven. Ze zijn eenvoudig te berekenen en laten toe het armoedeniveau te evalueren aan de hand van een uitgebreide klasse van armoedemaatstaven, zonder de maatstaven effectief te berekenen. De resultaten van de IGL-curven gaven dezelfde significante tendens weer als de armoedemaatstaven.

Bibliografie

Atkinson, A.B. (1987), “On the measurement of poverty”, Econometrica, vol. 55, n° 4, blz. 749-764.

Borooah, V.K. (1991), "Problems in the measurement of inequality and poverty", Indian economic journal, vol. 38, n° 4, blz. 12-38.

Cantillon, B., M. Marx, D. Proost & R. Van Dam (1993), "Sociale indicatoren: 1985-1992", Berichten Centrum voor Sociaal Beleid, Antwerpen, UFSIA (Universiteit Antwerpen), 51 blz.

Cantillon, B., M. Andries, B. Meulemans & B. Tan (1996), "Twintig jaar armoede en beleid inzake armoedebestrijding", Economisch en sociaal tijdschrift, jg. 50, n° 1, blz. 5-36.

Cantillon, B., L. De Lathouwer, I. Marx, R. Van Dam & K. Van den Bosch (1999), "Sociale Indicatoren 1976-1997", Berichten Centrum voor Sociaal Beleid, Antwerpen, UFSIA (Universiteit Antwerpen), 49 blz.

Chakravarty, S.R. (1983), “Ethically flexible measures of poverty”, Canadian journal of economics, vol. 16, blz. 74-85.

Clark, S., R. Hemming & D. Ulph (1981), “On indices for the measurement of poverty”, Economic journal, vol. 91, blz. 515-526.

Efron, B. & R.J. Tibshirani (1993), An introduction to the bootstrap, New York/London, Chapman and Hall, 436 blz.

Foster, J.E., J. Greer, E. Thorbecke (1984), “A class of decomposable poverty indices”, Econometrica, vol. 52, blz. 761-766.

Foster, J.E. & A.F. Shorrocks (1988), “Poverty orderings and welfare dominance”, Social choice and welfare, vol. 5, blz. 179-198.

Hagenaars, A.J.M. (1987), “A class of poverty indices”, International economic review, vol. 28, blz. 583-607.

Jenkins, S.P. & P.J. Lambert (1997), "Three 'I' 's of poverty curves, with an analysis of UK poverty trends", Oxford economic papers, vol. 49, blz. 317-327.

Jenkins, S.P. & P.J. Lambert (1998), “Three ‘I’ ’s of poverty curves and poverty dominance: tips for poverty analysis”, Research on economic inequality, vol. 8, blz. 39-56.

Johnson, D.T. (1988), “The measurement of poverty in Australia: 1981-1982 and 1985-1986”, Australian economic review, vol. 3, blz. 13-24.

Kalton, G. (1983), “Introduction to survey sampling”, Sage university paper series, n° 07-035, Beverly Hills/London/New Delhi, Sage publications, 96 blz.

Mills, J.A., S. Zandvakili (1995), “Statistical inference via bootstrapping for measures of inequality"” The Jerome Levy economics institute working paper series, n° 136, 30 blz.

Mills, J.A., S. Zandvakili (1997), "Statistical inference via bootstrapping for measures of inequality", Journal of applied econometrics, vol. 12, blz. 133-150.

Nationale Bank van België (2000), Indexcijfers van de consumptieprijzen – historisch overzicht, http://www.nbb.be/sdq/N/TAB/Dom2127/ORG/6W3DU01.HTM

Pyatt, G. (1987), “Measuring welfare, poverty and inequality”, Economic journal, vol. 97, blz. 459-467.

Seidl, C. (1986), "Poverty measurement: a survey", in: D. Bös, M. Rose & C. Seidl, eds., Welfare and efficiency in public economics, Berlin, Springer, 1986, blz. 71-147.

Sen, A. (1976), "Poverty: an ordinal approach to measurement", Econometrica, vol. 44, n° 2, blz. 219-231.

Sen, A. (1992), Inequality reexamined, Oxford, Clarendon press, 207 blz.

Shorrocks, A.F. (1995), “Revisiting the Sen poverty index”, Econometrica, vol. 63, blz. 1225-1230.

The MathWorks (1999), Matlab function reference (volume 1: language), version 5, 884 blz.

Thon, D. (1983), “A note on a troublesome axiom for poverty indices”, Economic journal, vol. 93, blz. 199-200.

Van den Bosch, K., T. Callan, J. Estivill, P. Hausman, B. Jeandidier, R. Muffels & J. Yfantopoulos (1993), “A comparison of poverty in seven European countries and regions using subjective and relative measures”, Journal of population economics, vol. 6, blz. 235-259.

Van den Bosch, K. (1999), Identifying the Poor, Using Subjective and Consensual Measures, doctoraal proefschrift, UFSIA (Universiteit Antwerpen), 672 blz.

Watts, H. (1968), “An economic definition of poverty”, in: D.P. Moynihan, eds., On understanding poverty: perspectives from the social sciences, New York, Basic books, 425 blz.

Zheng, B. (1993), Poverty measurement, statistical inference, and an application to the United States, Phd. thesis, West Virginia University, 168 blz.

Zheng, B. (1997), "Aggregate poverty measures", Journal of economic surveys, vol. 11, n° 2, blz. 123-162.

Appendix

In deze appendix wordt de methodologie voor het toetsen van hypothesen besproken. In de eerste paragraaf beschrijf ik de bootstrap methode voor het toetsen van hypothesen binnen *één* steekproef. Deze methode veronderstelt dat de observaties in de steekproef onafhankelijk zijn en dat de steekproef zelf een zuivere toevalssteekproef is. Geen van beide assumpties is echter van toepassing op de gebruikte gegevens. Op voorwaarde dat men de nodige aanpassingen doorvoert, kan men de bootstrap methode nog steeds gebruiken. Deze aanpassingen worden besproken in de tweede paragraaf. Tenslotte is het niet de bedoeling om statistische significantie na te gaan binnen *één* steekproef, maar over *verschillende* steekproeven. De bootstrap methode beschreven in de eerste twee paragrafen kan enkel toegepast worden indien de verschillende steekproeven onafhankelijk zijn. Het 'panel-karakter' van de data (dezelfde huishoudens worden in elk van de jaren bevraagd) wijst noodzakelijkerwijs op correlatie. Twee mogelijke werkwijzen om hieraan tegemoet te komen, worden beschreven in de laatste paragraaf.

A. De bootstrap methode[14]

In deze paragraaf beschrijf ik eerst de bootstrap methode voor het toetsen van hypothesen binnen *één* steekproef (hier: binnen één jaar). Nadien worden de voordelen van de bootstrap aangehaald.

De bootstrap laat toe de verdeling van een statistiek (hier: de armoedemaatstaven en de IGL-curven) te benaderen. Men gebruikt hiervoor geen asymptotische theorie (zoals gebruikelijk is binnen de klassiek statistische theorie), maar eerder simulatiemethodes. De procedure bestaat uit vier stappen.
In een eerste stap observeert men een steekproef X met n observaties x_i (in dit artikel is X de representatieve steekproef verzameld door het CSB en x_i het individueel equivalent inkomen). In de veronderstelling van onafhankelijke

[14] Voor een uitgebreid overzicht, zie Efron & Tibshirani (1993).

observaties en een zuivere toevalssteekproef, kan men zonder enige restrictie een uniforme dichtheidsfunctie opleggen, met name elke observatie heeft een kans $1/n$.

In een tweede stap, zal men uit X een nieuwe toevalssteekproef trekken, met teruglegging, waarbij elke observatie een kans op trekking heeft van $1/n$. Men bekomt dan de bootstrap-steekproef X^1 met n observaties x_i^1.

In de derde stap berekent men eerst de armoedegrens voor X^1 (de armoedegrens voor X^1 hoeft immers niet dezelfde te zijn als deze voor X en verwaarlozing hiervan zal leiden tot een overschatting van statistische significantie)[15] en nadien de statistiek voor X^1. De statistiek is één waarde in het geval van de armoedemaatstaven en één curve in het geval van de IGL-curven.

De vierde stap bestaat uit s replicaties van stap twee en drie. Zodoende bekomt men voor de armoedemaatstaven s waarden van de statistiek, waarbij de i^{de} waarde gebaseerd is op $X^i, i = 1,2,...,s$. Het histogram van de s waarden is dan een benadering voor de verdeling van de statistiek. Op basis van deze verdeling kan men dan betrouwbaarheidsintervallen berekenen (Mills & Zandvakili, 1995, blz. 5-6).[16] Voor de IGL-curven bekomt men s curven in plaats van s waarden. Op basis van deze s curven kan men een betrouwbaarheidsregio opstellen. Meer bepaald heeft men voor elk punt $k/n \in [0,1]$ op de horizontale as s overeenkomstige waarden, met name $IGL\left(r^i, k/n\right), i = 1,2,...,s$. Het histogram van deze s waarden kan gebruikt worden om een betrouwbaarheidsinterval (voor elk punt k/n) te berekenen. Al deze betrouwbaarheidsintervallen samen vormen dan de betrouwbaarheidsregio.[17]

In de toepassing van bovenstaande procedure heb ik gekozen voor 1000 replicaties (s=1000).[18] De betrouwbaarheidsintervallen werden berekend aan de hand van de percentielenmethode. Deze methode gebruikt de percentielen van een verdeling om

[15] Indien men zou opteren voor een absolute grens, moet men de grens niet herberekenen.

[16] Het is belangrijk op te merken dat deze betrouwbaarheidsintervallen enkel gebruikt kunnen worden voor het toetsen van hypothesen binnen *één* steekproef en niet over *verschillende* steekproeven. Zo kan men bijvoorbeeld nagaan of de Sen maatstaf voor 1992 statistisch verschillend is van nul, maar kan men op basis van de beschreven procedure niet nagaan of de Sen maatstaf statistisch verschillend is van deze voor 1985. In de derde paragraaf wordt de procedure hiervoor aangepast.

[17] Een bijkomend probleem bij het opstellen van een betrouwbaarheidsregio betreft de gewichten. Aangezien de steekproeven gewogen zijn, is elk van de s IGL-curven gedefinieerd voor andere punten op de horizontale as. Daarom was het noodzakelijk de s IGL-curven te interpoleren naar dezelfde waarden op de horizontale as. Ik gebruikte hiervoor cubic spline interpolatie (The MathWorks, 1999, blz. 434-438).

[18] De verdeling van de Sen index werd eveneens benaderd voor s=2000. De resulterende verdeling verschilde nauwelijks van deze op basis van s=1000.

een betrouwbaarheidsinterval op te stellen en is superieur (in vergelijking met intervallen gebaseerd op standaardfouten) indien de verdeling van de statistiek scheef is (Efron & Tibshirani, 1993, blz. 168-177).

De bootstrap methode heeft verscheidene voordelen in vergelijking met de klassieke statistische methodologie.

Ten eerste bestaat er nauwelijks asymptotische theorie omtrent de verdelingen van de armoedemaatstaven. Zulke theorie is extreem moeilijk te ontwikkelen aangezien de maatstaven meestal niet-lineaire functies van het inkomen zijn. De bootstrap daarentegen kan men toepassen op elke maatstaf.

Ten tweede houdt de bootstrap rekening met het theoretisch bereik van de armoedemaatstaven, met name $[0,1]$. Aangezien de verdeling van de armoedemaatstaven wordt gesimuleerd, zullen de betrouwbaarheidsintervallen steeds binnen het theoretisch bereik liggen. Dit is niet noodzakelijk het geval voor de klassieke statistische methodologie aangezien deze gebaseerd is op asymptotische theorie.

Ten derde blijkt uit Monte Carlo simulaties en uit empirische toepassingen dat de bootstrap methode superieur is aan klassieke statistische methoden (Mills & Zandvakili, 1997, blz. 136-137).

Ten vierde houdt de bootstrap automatisch rekening met de variatie in de armoedegrens (zie stap 3 in de procedure).

Tenslotte kan men de methode aanpassen zodat ze rekening houdt met bijkomende complicaties, zoals afhankelijke observaties, enz. (zie paragraaf B & C).

B. Afhankelijke observaties en geclusterde steekproef

De procedure in paragraaf A kan gebruikt worden bij het toetsen van hypothesen binnen *één* steekproef indien de observaties onafhankelijk zijn en indien de steekproef een zuivere toevalssteekproef is. Aangezien de gebruikte gegevens niet voldoen aan beide assumpties, moet men enkele aanpassingen doorvoeren.

Ten eerste zijn de observaties uit de steekproef van het CSB afhankelijk. De verschillende maatstaven en IGL-curven worden op individueel niveau berekend terwijl het inkomen enkel op het niveau van het huishouden kan worden vastgesteld.

Met andere woorden, alle individuen binnen één huishouden hebben hetzelfde equivalent inkomen. Indien men bij de bootstrap geen rekening houdt met deze afhankelijkheid zullen de betrouwbaarheidsintervallen onderschat worden. Dit probleem kan echter ondervangen worden door huishoudens in plaats van individuen te trekken (stap 2 in paragraaf A).

Ten tweede is de steekproef van het CSB geen zuivere toevalssteekproef. Het betreft een binnen gemeenten geclusterde steekproef waarbij de gemeenten zijn gestratificeerd. Implementatie van de stratificatie resulteert in enorme programmeertechnische problemen en wordt daarom niet geïncorporeerd. Toch is het onwaarschijnlijk dat deze verwaarlozing serieuze gevolgen met zich meebrengt. De betrouwbaarheidsintervallen zijn hooguit overschat. Bovendien blijkt uit vorig onderzoek dat de overschatting eerder laag is (Van den Bosch, 1999, blz. 385). Het 'cluster-aspect' van de steekproef wordt wel geïncorporeerd. Een verwaarlozing hiervan zou leiden tot een onderschatting van de betrouwbaarheidsintervallen. Deze onderschatting is in absolute termen veel belangrijker dan de overschatting omwille van stratificatie (Van den Bosch, 1999, blz. 385). Om dit te ondervangen, maakte ik gebruik van de 'ultimate cluster approximation' (Kalton, 1983, blz. 35). Deze benadering verwaarloost de variantie van de huishoudens binnen de clusters. Gegeven de steekproefopzet is deze verwaarlozing onbelangrijk (Van den Bosch, 1999, blz. 391). De 'ultimate cluster approximation' houdt concreet in dat men clusters in plaats van huishoudens trekt (stap 2 in paragraaf A). Aangezien een cluster een groep van huishoudens vormt, wordt eveneens gecorrigeerd voor de afhankelijkheid van de observaties binnen één huishouden.

C. Panelgegevens

De bootstrap procedure beschreven in paragraaf A en B is enkel geschikt om hypothesen te toetsen binnen één steekproef. Toch kan men op een eenvoudige wijze hypothesen toetsen over *verschillende* steekproeven indien de steekproeven onafhankelijk zijn. Men past dan de bootstrap procedure toe op de verschillende steekproeven afzonderlijk en de resulterende verdelingen kunnen gebruikt worden om een test uit te voeren op verschillen tussen armoedemaatstaven.

Indien de steekproeven daarentegen gecorreleerd zijn, kan men bovenvernoemde test niet langer gebruiken aangezien dit zou resulteren in een overschatting van de betrouwbaarheidsintervallen. Aangezien panelgegevens steeds gecorreleerd zijn over de jaren heen, moeten alternatieve procedures gebruikt worden. In dit artikel gebruik ik twee verschillende werkwijzen. De eerste werkwijze werd ontwikkeld door Mills en Zandvakili (1997). De tweede werkwijze is gebaseerd op gepaarde waarnemingen.

De methode van Mills en Zandvakili (1997) bestaat uit twee stappen.[19] In een eerste stap berekent men een armoedemaatstaf of IGL-curve afzonderlijk voor elk jaar op basis van de originele steekproeven en past men de bootstrap procedure toe (opnieuw afzonderlijk voor elk jaar) om de verdeling van de statistiek (zijnde maatstaf of curve) te bekomen. In een tweede stap berekent men het verschil tussen de maatstaven (maatstaf in jaar 1 minus maatstaf in jaar 2) of curven op basis van de originele steekproeven. De resulterende waarde of curve is de teststatistiek. De verdeling van deze teststatistiek wordt dan bekomen door het verschil te nemen van de afzonderlijke verdelingen, berekend in stap 1. Op basis van de resulterende verdeling kan men een betrouwbaarheidsinterval of -regio berekenen.

De methode gebaseerd op de gepaarde waarnemingen bestaat uit één stap. De teststatistiek zelf wordt berekend zoals in de methode van Mills en Zandvakili. De verdeling van de teststatistiek wordt bekomen door de teststatistiek zelf te bootstrappen in plaats van deze in twee stappen te berekenen. Meer bepaald worden in elke trekking in de bootstrap procedure (stap 2 in paragraaf A) dezelfde clusters getrokken uit de verschillende steekproeven. Op deze wijze wordt immers rekening gehouden met de correlatie tussen de steekproeven. Een voorbeeld: stel dat men 1988 en 1985 wil vergelijken aan de hand van de Sen maatstaf. De teststatistiek bekomt men door het verschil te nemen van de Sen maatstaven van 1988 en 1985. De verdeling van deze teststatistiek bekomt men door paarsgewijze trekking in de bootstrap procedure. Meer bepaald trekt men in stap 2 (paragraaf A) dezelfde clusters in 1985 en 1988. In stap 3 berekent men dan het verschil tussen de Sen maatstaven

[19] Deze werkwijze is enkel geldig onder bepaalde assumpties. Voor een uitgebreide discussie, zie Mills & Zandvakili (97).

van 1988 en 1985. Voldoende replicaties van deze procedure leveren dan de verdeling van de teststatistiek.

De Hasse diagrammen in de tekst (figuur 2 en 4) werden opgesteld volgens de methode van gepaarde waarnemingen. Deze methode werd verkozen omdat ze minder assumpties vereist (zie voetnoot 19). De resultaten van de methode van Mills en Zandvakili worden weergegeven in onderstaande tabellen en figuren. Voor figuur 2 is dit onbelangrijk aangezien beide methodes tot dezelfde conclusies leiden. Voor figuur 4 is dit wel belangrijk, aangezien beide methodes tot verschillende resultaten leiden.

In tabel A.1 worden de teststatistieken en betrouwbaarheidsintervallen van het verschil tussen de armoedegrenzen weergegeven. Indien nul tot het betrouwbaarheidsinterval behoort, is het verschil tussen de armoedegrenzen niet statistisch significant.

[tabel A.1]

In tabel A.2. worden de teststatistieken en betrouwbaarheidsintervallen van het verschil tussen de armoedemaatstaven weergegeven. Indien nul tot het betrouwbaarheidsinterval behoort, is het verschil tussen de armoedemaatstaven niet statistisch significant.

[tabel A.2]

In figuur A.1-A.6 worden de teststatistieken en betrouwbaarheidsregio’s van het verschil tussen de IGL-curven weergegeven. De stippellijnen duiden de boven- en ondergrenzen van de betrouwbaarheidsregio’s aan. De volle lijn is de test statistiek, met name het verschil tussen de IGL-curven. Indien de vette horizontale lijn tot de betrouwbaarheidsregio behoort, is het verschil tussen de IGL-curven niet statistisch significant.

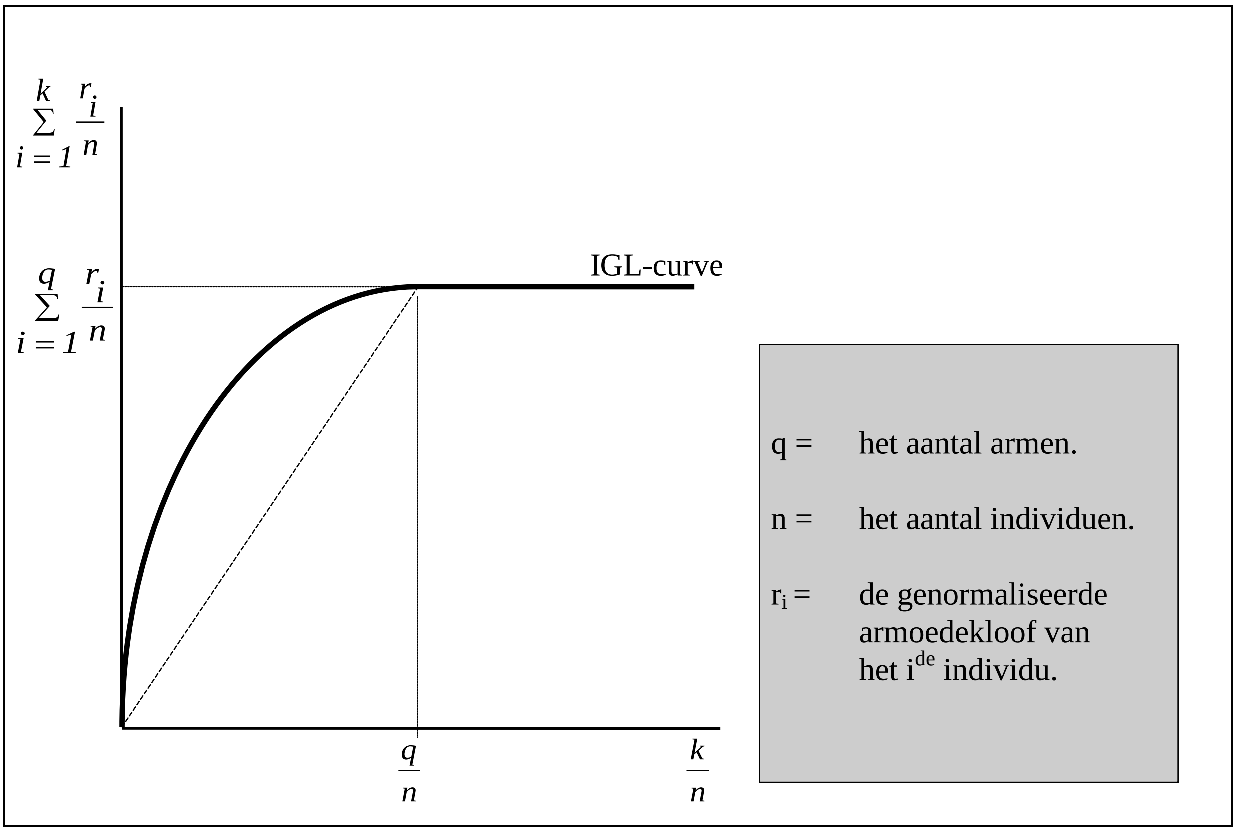
[figuur A.1-A.6]

Abstract

Poverty measures, IGL-curves and the bootstrap: an empirical study of the evolution of the poverty level in Belgium between 1985 and 1997

The headcount ratio, measuring the number of poor as a percentage of the population, is the most popular poverty measure in empirical research. In the seventies however, A. Sen, noticed that other characteristics of poverty can be as important as the number of poor, i.e. the average income of the poor and the income inequality among the poor. In this paper, I compare the headcount ratio with other poverty measures that take into account these other characteristics of poverty. In addition, I focus on statistical inference; are the changes in the Belgian poverty level between 1985 and 1997 significant?

Figuur 1: IGL-curve



Bron: Jenkins, S.P. & P.J. Lambert (1998), blz. 42

Tabel 1: de armoedegrenzen (EG-grens)

	armoedegrens	**armoedegrens (88)**
1985	12820 BEF	13343 BEF
1988	13943 BEF	13943 BEF
1992	17217 BEF	15275 BEF
1997	20140 BEF	16133 BEF

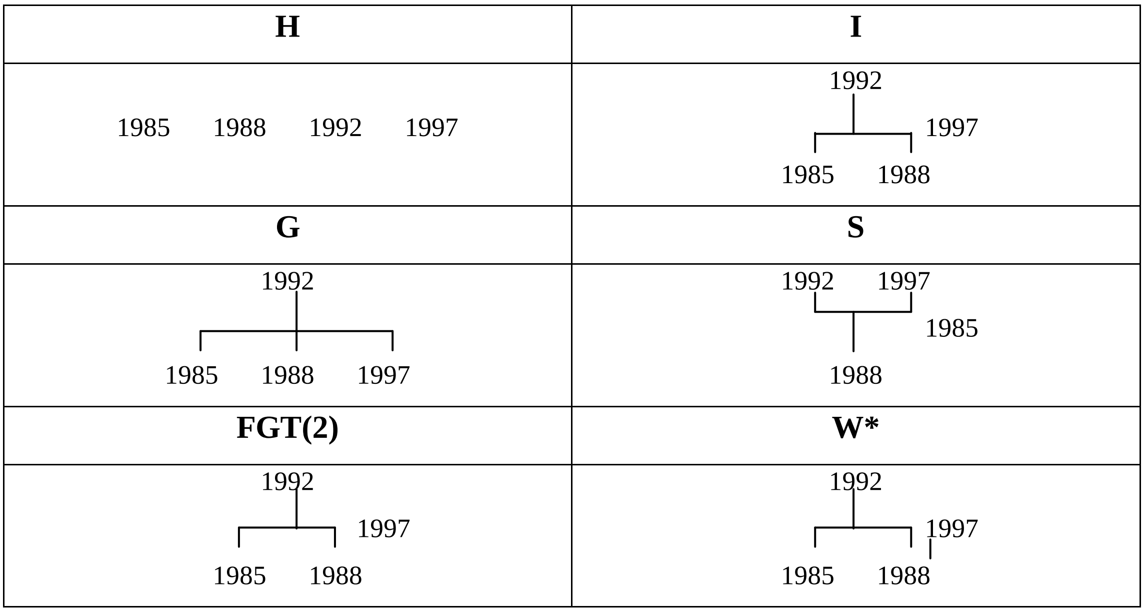
Armoedegrens (88) werd berekend met behulp van de consumptieprijsindex (1988=100), bron: Nationale Bank van België (2000).

Tabel 2: de resultaten van de armoedemaatstaven

	1985	**1988**	**1992**	**1997**
H	0.0563 (3)	0.0518 (1)	0.0558 (2)	0.0687 (4)
I	0.1559 (1)	0.1563 (2)	0.2255 (4)	0.1700 (3)
G	0.0968 (3)	0.0838 (1)	0.1618 (4)	0.0958 (2)
S	0.0132 (2)	0.0114 (1)	0.0191 (4)	0.0168 (3)
FGT(2)	0.0028 (2)	0.0020 (1)	0.0063 (4)	0.0035 (3)
W	0.0110 (2)	0.0095 (1)	0.0302* (4)	0.0169* (3)

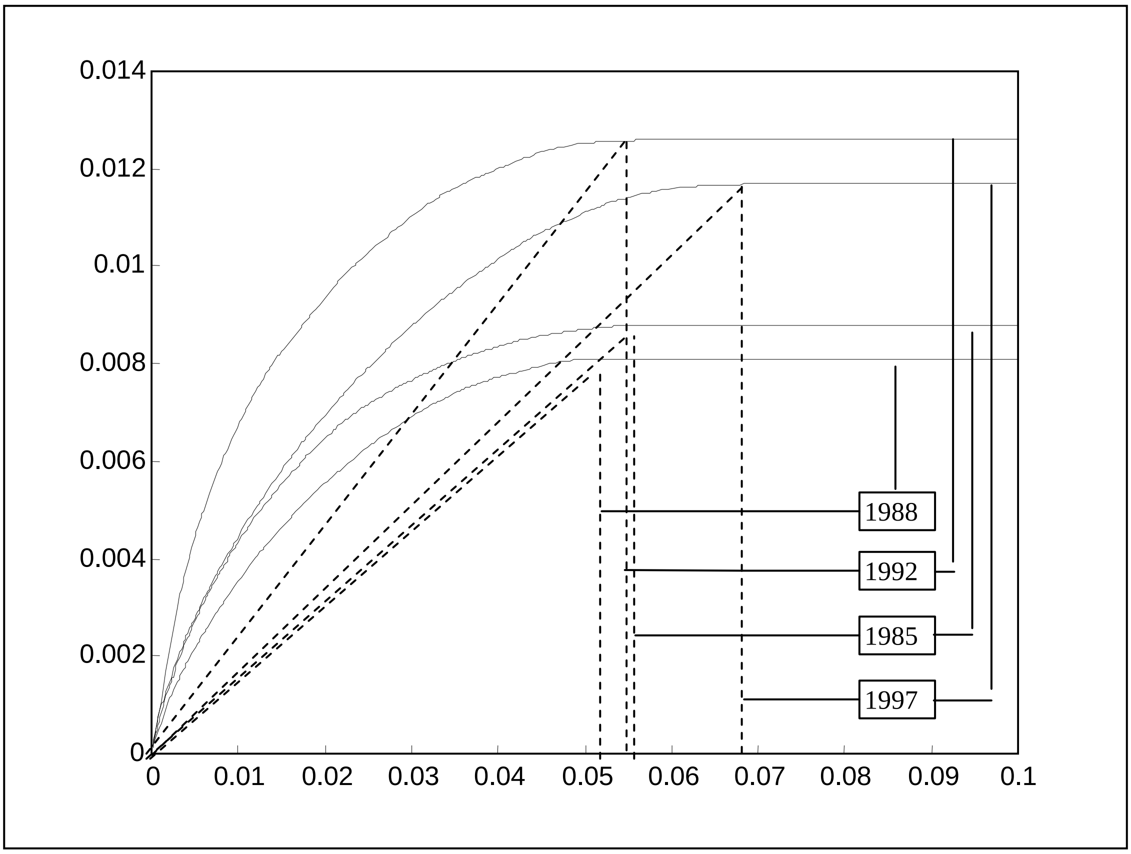
Tussen haken is de armoedeordening weergegeven. Aangezien de steekproef voor 1992 en 1997 nulinkomens bevat werden deze in * vervangen door een inkomen van 1 BEF. Ik berekende de FGT maatstaf voor α=2-5, maar de armoedeordening bleef ongewijzigd. Daarom presenteer ik enkel de resultaten voor α=2.

Figuur 2: statistische significantie van de armoedemaatstaven: Hasse diagrammen



Alle Hasse diagrammen zijn gebaseerd op testen met een significantieniveau van 99 percent. Aangezien de steekproef voor 1992 en 1997 nulinkomens bevat werden deze in * vervangen door een inkomen van 1 BEF. Ik berekende de FGT maatstaf voor α=2-5, maar de significantie van de armoedeordening bleef voor α=2-5 ongewijzigd. Daarom presenteer ik enkel de resultaten voor α=2.

Figuur 3: de genormaliseerde IGL-curven



Tabel 3: IGL-dominantie en de armoedegrens

dominerend jaar (§) vs. gedomineerd jaar (#)	**grens van het gedomineerd jaar (#)**	**bereik grens van het dominerend jaar (§)**
1985 (§) vs. 1988 (#)	13397	12637 – 12820
1992 (§) vs. 1985 (#)	15039	15825 – 17217
1992 (§) vs. 1988 (#)	15715	15597 – 17217
1992 (§) vs. 1997 (#)	18183	16839 – 17217
1997 (§) vs. 1985 (#)	16657	19166 – 20140
1997 (§) vs. 1988 (#)	17407	18948 – 20140

De grens van het gedomineerd jaar is steeds uitgedrukt in prijzen van het dominerend jaar met behulp van de consumptieprijsindex, bron: Nationale Bank van België (2000).

Figuur 4: een Hasse diagram voor de IGL-curven

1992

1985 1997

1988

Het Hasse diagram is gebaseerd op testen met een significantieniveau van 99 percent. De testen werden uitgevoerd met de methode van gepaarde waarnemingen (zie appendix C).

Tabel A.1: testen van het verschil tussen armoedegrenzen (in prijzen van 1988)

	verschil van armoedegrenzen
1988 vs. 1985	600$^{(**)}$ (-146, 1236) [281, 958]
1992 vs. 1985	1933$^{(*), (**)}$ (1293, 2587) [1485, 2308]
1992 vs. 1988	1332$^{(*), (**)}$ (608, 2016) [904, 1797]
1997 vs. 1985	2790$^{(*), (**)}$ (2244, 3341) [2332, 3204]
1997 vs. 1988	2190$^{(*), (**)}$ (1539, 2862) [1676, 2625]
1997 vs. 1992	858$^{(*), (**)}$ (182, 1488) [432, 1349]

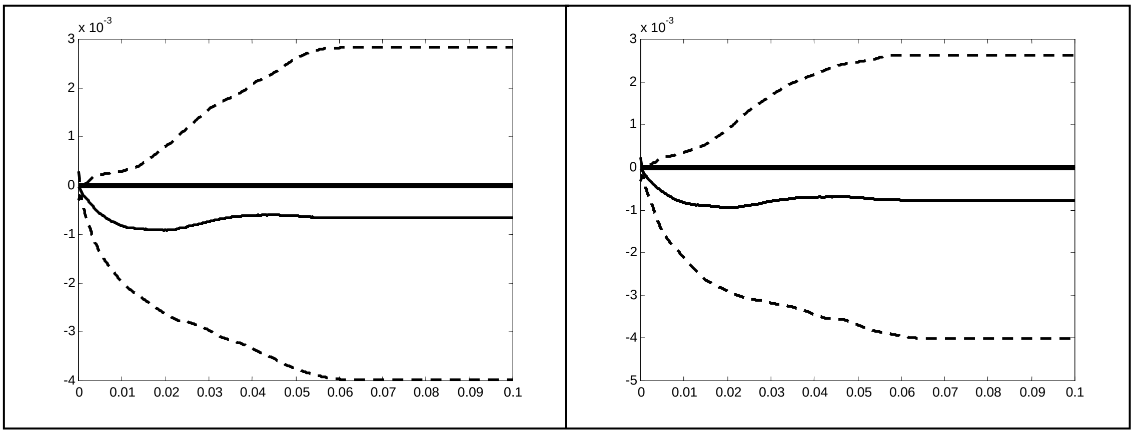
(): 99% betrouwbaarheidsintervallen volgens de methode van Mills & Zandvakili; []: 99% betrouwbaarheidsintervallen volgens de methode van gepaarde waarnemingen; *: significant volgens de methode van Mills & Zandvakili; **: significant volgens de methode van gepaarde waarnemingen. Alle grenzen zijn in prijzen van 1988.

Tabel A.2: testen van het verschil tussen armoedemaatstaven

	H	**I**	**G**	**S**
1988 vs. 1985	-0.0046 (-0.0230, 0.0127) [-0.0181, 0.0118]	0.0004 (-0.0407, 0.0393) [-0.0405, 0.0375]	-0.0130 (-0.0418, 0.0159) [-0.0434, 0.0157]	-0.0018 (-0.0071, 0.0028) [-0.0065, 0.0028]
1992 vs. 1985	-0.0005 (-0.0193, 0.0150) [-0.0163, 0.0168]	0.0696(*), (**) (0.0023, 0.1443) [0.0065, 0.1454]	0.0651(*), (**) (0.0082, 0.1331) [0.0104, 0.1351]	0.0059 (-0.0021, 0.0148) [-0.0015, 0.0155]
1992 vs. 1988	0.0040 (-0.0164, 0.0213) [-0.0132, 0.0184]	0.0692(*), (**) (0.0041, 0.1452) [0.0069, 0.1484]	0.0780(*), (**) (0.0200, 0.1517) [0.0226, 0.1396]	0.0077(*), (**) (0.0008, 0.0161) [0.0005, 0.0156]
1997 vs. 1985	0.0124 (-0.0075, 0.0300) [-0.0026, 0.0277]	0.0141 (-0.0271, 0.0537) [-0.0249, 0.0538]	-0.0010 (-0.0302, 0.0267) [-0.0291, 0.0291]	0.0036 (-0.0021, 0.0094) [-0.0017, 0.0082]
1997 vs. 1988	0.0169 (-0.0017, 0.0370) [-0.0021, 0.0350]	0.0137 (-0.0235, 0.0612) [-0.0277, 0.0602]	0.0120 (-0.0188, 0.0408) [-0.0201, 0.0444]	0.0054(*), (**) (0.0001, 0.0116) [0.0003, 0.0105]
1997 vs. 1992	0.0129 (-0.0066, 0.0337) [-0.0058, 0.0310]	-0.0555 (-0.1323, 0.0162) [-0.1414, 0.0118]	-0.0660(*), (**) (-0.1415, -0.0098) [-0.1382, -0.0130]	-0.0023 (-0.0123, 0.0063) [-0.0133, 0.0058]
	FGT(2)	**W*** **		
1988 vs. 1985	-0.0007 (-0.0024, 0.0006) [-0.0023, 0.0007]	-0.0016 (-0.0063, 0.0027) [-0.0060, 0.0028]		
1992 vs. 1985	0.0035(*), (**) (0.0000, 0.0087) [0.0002, 0.0089]	0.0192(*), (**) (0.0016, 0.0496) [0.0013, 0.0507]		
1992 vs. 1988	0.0042(*), (**) (0.0010, 0.0092) [0.0009, 0.0089]	0.0208(*), (**) (0.0035, 0.0494) [0.0035, 0.0491]		
1997 vs. 1985	0.0008 (-0.0010, 0.0028) [-0.0010, 0.0026]	0.0059 (-0.0009, 0.0148) [-0.0007, 0.0140]		
1997 vs. 1988	0.0015 (-0.0002, 0.0034) [-0.0001, 0.0032]	0.0075(*), (**) (0.0008, 0.0166) [0.0006, 0.0155]		
1997 vs. 1992	-0.0027 (-0.0078, 0.0011) [-0.0083, 0.0011]	-0.0133 (-0.0426, 0.0059) [-0.0472, 0.0072]		

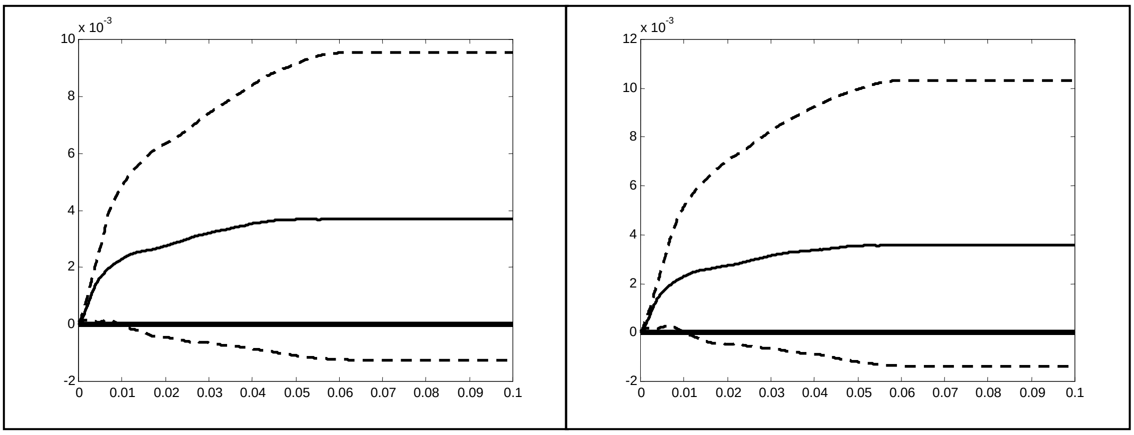
(): 99% betrouwbaarheidsintervallen volgens de methode van Mills en Zandvakili; []: 99% betrouwbaarheidsintervallen volgens de methode van gepaarde waarnemingen; *: significant volgens de methode van Mills en Zandvakili; **: significant volgens de methode van gepaarde waarnemingen; aangezien de steekproef voor 1992 en 1997 nulinkomens bevat werden deze in *** vervangen door een inkomen van 1 BEF.

Figuur A.1: testen van het verschil tussen IGL-curven: 1988 vs. 1985 (links: methode van Mills & Zandvakili, rechts: methode volgens gepaarde waarnemingen)



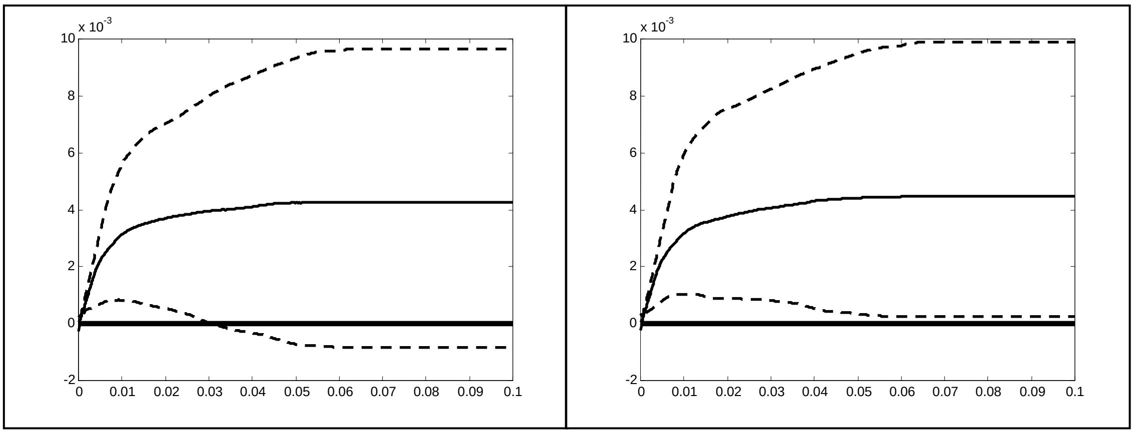
significantieniveau van 99 percent

Figuur A.2: testen van het verschil tussen IGL-curven: 1992 vs. 1985 (links: methode van Mills & Zandvakili, rechts: methode volgens gepaarde waarnemingen)



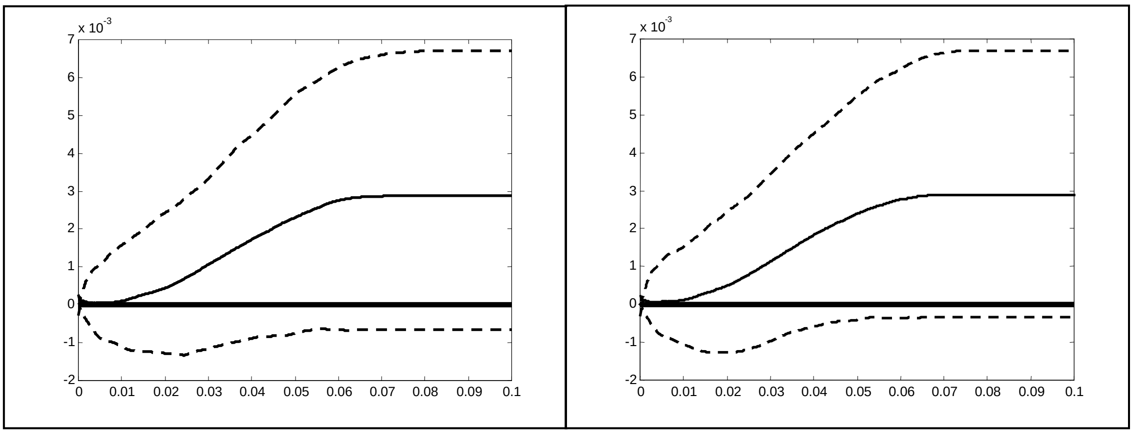
significantieniveau van 99 percent

Figuur A.3: testen van het verschil tussen IGL-curven: 1992 vs. 1988 (links: methode van Mills & Zandvakili, rechts: methode volgens gepaarde waarnemingen)



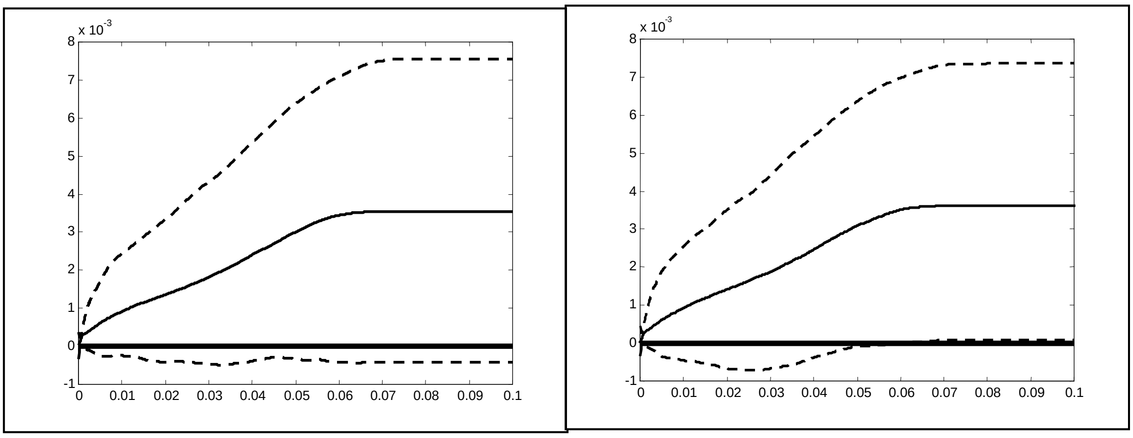
significantieniveau van 99 percent

Figuur A.4: testen van het verschil tussen IGL-curven: 1997 vs. 1985 (links: methode van Mills & Zandvakili, rechts: methode volgens gepaarde waarnemingen)



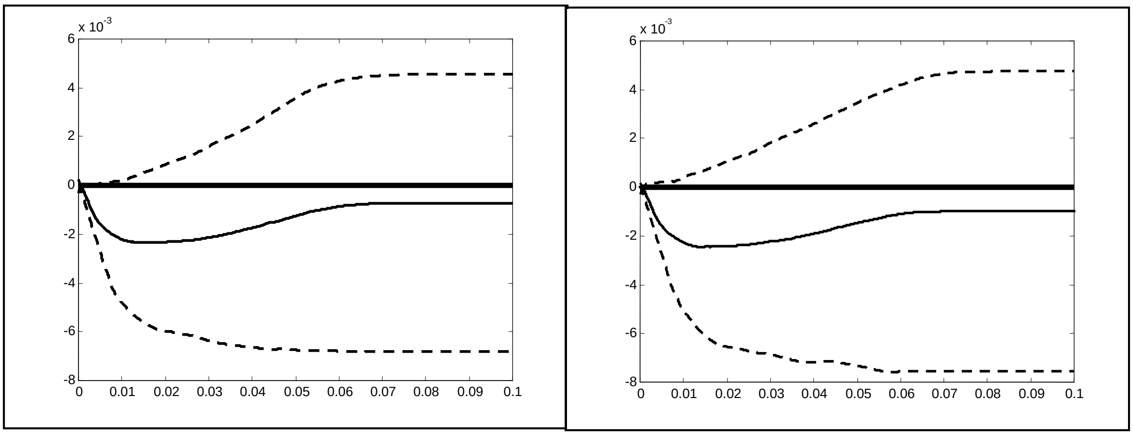
significantieniveau van 99 percent

Figuur A.5: testen van het verschil tussen IGL-curven: 1997 vs. 1988 (links: methode van Mills & Zandvakili, rechts: methode volgens gepaarde waarnemingen)



significantieniveau van 99 percent

Figuur A.6: testen van het verschil tussen IGL-curven: 1997 vs. 1992 (links: methode van Mills & Zandvakili, rechts: methode volgens gepaarde waarnemingen)



significantieniveau van 99 percent