Publicatiedatum CBS-website: 31 augustus 2007

Meten van inkomensongelijkheid

Methoden en definities
Inkomen en bestedingen

Marion van den Brakel-Hofmans

Centraal Bureau voor de Statistiek

Voorburg/Heerlen 2007

Verklaring der tekens

.	=	gegevens ontbreken
*	=	voorlopig cijfer
x	=	geheim
–	=	nihil
–	=	(indien voorkomend tussen twee getallen) tot en met
0 (0,0)	=	het getal is minder dan de helft van de gekozen eenheid
niets (blank)	=	een cijfer kan op logische gronden niet voorkomen
2006–2007	=	2006 tot en met 2007
2006/2007	=	het gemiddelde over de jaren 2006 tot en met 2007
2006/'07	=	oogstjaar, boekjaar, schooljaar enz. beginnend in 2006 en eindigend in 2007
2004/'05–2006/'07	=	boekjaar enz., 2004/'05 tot en met 2006/'07

In geval van afronding kan het voorkomen dat de totalen niet geheel overeenstemmen met de som der opgetelde getallen.

Verbeterde cijfers in de staten en tabellen zijn niet als zodanig gekenmerkt.

Colofon

Uitgever
Centraal Bureau voor de Statistiek
Prinses Beatrixlaan 428
2273 XZ Voorburg

Prepress
Centraal Bureau voor de Statistiek - Facilitair bedrijf

Inlichtingen
Tel.: (088) 570 70 70
Fax: (070) 337 59 94
Via contactformulier: www.cbs.nl/infoservice

Bestellingen
E-mail: verkoop@cbs.nl
Fax: (045) 570 62 68

Internet
www.cbs.nl

1. Inleiding

Voor het in kaart brengen van inkomensverschillen binnen een populatie [1] bestaat een groot aantal ongelijkheidsmaten. Dergelijke maten hebben als doel de inkomensongelijkheid in de populatie in één getal samen te vatten. Elke maat kent daarbij zijn eigen interpretatie en heeft zowel voor- als nadelen. Bij het beschrijven van inkomensongelijkheid is het daarom belangrijk diverse maten naast elkaar te zetten om daarmee een zo compleet mogelijk beeld van de ongelijkheid te krijgen. Het is echter onmogelijk om hierin uitputtend te zijn. Daarom concentreert dit artikel zich op de meest gangbare ongelijkheidsmaten die samen een gevarieerd beeld geven van de inkomensongelijkheid in de populatie.

Veel ongelijkheidsmaten kunnen worden geïllustreerd met één van de drie bekende grafische manieren om inkomensverdelingen weer te geven: de frequentieverdeling, de Lorenz-curve en de parade van Pen. Paragraaf 2 geeft allereerst een toelichting op deze grafieken. Paragraaf 3 bespreekt vervolgens vijf ongelijkheidsmaten. Daarbij wordt zoveel mogelijk gebruik gemaakt van illustraties met deze grafieken. In de paragrafen 3.1 en 3.2 komen eerst de *ratio 80/20* en de *relatieve interkwartielafstand* aan de orde. Daarna volgen in de paragrafen 3.3 en 3.4 de *Ginicoëfficiënt* en de *polarisatie-index*. Paragraaf 3.5 bespreekt tot slot de *Theilcoëfficiënt*. De bespreking van de vijf ongelijkheidsmaten en bijbehorende grafische illustraties wordt ondersteund met inkomensgegevens van personen afkomstig uit de Inkomensstatistiek [2] van het CBS.

In paragraaf 4 komen ter afsluiting van dit artikel de meest gangbare criteria voor ongelijkheidsmaten aan de orde. Met deze criteria kunnen de maten op hun specifieke bruikbaarheid, voor bijvoorbeeld de Inkomensstatistiek, beoordeeld worden.

2. Grafische weergaven van inkomensverdelingen

Voor het grafisch weergeven van inkomensverdelingen wordt doorgaans gebruik gemaakt van de frequentieverdeling, de Lorenz-curve of de parade van Pen. Deze paragraaf geeft een nadere toelichting op de drie grafieken aan de hand van inkomensgegevens van de Nederlandse bevolking afkomstig uit de Inkomensstatistiek van het CBS. Daarbij is voor het inkomen van een persoon uitgegaan van het gestandaardiseerd besteedbare huishoudensinkomen [3].

---

1) Dit artikel gaat uit van een populatie van personen. Evenzo kan uitgegaan worden van een populatie bestaande uit huishoudens, gemeenten, of nog andere eenheden.

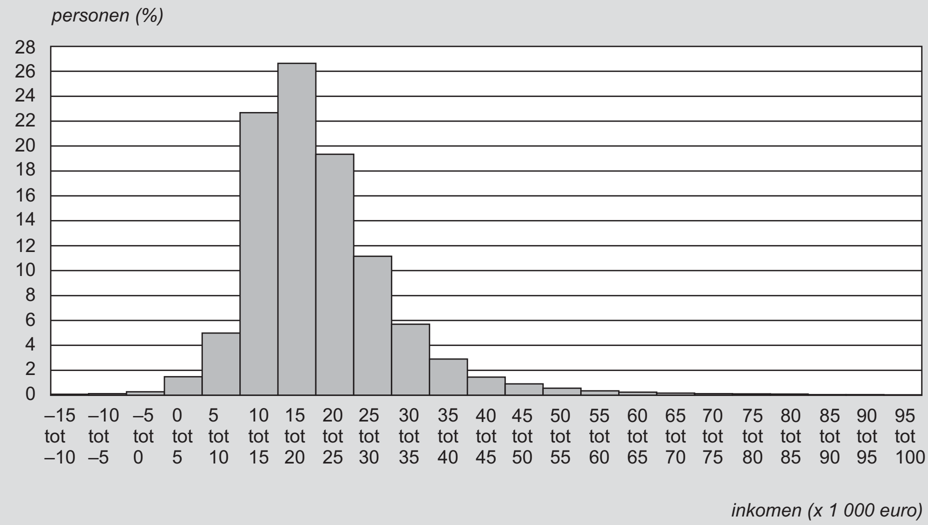
2) De Inkomensstatistiek van het CBS is gebaseerd op het Inkomenspanelonderzoek (IPO). Het IPO is een steekproefonderzoek op basis van een panel, dat sinds 1989 bestaat. Dit panel van zogenaamde kernpersonen bestaat momenteel uit circa 80 000 personen. Aan dit panel worden de huishoudensleden van de kernpersonen toegevoegd, waardoor de totale IPO steekproef uit circa 250 000 steekproefpersonen bestaat. Van de kernpersonen en hun huishoudensleden worden, in samenwerking met de Belastingdienst, zowel demografische gegevens als gegevens over het inkomen verzameld. Het CBS publiceert jaarlijks inkomenscijfers op basis van het IPO. Deze cijfers hebben betrekking op inkomensgegevens van 1 januari tot en met 31 december van het betreffende onderzoeksjaar en gaan over de totale bevolking exclusief de institutionele bevolking.

3) Uitgangspunt in de beschrijving van de personele inkomensverdeling is de vrij besteedbare inkomensruimte ('mate van welvaart') van iedere inwoner van Nederland. Deze besteedbare ruimte komt tot uitdrukking in het aan de persoon toegekende gestandaardiseerd besteedbare huishoudensinkomen. De levensstandaard van een persoon is immers verbonden met het inkomen van zijn huishouden. Het besteedbare huishoudensinkomen is opgebouwd uit lonen van werkende huishoudensleden, winst uit eigen bedrijf en inkomen uit vermogen vermeerderd met ontvangen uitkeringen en andere toelagen, en verminderd met de betaalde premies en belastingen. Om de inkomens van verschillende typen huishoudens onderling vergelijkbaar te maken, wordt het besteedbare huishoudensinkomen gecorrigeerd voor verschillen in grootte en samenstelling van het huishouden. Het aldus gestandaardiseerde huishoudensinkomen wordt vervolgens toegekend aan ieder lid van het desbetreffende huishouden als zijnde diens vrij besteedbare inkomensruimte. Wanneer in de vervolgtekst over inkomen gesproken wordt, wordt altijd het aan de persoon toegekende gestandaardiseerd besteedbare huishoudensinkomen bedoeld, tenzij dit expliciet anders is vermeld. Opgemerkt zij dat in een populatie van huishoudens naast onderzoek naar ongelijkheid van het (gestandaardiseerd) besteedbare huishoudensinkomen ook naar ongelijkheid bij andere inkomensbegrippen onderzoek gedaan kan worden, zoals ongelijkheid bij het bruto huishoudensinkomen.

2.1 Frequentieverdeling

In de frequentieverdeling (ook wel histogram genoemd) worden aantallen personen uitgezet tegen het inkomen in klassen. Het aantal inkomensklassen kan verschillen al naar gelang de gewenste detaillering. Ditzelfde geldt ook voor de gehanteerde klassenbreedte. Uit de frequentieverdeling wordt snel duidelijk waar het gemiddelde ongeveer ligt, hoe groot de spreiding van de inkomens is en of er extreme inkomens zijn. Figuur 1 geeft de frequentieverdeling van het inkomen van personen in 2004. Inkomens zijn ingedeeld in klassen van 5 duizend euro. Inkomens boven 100 duizend euro (0,3 procent van de bevolking) zijn niet weergegeven in het histogram, evenals de inkomens kleiner of gelijk aan min 15 duizend euro (0,2 procent van de bevolking). Voor elke inkomensklasse is aangegeven hoeveel personen (in procenten van de totale bevolking) zich in deze klasse bevindt.

1. Frequentieverdeling van het inkomen van personen, 2004



Bron: CBS (Inkomensstatistiek)

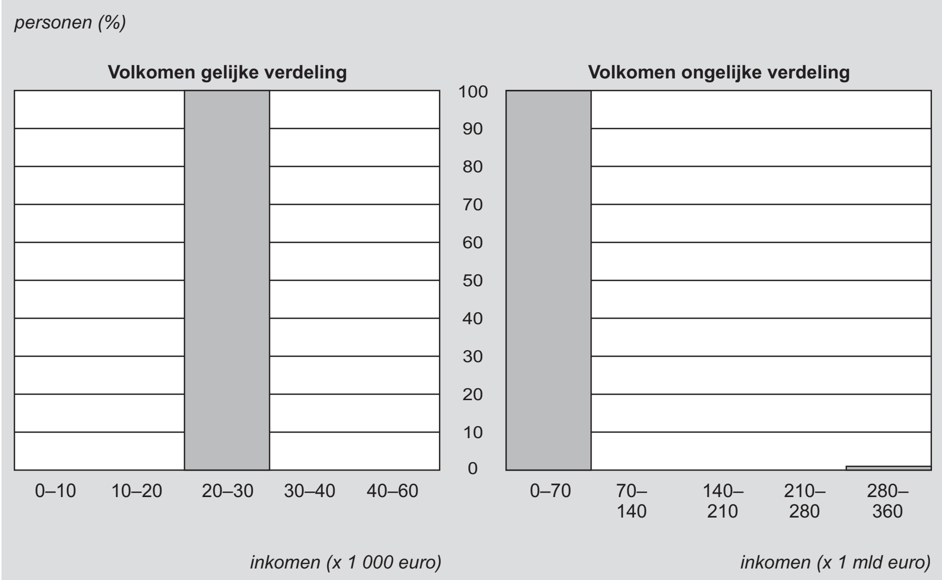
In 2004 bedroeg het gemiddelde inkomen van personen bijna 21 duizend euro. De mediaan, dat is bij een ordening van alle personen naar hoogte van het inkomen het inkomen van de middelste persoon (of het gemiddelde inkomen van de twee middelste personen), lag met bijna 19 duizend euro onder dit gemiddelde. Dat is kenmerkend voor inkomensverdelingen, die doorgaans scheef naar links verdeeld zijn. De meeste mensen (ruim 26 procent) bevinden zich in de klasse van 15 tot 20 duizend euro. Ruim 0,5 procent van de bevolking had in 2004 een negatief inkomen. Dit betreft voornamelijk zelfstandigen die verlies hebben geleden. Anderzijds had ruim 2 procent van de bevolking (ruim 300 duizend mensen) een inkomen boven 50 duizend euro.

Bij een volkomen gelijke inkomensverdeling heeft iedereen eenzelfde inkomen. In dit geval geeft het histogram in één klasse alle personen aan, terwijl de andere klassen geen personen bevatten. Wanneer in 2004 iedereen een gelijk inkomen zou hebben gehad, zou iedereen tot de klasse van 20 tot 25 duizend hebben behoord. In die klasse valt immers het gemiddelde inkomen. In tegenstelling hiermee is sprake van een totaal ongelijke verdeling wanneer één persoon al het inkomen (bijna 332 miljard euro) bezit en de overige personen geen inkomen hebben. Bijna iedereen behoort dan tot de laagste klasse in het histogram (zie figuur 2).

2.2 De Lorenz-curve

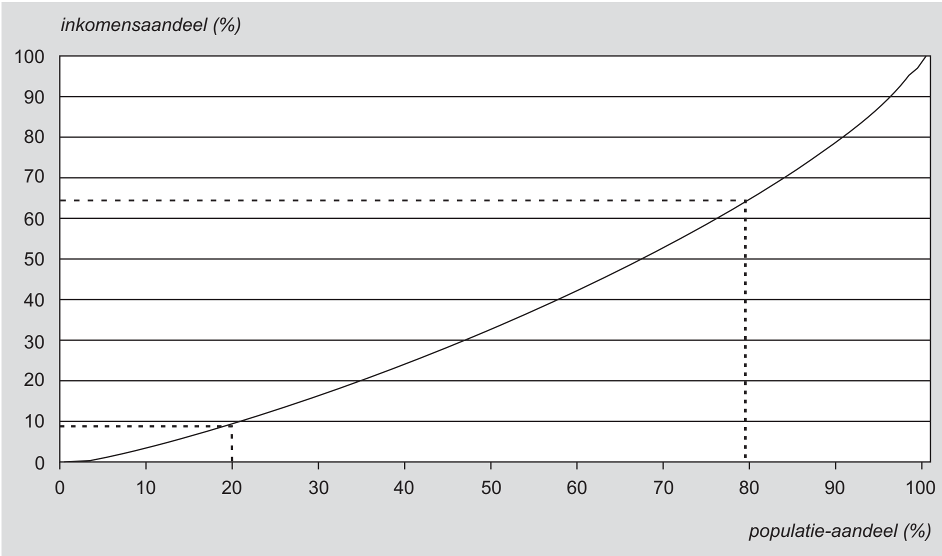
De Lorenz-curve is in 1905 ontwikkeld door Max. O. Lorenz om inkomensverdelingen weer te geven. In feite is de Lorenz-curve niets anders dan de cumulatieve verdelings-

2. Frequentieverdeling bij volkomen gelijke en volkomen ongelijke verdeling



functie van het inkomen: de cumulatieve inkomensaandelen op de Y-as worden uitgezet tegen de cumulatieve populatie-aandelen op de X-as. Voor elk percentage personen (*x* %) met de laagste inkomens geeft de curve aldus aan welk percentage *y* % van het totale inkomen zij bezitten. Uit figuur 3 is af te lezen dat 20 procent van de personen met de laagste inkomens in 2004 ruim 9 procent van het totale inkomen bezitten. Evenzo bezat 80 procent van de bevolking bijna 65 procent van het totale inkomen, zodat 20 procent van de personen met de hoogste inkomens ruim 35 procent van het totale inkomen bezaten.

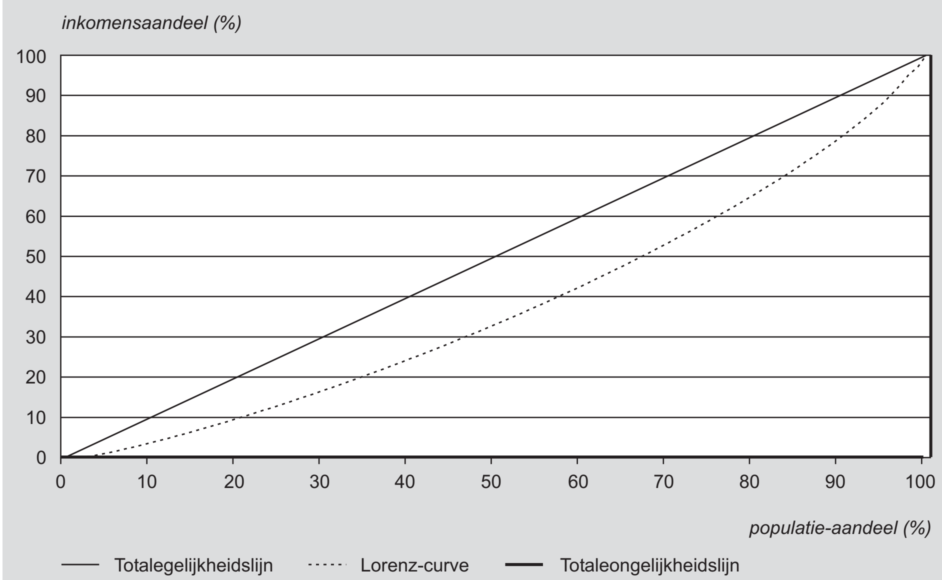
3. Lorenz-curve van het inkomen van personen, 2004



Bron: CBS (Inkomensstatistiek)

Bij een volkomen gelijke inkomensverdeling (iedereen heeft hetzelfde inkomen) bezit *p* procent van de personen *p* procent van het totale inkomen. Dit wordt weergegeven door de lijn van totale gelijkheid in figuur 4. In geval van een totaal ongelijke verdeling (één persoon bezit al het inkomen) geldt dat het inkomensaandeel gelijk aan nul is voor alle populatie-aandelen kleiner dan 100 procent. Bij het populatie-aandeel van 100 procent is het inkomensaandeel eveneens gelijk aan 100 procent. Hoe dichter de Lorenz-curve van een inkomensverdeling bij de totalegelijkheidslijn ligt, hoe kleiner de inkomensverschillen, en omgekeerd.

4. Lorenz-curve bij een volkomen gelijke en volkomen ongelijke verdeling



2.3 Parade van Pen

Een beeldende manier om een inkomensverdeling te beschrijven is de bekende 'Parade van Reuzen en Dwergen' van Pen (1971), kortweg aangeduid met de parade van Pen. In deze parade marcheren mensen achter elkaar op volgorde van de hoogte van hun inkomen. Daarbij is denkbeeldig iemands lengte evenredig gemaakt aan zijn of haar inkomen. Iemand met een gemiddeld inkomen (20,8 duizend euro in 2004) krijgt de gemiddelde lengte (1,80 meter). Mensen met een lager inkomen worden in elkaar geduwd en mensen met een hoger inkomen worden uitgerekt. Aan iemands uiterlijk is zo te zien wat zijn of haar levensstandaard is. In figuur 5 is de parade van Pen weergegeven voor 2004. Op de horizontale as staan personen in volgorde van hun inkomen [4], op de verticale as staan inkomens. Om presentatieredenen zijn in figuur 5 de personen samengevoegd tot honderd 1%-groepen. De hoogte van een staafje in de figuur representeert het gemiddelde inkomen van een 1%-groep.

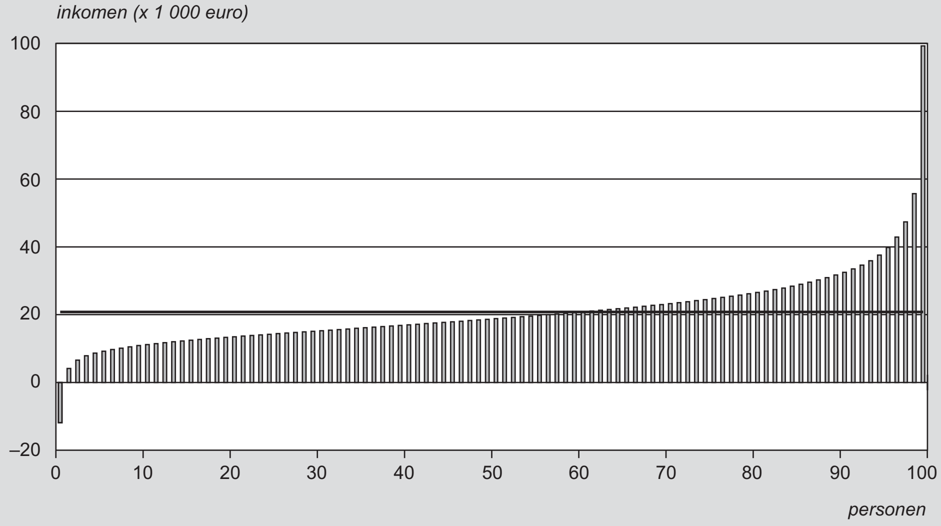
In de voorbijtrekkende parade komen vooraan in de stoet alleen onder de grond mensen voorbij, ze lopen als het ware ondersteboven. Deze mensen, veelal zelfstandigen, hebben een negatief inkomen. Daarna volgen de dwergen met een laag inkomen, veelal uit een uitkering of uit een laaggeschoolde baan, en gaat de stoet verder met steeds groter wordende dwergen. Ook Jan Modaal met zijn gezin [5] en een besteedbaar huishoudensinkomen van ruim 21 duizend euro trekt dan voorbij. Het gestandaardiseerd besteedbare huishoudensinkomen van dit gezin bedraagt ruim 11 duizend euro, en de familie Modaal is dus amper een meter lang. Op het moment dat Jan Modaal voorbij komt, is bijna 11 procent van de bevolking gepasseerd. Als de mediaan voorbij komt is precies de helft van de bevolking gepasseerd en bij het voorbij komen van de mensen met het gemiddelde inkomen is dat bijna 60 procent. Van de reuzen aan het einde van de parade is het merendeel zelfstandige of directeur.

Uit de parade van Pen kan dus eenvoudig een indruk verkregen worden van bekende kengetallen van een inkomensverdeling. Zo is de mediaan van de inkomensverdeling (bijna 19 duizend euro in 2004) af te lezen bij het percentage personen gelijk aan 50 pro-

[4] Pen (1971) koppelt aan de voorbijtrekkende parade een tijdsaspect. De stoet trekt in precies een uur voorbij, zodat iedere minuut ruim een kwart miljoen mensen langskomen. Met dit tijdsaspect kan de inkomensverdeling nog beeldender beschreven worden omdat de scheefheid van de verdeling tot uitdrukking komt. Zo passeren pas na 36 minuten (dus niet na een half uur) de mensen van gemiddelde lengte.

[5] Het begrip modaal inkomen wordt door het Centraal Planbureau gebruikt voor koopkrachtoverzichten. Het is vastgesteld als het brutoloon van een alleenverdiener in de marktsector met een partner en twee kinderen en valt (voor de wijziging van de ziektekostenwet in 2006) net onder de premiegrens Ziekenfondswet.

5. Parade van Pen voor het inkomen van personen, 2004

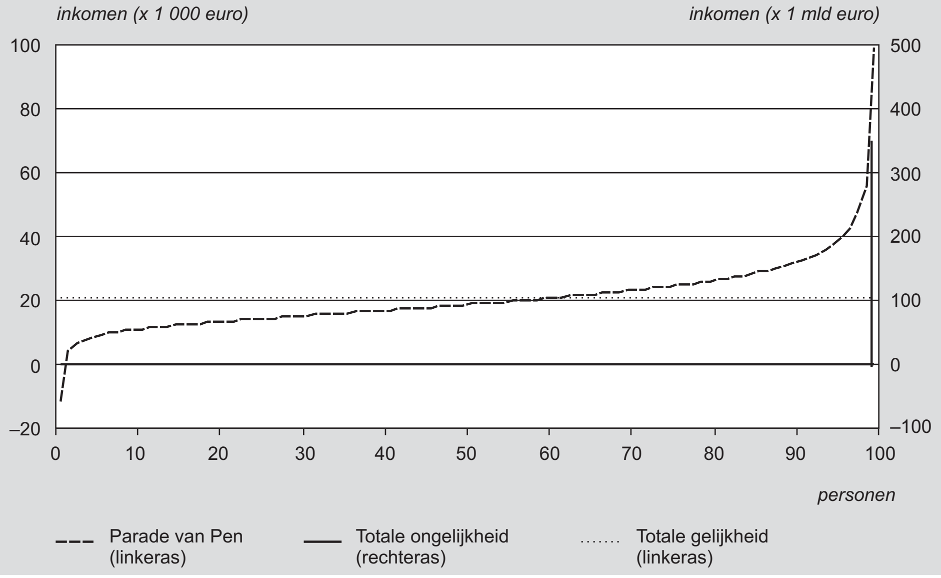


Bron: CBS (Inkomenstatistiek)

cent. Ook decielgrenzen[6)] en en kwartielgrenzen kunnen uit de parade van Pen gehaald worden. Het gemiddelde inkomen van bijna 21 duizend euro in 2004 is als een horizontale lijn weergegeven in figuur 5. Opgemerkt zij dat wanneer op de verticale as van de parade van Pen in plaats van inkomens cumulatieve inkomensaandelen worden weergegeven, de eerder besproken Lorenz curve ontstaat.

Bij een verdeling waarin iedereen hetzelfde inkomen heeft, toont de parade van Pen een horizontale lijn ter hoogte van het gemiddelde inkomen. Bij een volkomen ongelijke verdeling is de parade van Pen overal nul, behalve bij de ene persoon die al het inkomen bezit (figuur 6).

6. Parade van Pen bij een volkomen gelijke en volkomen ongelijke verdeling



---

6) Een indeling van een populatie in decielgroepen ontstaat door personen oplopend te ordenen naar hoogte van hun inkomen en vervolgens tien groepen personen van gelijke omvang te vormen. De 10 procent personen met de laagste inkomens worden doorgaans aangeduid als het laagste deciel, de 10 procent personen met de hoogste inkomens als het hoogste deciel. Decielgrenzen zijn achtereenvolgens: het hoogste inkomen van het laagste deciel, het hoogste inkomen van het tweede deciel, het hoogste inkomen van het derde deciel, …, het hoogste inkomen (van het hoogste deciel).

3. Ongelijkheidsmaten

Grafieken van inkomensverdelingen, zoals de parade van Pen, de Lorenz-curve en de frequentieverdeling, geven enig inzicht in de mate van ongelijkheid van inkomens in een populatie. Vooral bij extreme verdelingen (zoals de volkomen gelijke en de volkomen ongelijke verdeling) is het beeld van ongelijkheid in één oogopslag duidelijk uit de grafieken. In de praktijk komen dergelijke verdelingen echter zelden voor en bestaat naast de behoefte aan het presenteren van inkomensverdelingen in grafiek de behoefte aan een kwantitatieve uitdrukking van de bijbehorende ongelijkheid in één getal. Deze paragraaf bespreekt inkomensongelijkheid in getal en grafiek door vijf ongelijkheidsmaten te presenteren en waar mogelijk te illustreren aan de hand van de parade van Pen, de Lorenz-curve of de frequentieverdeling.

3.1 Ratio 80/20

Eenvoudige en daarom populaire ongelijkheidsmaten zijn maten gebaseerd op ratio's van inkomensaandelen. Zo worden decielverhoudingen, zoals de ratio van de inkomensaandelen van het hoogste deciel (hoogste 10 procent) en de drie laagste decielen (laagste 30 procent), vaak gebruikt als ongelijkheidsmaat. Ook het aandeel van de 50 procent laagste inkomens in de totale inkomenssom is een bekende ongelijkheidsmaat gebaseerd op een ratio van inkomensaandelen.

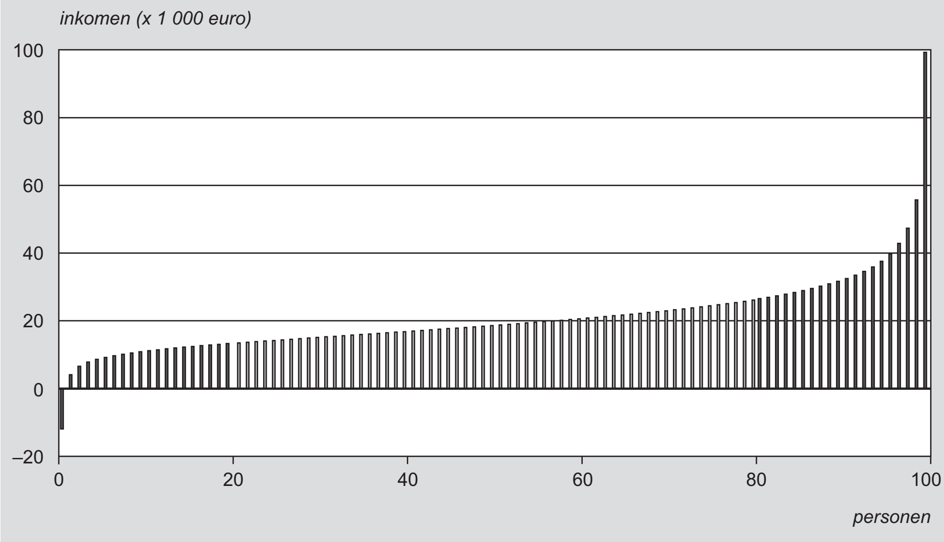
In Europees verband wordt vaak de ratio van de inkomensaandelen van de 20 procent hoogste inkomens (hoogste quintiel) en de 20 procent laagste inkomens (laagste quintiel), de ratio 80/20, gebruikt. De interpretatie van deze maat is vrij eenvoudig: ze geeft aan hoeveel meer van de 'inkomenskoek' naar het hoogste quintiel gaat. Precies gedefinieerd is de ratio 80/20 de ratio van het totale inkomen van de hoogste quintielgroep (Q_{80}) en het totale inkomen van de laagste quintielgroep (Q_{20}):

$$\text{ratio 80/20} = \frac{\sum_{i \in Q_{80}} w_i x_i}{\sum_{i \in Q_{20}} w_i x_i}$$

met x_i het inkomen van waarneming i, w_i het gewicht van waarneming i.

Ongelijkheidsmaten gebaseerd op ratio's van inkomensaandelen kunnen inzichtelijk geïllustreerd worden met de parade van Pen. In de parade van Pen is de ratio 80/20 de verhouding van de donkergekleurde oppervlakken (figuur 7). In 2004 was de ratio 80/20

7. Parade van Pen voor het inkomen van personen, 2004

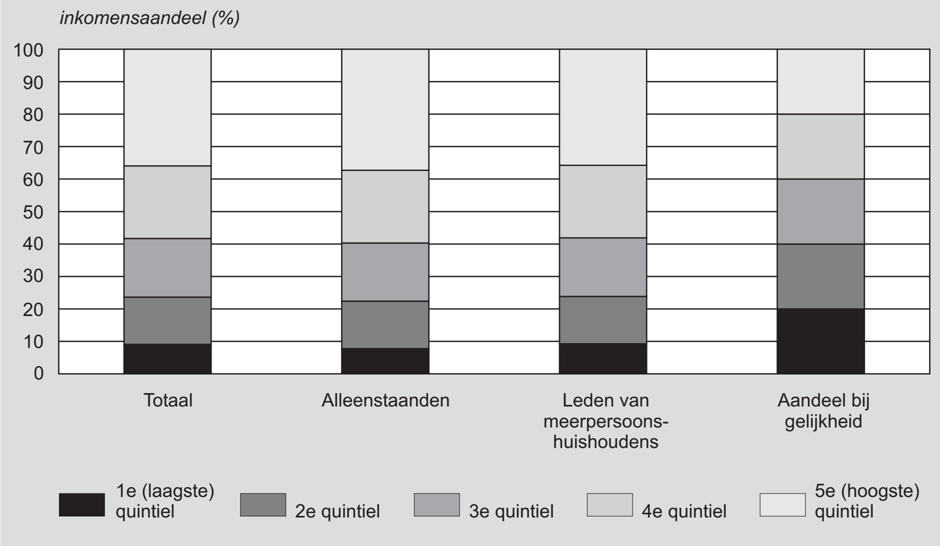


Bron: CBS (Inkomensstatistiek)

ruim 4 (tabel 1). Het inkomensaandeel van de rijkste 20 procent was dus ruim vier keer groter dan het aandeel van de armste 20 procent. Voor alleenstaanden was de ratio 80/20 gelijk aan 4,80 en voor leden van een meerpersoonshuishouden 3,87 (tabel 1). Bij de alleenstaanden liggen de inkomens dus verder uit elkaar en is er op basis van de ratio 80/20 meer sprake van ongelijkheid in inkomen.

In figuur 8 wordt de ongelijkheid op basis van de ratio 80/20 nog eens grafisch geïllustreerd door de gestapelde staafdiagrammen, die de inkomensaandelen van de quintielen laat zien voor alleenstaanden en leden van een meerpersoonshuishouden. De laatste staaf in figuur 8 geeft de quintielaandelen van een gelijke inkomensverdeling weer. Voor deze verdeling is de ratio 80/20 gelijk aan 1. Voorts laat figuur 8 zien dat het inkomensaandeel van de hoogste quintielgroep ruim anderhalf maal hoger is dan het aandeel van deze quintielgroep in een gelijke inkomensverdeling. Dit geldt zowel voor alleenstaanden als voor leden van een meerpersoonshuishouden.

8. Inkomensaandeel van personen per quintiel naar huishoudenssamenstelling, 2004



Bron: CBS (Inkomensstatistiek)

De ratio 80/20 concentreert zich op de staarten van de verdeling en is daarmee ongevoelig voor de tussenliggende verdeling. De ratio 80/20 kan sterk reageren op extreme inkomens in de staarten van de verdeling. De volgende paragraaf bespreekt een ongelijkheidsmaat die niet beïnvloed wordt door extreme inkomens.

3.2 Relatieve interkwartielafstand

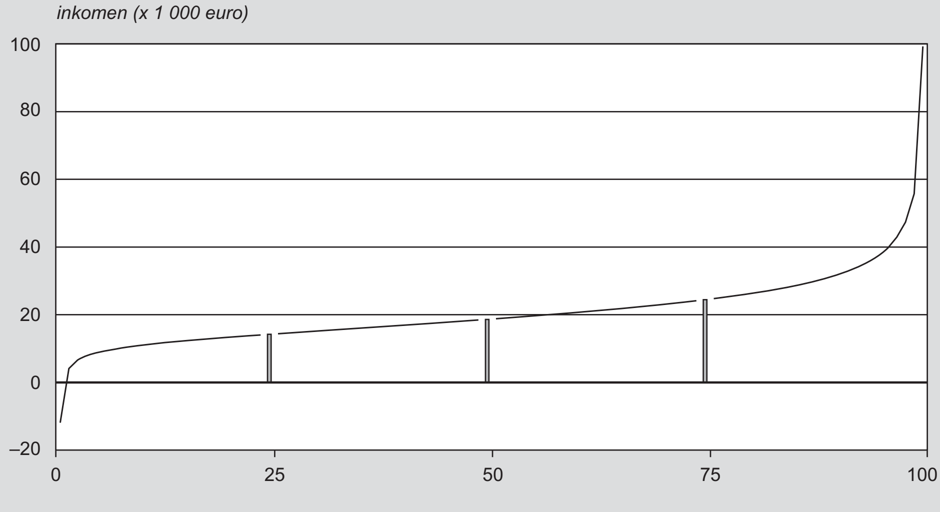
Een ongelijkheidsmaat die niet gevoelig is voor extreme inkomens in de staarten van een inkomensverdeling is de relatieve interkwartielafstand (Relatieve IKA). Deze maat concentreert zich op het midden van de verdeling en is gebaseerd op kwartielen. De maat is gedefinieerd als het verschil tussen het hoogste inkomen in het derde kwartiel (I_{75}) en het hoogste inkomen in het eerste kwartiel (I_{25}) gedeeld door de mediaan (m):

$$\text{relatieve IKA} = \frac{I_{75} - I_{25}}{m} = \frac{IKA}{m}$$

Naarmate de Relatieve IKA hoger is, liggen de inkomens verder uit elkaar en is er dus sprake van een hogere inkomensongelijkheid.

Evenals de ratio 80/20 kan de Relatieve IKA grafisch geïllustreerd worden door de parade van Pen. In de parade van Pen is de Relatieve IKA gebaseerd op de inkomens bij de percentages personen 75% (I_{75}), 25% (I_{25}) en 50% (mediaan m), zie figuur 9.

9. Parade van Pen voor het inkomen van personen, 2004



Bron: CBS (Inkomensstatistiek)

De Relatieve IKA geeft de mate van spreiding rond het mediane inkomen weer. Helaas gaat echter nogal wat informatie verloren doordat het tweede en vierde kwartiel van de betreffende inkomensverdeling buiten beschouwing blijven. Zo heeft de Relatieve IKA van een volkomen gelijke verdeling dezelfde waarde als die van een volkomen ongelijke verdeling, namelijk nul.

Voor 2004 komt de Relatieve IKA voor de Nederlandse bevolking uit op 0,57. Dit is ook zo voor alleenstaanden. Voor leden van een meerpersoonshuishouden is de Relatieve IKA in 2004 gelijk aan 0,55 (tabel 1). Op basis hiervan is dus enigszins sprake van verschil in ongelijkheid.

3.3 Ginicoëfficiënt

De meest gebruikte maat voor inkomensongelijkheid is de Ginicoëfficiënt. In tegenstelling tot de ratio 80/20 en de Relatieve IKA beschouwt de Ginicoëfficiënt alle inkomens van de betreffende populatie. De maat is ontwikkeld door de Italiaanse statisticus Corrado Gini en gepubliceerd in 1912. De Ginicoëfficiënt is eenvoudig te interpreteren dankzij de grafische ondersteuning door de Lorenz-curve (zie hieronder). Daarnaast ligt de waarde van de coëfficiënt tussen 0 en 1, waarbij 0 correspondeert met totale gelijkheid (iedereen heeft hetzelfde inkomen) en 1 correspondeert met totale ongelijkheid (één persoon bezit al het inkomen). Ook dit draagt bij aan de populariteit van de Ginicoëfficiënt.

De Ginicoëfficiënt is gelijk aan de helft van het relatieve gemiddelde verschil:

$$G = \frac{\Delta}{2\mu}$$

waarbij Δ gelijk is aan $\Delta = \sum_{i=1}^{n} \sum_{j=1}^{n} \left| x_i - x_j \right| w_i \, w_j$ met n de steekproefomvang, x_i het inkomen van persoon i, w_i het gewicht van persoon i $(\sum_i w_i = 1)$ en μ het gewogen gemiddelde inkomen $(\sum_{i=1}^{n} w_i \, x_i)$.

In de teller van de Ginicoëfficiënt wordt de som berekend van de absolute waarden van de inkomensverschillen van elke persoon met alle andere personen. Hoe kleiner deze som, hoe geringer de ongelijkheid. Omdat het inkomensverschil tussen elk tweetal personen tweemaal bepaald wordt, wordt in de formule het resultaat door 2 gedeeld. Delen door het gemiddeld μ zorgt er vervolgens voor dat de Ginicoëfficiënt een waarde tussen

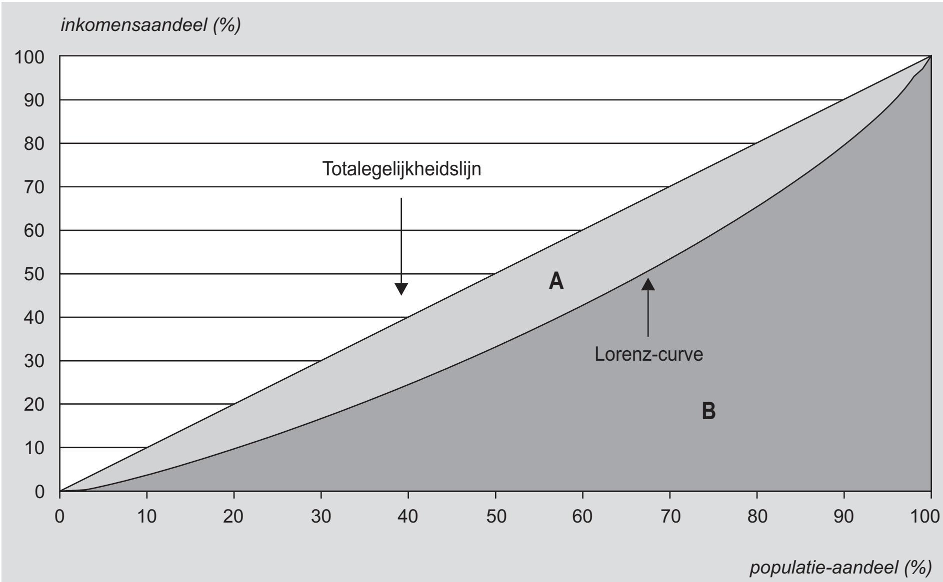
0 en 1 heeft. In de berekening van de Ginicoëfficiënt worden nulinkomens en negatieve inkomens gewoon meegenomen.

De Lorenz-curve biedt een eenvoudige interpretatie van de Ginicoëfficiënt. In figuur 10 is nogmaals de Lorenz-curve voor het inkomen van personen (2004) gegeven. Bovendien is de Lorenz-curve van een volkomen gelijke verdeling gegeven (totalegelijkheidslijn). De Ginicoëfficiënt *G* wordt berekend op basis van de Lorenz-curve als de ratio

$$G = \frac{oppervlakte\ A}{oppervlakte\ A + oppervlakte\ B}$$

Hoe dichter de Lorenz-curve van een inkomensverdeling bij de totalegelijkheidslijn ligt, hoe lager de inkomensongelijkheid én hoe lager de Ginicoëfficiënt. Merk op dat *oppervlakte A + oppervlakte B* gelijk is aan ½, zodat *G = 2A = 1–2B*.

10. Lorenz-curve en Ginicoëfficiënt



Gelijke Ginicoëfficiënten betekenen niet automatisch gelijke inkomensverdelingen. Twee Lorenz-curven kunnen verschillend van vorm zijn, terwijl ze toch dezelfde Ginicoëfficiënt geven. Een extreem voorbeeld is een inkomensverdeling waarbij de helft van de personen geen inkomen heeft en bij de andere helft het totale inkomen gelijkwaardig verdeeld is. Dan is *G* = ½. Een inkomensverdeling waarbij één rijk persoon de helft van het totale inkomen bezit en de andere helft van het totale inkomen gelijkwaardig verdeeld is onder de overige bevolking, heeft eveneens een *G* gelijk aan ½. De Lorenz-curven van beide verdelingen zijn echter volkomen verschillend.

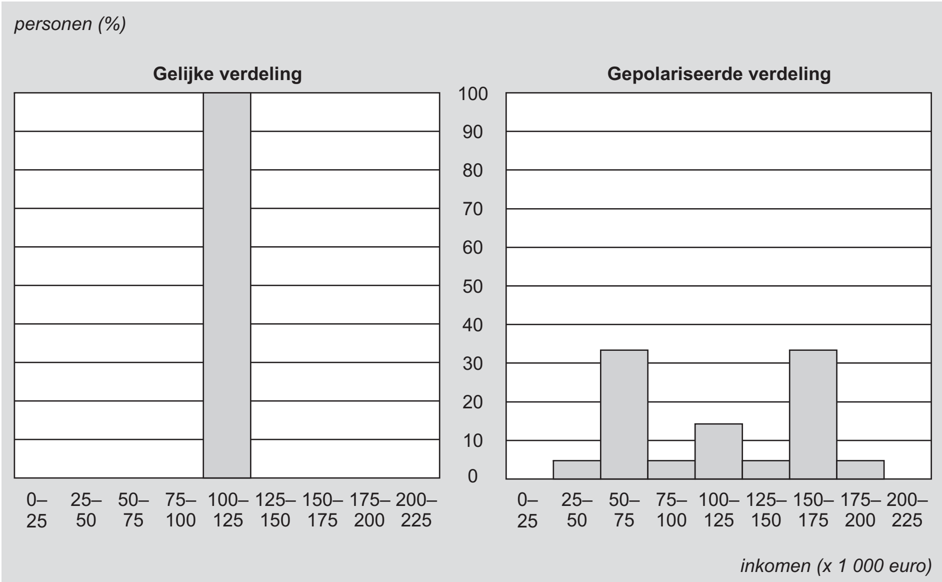
De Ginicoëfficiënt is heel geschikt om groepen onderling te vergelijken, bijvoorbeeld inkomensongelijkheden in provincies of bij verschillende typen huishoudens. Daarnaast kan met de Ginicoëfficiënt goed in beeld gebracht worden hoe de inkomensongelijkheid in een bepaalde periode verandert, bijvoorbeeld de verandering van de inkomensongelijkheid van 2001 tot 2005 in Nederland. Nadeel van de Ginicoëfficiënt is dat deze vooral gevoelig is voor veranderingen die plaatsvinden rondom het gemiddelde van de inkomensverdeling en wat minder voor veranderingen in de staarten van de verdeling (Allison, 1978). Hierdoor varieert de Ginicoëfficiënt relatief weinig wanneer bijvoorbeeld in een reeks van jaren de welvaart van de rijkeren afneemt ten gunste van de armeren. Daarom is het belangrijk naast de Ginicoëfficiënt ook andere ongelijkheidsmaten te presenteren.

In tabel 1 zijn de Ginicoëfficiënten gegeven voor het inkomen van personen in 2004. De waarde van de coëfficiënt is voor de totale bevolking gelijk aan 0,27. Voor alleenstaanden is die waarde 0,29 en voor leden van een meerpersoonshuishouden 0,26. Dat is een verschil van drie procentpunten. Onder alleenstaanden bestaat duidelijk een hogere inkomensongelijkheid.

3.4 Polarisatie-index

Polarisatie houdt in dat er steeds meer een groep met hoge inkomens en een groep met lage inkomens ontstaat en dat middeninkomens verdwijnen. In figuur 11 is dit geïllustreerd aan de hand van het histogram van een verdeling met alleen middeninkomens (gelijke verdeling) en het histogram van een meer gepolariseerde inkomensverdeling. Deze laatste verdeling kent ten opzichte van de gelijke verdeling meer lage en hoge inkomens en de groep middeninkomens is veel kleiner.

11. Frequentieverdeling van een gelijke en een gepolariseerde inkomensverdeling.



De mate van polarisatie komt tot uitdrukking in de polarisatie-index *P*, die verband houdt met de Ginicoëfficiënt *G*:

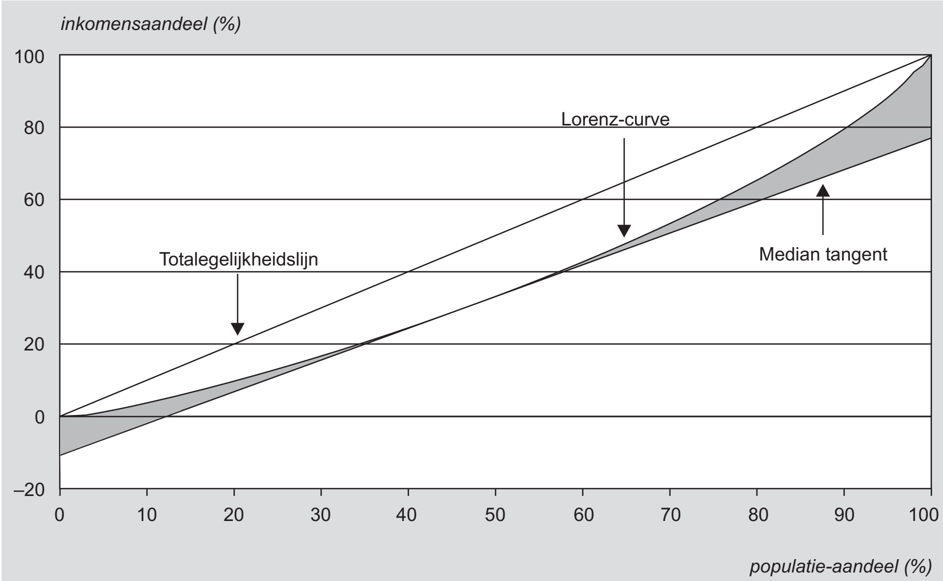
$$P = 2\frac{\mu}{m}(1 - G - 2L_{½})$$

met μ het gemiddelde inkomen, *m* de mediaan, *G* de Ginicoëfficiënt en $L_{½}$ het inkomensaandeel van de laagste helft van de bevolking.

De polarisatie-index ligt, net als de Ginicoëfficiënt *G*, tussen 0 en 1. *P* is, evenals *G*, gelijk aan 0 voor een verdeling met volkomen gelijke inkomens (het inkomensaandeel van de laagste helft van de bevolking is dan ½). Als één persoon al het inkomen bezit, is *P* strikt genomen niet gedefinieerd omdat de mediaan dan gelijk is aan 0. Voor verdelingen die deze extreme verdeling benaderen (één persoon heeft bijna alles en de rest van de bevolking bijna niets) is *P*, net als *G*, bijna gelijk aan 1. Voor dergelijke extreme inkomensverdelingen ontlopen *P* en *G* elkaar dus niet. De polarisatie-index heeft echter in bepaalde opzichten een toegevoegde waarde ten opzichte van andere ongelijkheidsmaten. Dit wordt duidelijk gemaakt door *P* eerst grafisch toe te lichten aan de hand van de Lorenz-curve en vervolgens een voorbeeld te geven waarin polarisatie niet samengaat met een grotere inkomensongelijkheid.

Figuur 12 geeft de Lorenz-curve voor een willekeurige inkomensverdeling. In Wolfson (1997) is afgeleid dat het verschijnsel polarisatie wordt weergegeven door de oppervlakte tussen de Lorenz-curve en de zogenaamde mediantangentlijn (donkergekleurde deel in figuur 12). De mediantangentlijn is de lijn met een helling gelijk aan m/μ die raakt aan de Lorenz-curve in het punt waar x = 50 procent. Vermenigvuldiging van deze gearceerde oppervlakte met vier zorgt er vervolgens voor dat de index *P* tussen 0 en 1 ligt. Hoe dichter de Lorenz-curve van een inkomensverdeling bij de totaleongelijkheidslijn ligt, hoe kleiner de oppervlakte tussen de curve en de mediantangentlijn, en hoe kleiner dus ook *P*.

12. Grafische weergave polarisatie



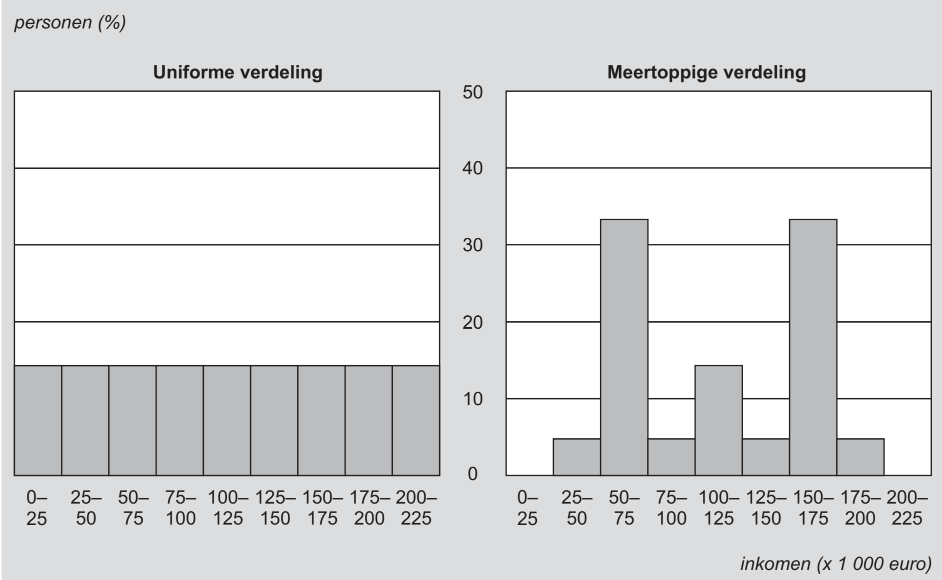
Het volgende voorbeeld toont aan dat meer polarisatie niet altijd samengaat met een grotere inkomensongelijkheid. Staat 1 bevat de gegevens van twee inkomensverdelingen: een uniforme verdeling en een meertoppige verdeling. In een uniforme inkomensverdeling komt ieder inkomen even vaak voor. In de uniforme verdeling van dit voorbeeld komen de verschillende weergegeven inkomens tussen 25 duizend en 175 duizend telkens 3 duizend keer voor. De meertoppige verdeling is als het ware ontstaan uit de uniforme verdeling doordat een aantal middeninkomens verdwenen is en er lage en hoge inkomens zijn bijgekomen. Het totale inkomen van de populatie is hierbij gelijk gebleven.

Staat 1
Voorbeeld van twee inkomensverdelingen

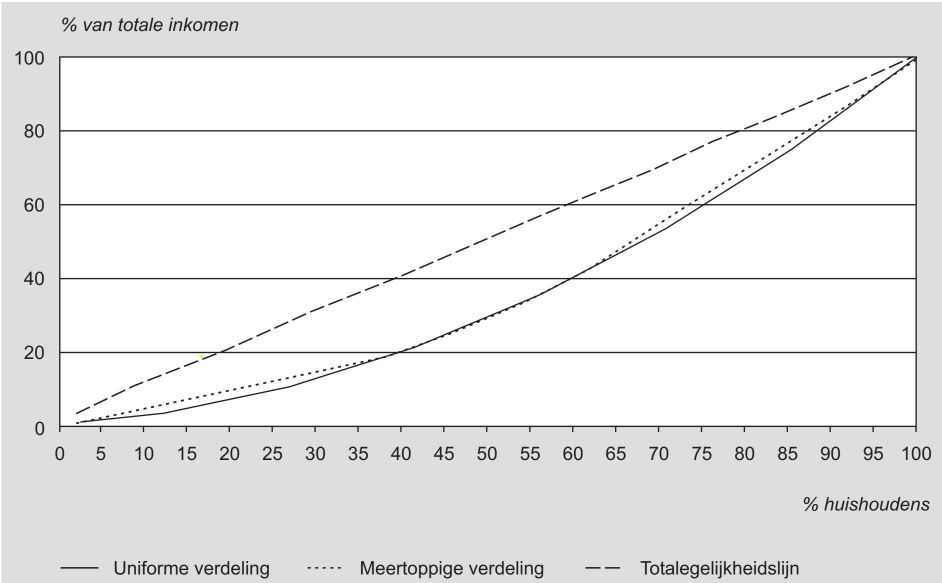
	Uniforme verdeling		Meertoppige verdeling	
	aantal personen	inkomen	aantal personen	inkomen
	x 1 000	*1 000 euro*	*x 1 000*	*1 000 euro*
	3	25	1	25
	3	50	7	50
	3	75	1	75
	3	100	3	100
	3	125	1	125
	3	150	7	150
	3	175	1	175
Totaal	21	2 100	21	2 100

Figuur 13 laat van beide verdelingen de bijbehorende histogrammen zien. Duidelijk is dat de meertoppige verdeling meer gepolariseerd is dan de uniforme verdeling. De meer gepolariseerde meertoppige verdeling is echter ook een meer gelijke verdeling, wat tot uitdrukking komt in figuur 14. In deze figuur zijn voor beide verdelingen de Lorenz-curven weergegeven. Er blijkt dat de Lorenz-curve van de meertoppige verdeling dichter bij de totalegelijkheidslijn ligt dan de Lorenz-curve van de uniforme verdeling. Dit betekent onder meer dat de Ginicoëfficiënt van de meertoppige verdeling lager is dan die van de uniforme verdeling (zie ook staat 2). Dit geldt ook voor de ratio 80/20. In dit geval gaat meer polarisatie dus gelijk op met minder inkomensongelijkheid.

13. Frequentieverdeling van uniforme en meertoppige verdeling



14. Lorenz-curve van de uniforme en meertoppige verdeling



Staat 2
Kengetallen van inkomensongelijkheid van uniforme en meertoppige verdeling

	Uniforme verdeling	Meertoppige verdeling
Ginicoëfficiënt	0,29	0,26
Ratio 80/20	5,40	3,57
Relatieve IKA	1,00	1,00
Polarisatie-index	0,38	0,43

3.5 Theilcoëfficiënt

In het voorgaande zijn ongelijkheidsmaten besproken die geïllustreerd kunnen worden met bekende grafische weergaven van inkomensverdelingen. Een geheel andere groep ongelijkheidsmaten vormen maten die gebaseerd zijn op het entropieconcept uit de informatietheorie. Deze maten hebben geen eenvoudige illustratiemogelijkheid. Zeer bekend in dit kader is de Theilcoëfficiënt (Theil, H., 1967).

De Theilcoëfficiënt wordt, net als de Ginicoëfficiënt, zeer veel gebruikt bij de beschrijving van inkomensongelijkheid, hoewel het door het onderliggende entropieconcept een iets lastiger te interpreteren maat is. De formule voor de Theilcoëfficiënt T is

$$T=\sum_{i=1}^{n}\frac{x_i w_i}{\mu}\ln\left(\frac{x_i}{\mu}\right)=\frac{1}{\mu}\left(\sum_{i=1}^{n}w_i x_i\ln(x_i)\right)-\ln(\mu)$$

met n de steekproefomvang, x_i het inkomen van persoon i, w_i het gewicht van persoon i ($\sum_i w_i=1$) en μ het (gewogen) gemiddeld inkomen.

De Theilcoëfficiënt is dus gelijk aan het gemiddelde van de logaritme van alle relatieve inkomensaandelen, gewogen met de inkomensaandelen. In het geval van negatieve inkomens wordt de methode van Odink en Van Imhoff (1984) toegepast: de totale bijdrage aan de ongelijkheid van de inkomens waarvoor het cumulatieve inkomen negatief is, wordt gelijk aan nul gesteld. Sommatie vindt dus plaats over de eenheden waarvoor het cumulatieve inkomen positief is, uitgaande van een oplopende sortering naar hoogte van inkomen.

Bij een volkomen gelijke inkomensverdeling is T = 0. Als één iemand al het inkomen bezit (totale inkomensongelijkheid) dan is T gelijk aan $ln(n)$ met n de steekproefomvang. De Theilcoëfficiënt kent dus geen vast maximum. Theil (1967, p. 92) stelt dat de afhankelijkheid van n een wenselijke eigenschap is van de coëfficiënt. Een groep mensen bestaande uit twee personen waarin één persoon al het inkomen geniet is immers intuïtief minder ongelijk dan een grote groep mensen waarin één persoon alles bezit. Het vergelijken van de inkomensongelijkheid van twee populaties (bijvoorbeeld twee landen) met behulp van de Theilcoëfficiënten gebaseerd op twee van omvang verschillende steekproeven kan echter tot verkeerde conclusies leiden. In dergelijke situaties kan de Theilcoëfficiënt gedeeld worden door zijn maximale waarde. De waarde van de coëfficiënt varieert dan tussen 0 en 1.

In tabel 1 staan de Theilcoëfficiënten voor het inkomen van personen in 2004. Voor alleenstaanden is de waarde van de Theilcoëfficiënt gelijk aan 0,18, voor leden van een meerpersoonshuishouden is dat 0,14. Evenals bij de andere ongelijkheidsmaten is op basis van de Theilcoëfficiënt sprake van aanzienlijk verschil in ongelijkheid tussen alleenstaanden en meerpersoonshuishoudens.

De Theilcoëfficiënt reageert sterker op veranderingen in de staarten van de inkomensverdeling dan de Ginicoëfficiënt (Allison, 1978) en de Relatieve IKA. Daarmee vullen deze maten elkaar bij onderzoek naar inkomensongelijkheid goed aan. Nog een meerwaarde van de Theilcoëfficiënt is dat deze te schrijven is als een gewogen som van ongelijkheid binnen subgroepen:

$$T=\sum_{k=1}^{m}S_k T_k+\sum_{k=1}^{m}S_k\ln\left(\frac{\mu_k}{\mu}\right).$$

Hierbij is de populatie verdeeld in m subgroepen en is S_k het inkomensaandeel van groep k, T_k de Theilcoëfficiënt van groep k en μ_k het gemiddelde inkomen van groep k.
De eerste term aan de rechterzijde geeft de binnengroepsongelijkheid weer en de tweede term de tussengroepsongelijkheid. Zo is de inkomensongelijkheid in Nederland de gewogen som van ongelijkheden binnen de verschillende huishoudenstypen en ongelijkheden tussen de verschillende huishoudenstypen, waarbij wordt gewogen met de inkomensaandelen van de verschillende huishoudenstypen. In staat 3 zijn naast de gemiddelde inkomens van de verschillende huishoudenstypen waartoe personen kunnen behoren ook de Theilcoëfficiënten en de inkomensaandelen in procenten voor deze huishoudenstypen weergegeven. Gebruik makend van de decompositieformule kan met deze gegevens berekend worden dat de binnengroepsongelijkheid gelijk is aan 0,14 en de tussengroepsongelijkheid aan 0,01. Dit betekent dat 93 procent van de totale inkomensongelijkheid in Nederland (T=0,15) verklaard wordt door ongelijkheid bínnen de verschillende typen huishoudens. Slechts 7 procent is toe te schrijven aan ongelijkheid tussen huishoudenstypen. Deze resultaten zijn uiteraard anders wanneer de bevolking wordt ingedeeld in andere groepen (bijvoorbeeld naar leeftijd of naar voornaamste inkomensbron van het huishouden).

Staat 3
Decompositie van de Theilcoëfficiënt naar huishoudenstype, 2004

	Gemiddeld inkomen (μ_k)	Inkomensaandeel (S_k)	Theilcoëfficiënt (T_k)
	x 1 000		
Totaal	20,8	1,00	0,15
Eenpersoonshuishouden	16,9	0,12	0,18
Paar zonder kinderen	23,7	0,29	0,14
Paar met uitsluitend minderjarige kinderen	19,5	0,31	0,14
Paar met minstens één meerderjarig kind	24,7	0,18	0,11
Eenoudergezin	16,4	0,05	0,14
Overig	21,1	0,05	0,13

Bron: CBS (Inkomensstatistiek).

4. Criteria voor ongelijkheidsmaten

Voor een zo gevarieerd mogelijke beschrijving van de inkomensongelijkheid in een populatie kan uitgegaan worden van de meest gangbare criteria die gesteld worden aan ongelijkheidsmaten. Het betreft vijf criteria die veelvuldig in de literatuur worden aangetroffen.

1. Ongelijkheidsmaten dienen een positieve waarde te hebben. De reden voor deze eis is het gemak en de eenvoud van de interpretatie. Als alle inkomens gelijk zijn dan is de ongelijkheid nul. Iedere afwijking van een toestand van gelijkheid komt dan tot uitdrukking in een positieve waarde van de inkomensongelijkheidsmaat. Naarmate de inkomensverschillen in een populatie groter worden, wordt ook de waarde van de inkomensongelijkheidsmaat groter [7]. De bovengrens van een ongelijkheidsmaat is vast of variabel. Soms wordt deze bepaald door het aantal waarnemingen (zoals bij de Theilcoëfficiënt). Twee of meer verdelingen met verschillende omvang kunnen dan vergeleken worden door het standaardiseren van ongelijkheidsmaten met behulp van transformaties.
2. Een ongelijkheidsmaat dient symmetrisch ofwel anoniem te zijn. Bij verwisseling van twee willekeurige inkomens, blijft de maat onveranderd.
3. Ongelijkheidsmaten dienen schaalinvariant te zijn. Dit houdt dit in dat vermenigvuldiging van ieder inkomen met eenzelfde constante de mate van ongelijkheid onveranderd laat. Dit impliceert onafhankelijkheid van de meeteenheid, zodat bijvoorbeeld bij internationale vergelijkingen van inkomensongelijkheid niet met koersverhoudingen (euro/dollar) rekening gehouden hoeft te worden. Een ongelijkheidsmaat verandert bovendien niet wanneer alle inkomens met eenzelfde proportie (bijvoorbeeld 10 procent) toenemen. Hoewel rijkeren meer voordeel hebben van een proportionele verhoging, blijven de onderlinge verhoudingen van alle paren inkomens constant. Maten gebaseerd op verhoudingen van inkomens blijven hierbij dus constant.
4. De ongelijkheidsmaat moet groter worden indien inkomen wordt overgeheveld van een armer naar een rijker persoon. Omgekeerd dient de maat kleiner te worden wanneer inkomen wordt overgeheveld van een rijker naar een armer persoon, maar zo dat de rijkere niet armer wordt dan de armere. Dit criterium, geformuleerd door Pigou in 1912 en toegepast door Dalton in 1920, is een fundamenteel theorema in studies naar inkomensverdelingen (Pigou-Dalton criterium). Er kan onderscheid gemaakt worden naar de mate van gevoeligheid van de ongelijkheidsmaat voor de posities van de inkomens in de verdeling waartussen de overdracht plaatsvindt. De ratio 80/20 bijvoorbeeld is alleen gevoelig voor overdrachten naar het laagste quintiel vanuit de andere quintielen (en overdrachten naar het hoogste quintiel vanuit andere quintielen). Overdrachten tussen het tweede, derde en vierde quintiel onderling en overdrachten binnen het laagste of binnen het hoogste quintiel worden niet geregistreerd door de ratio 80/20. De maat voldoet daarom in beperkte mate aan het Pigou-Dalton criterium. Dat geldt ook voor de

[7] Voor sommige ongelijkheidsmaten neemt de waarde juist af, wanneer de inkomensverschillen in de populatie groter worden. Dit geldt bijvoorbeeld voor het aandeel van de 50 procent laagste inkomens in de totale inkomenssom.

Relatieve IKA. De Ginicoëfficiënt en de Theilcoëfficiënt voldoen beide aan het Pigou-Dalton criterium. Allison (1978) toont echter aan dat de Ginicoëfficiënt vooral gevoelig is voor overdrachten in het middengedeelte van de inkomensverdeling en minder voor overdrachten in de staarten van de verdeling. De Theilcoëfficiënt is juist gevoelig voor overdrachten in de gehele inkomensverdeling, maar vooral voor overdrachten bij de lage inkomens. De polarisatie-index tot slot voldoet niet aan het Pigou-Dalton criterium. Bij een inkomensoverdracht van een rijk persoon naar een minder rijk persoon in de bovenste helft van de inkomensverdeling neemt de index juist toe in plaats van af. De Ginicoëfficiënt neemt immers toe, terwijl het gemiddelde, de mediaan en het inkomensaandeel van de laagste helft van de bevolking gelijk blijven.

5. Decompositie van de inkomensongelijkheidsmaat heeft betrekking op de (gewenste) eigenschap dat de totale inkomensongelijkheid van een populatie kan worden uitgedrukt als een (met inkomensaandelen) gewogen som van de ongelijkheid binnen subgroepen en de ongelijkheid tussen de subgroepen. Van de besproken ongelijkheidsmaten heeft alleen de Theilcoëfficiënt deze eigenschap.

Schema 1 bevat een overzicht van de eigenschappen van de inkomensongelijkheidsmaten besproken in dit artikel. De Theilcoëfficiënt voldoet aan alle hierboven genoemde eigenschappen en is daarmee een zeer bruikbare maat. Op decompositie na voldoet de Ginicoëfficiënt eveneens aan de vijf eigenschappen. Beide coëfficiënten zijn daarom geschikte maten in studies naar inkomensongelijkheid. De ratio 80/20 en de Relatieve IKA voldoen niet aan alle eigenschappen, maar hebben als grote voordeel hun eenvoud en inzichtelijkheid. Daar de Ratio 80/20 vaak in Europees verband gebruikt wordt, is het voor de Inkomensstatistieken van het CBS een voor de hand liggende maat om naast de Gini- en Theilcoëfficiënt te presenteren. De polarisatie-index kan voor bepaalde onderzoeksdoeleinden een toegevoegde waarde hebben.

Schema 1
Overzicht van eigenschappen van diverse ongelijkheidsmaten

	Ondergrens	Bovengrens	Symmetrie	Schaal-invariantie	Pigou-Dalton criterium	Decompositie
Ratio 80/20	1 [1)]	onbegrensd	ja	ja	beperkt	nee
Relatieve IKA	0	onbegrensd	ja	ja	beperkt	nee
Ginicoëfficiënt	0	1	ja	ja	ja	nee
Polarisatie-index	0	1	ja	ja	nee	nee
Theilcoëfficiënt	0	ln(n) [2)]	ja	ja	ja	ja

1) Ervan uitgaande dat de inkomenstotalen van het hoogste en het laagste quintiel beide positief (of negatief) zijn.
2) Met *n* de steekproefomvang.

5. Samenvatting

In de literatuur is een groot aantal inkomensongelijkheidsmaten te vinden. Dit artikel bespreekt vijf gangbare ongelijkheidsmaten, de ratio 80/20, de relatieve interkwartielafstand, de Ginicoëfficiënt, de polarisatie-index en de Theilcoëfficiënt. Voor elk van de maten wordt aandacht besteed aan de specifieke eigenschappen en de interpretatie. In een beschrijving van inkomensongelijkheid in een populatie kunnen de Ginicoëfficiënt en de Theilcoëfficiënt elkaar completeren, omdat zij samen de ongelijkheid van de gehele inkomensverdeling beschouwen. Bovendien kan met de Theilcoëfficiënt de totale ongelijkheid in een populatie uitgesplitst worden naar een ongelijkheid tussen groepen en een ongelijkheid binnen groepen. Hiermee kan de totale ongelijkheid dus nader verklaard worden. De ratio 80/20 en de Relatieve IKA zijn vanwege hun eenvoud een goede aanvulling op de Ginicoëfficiënt en de Theilcoëfficiënt.

Literatuur

Allison, P.D. (1978). *Measures of inequality*, American Sociological Review, vol. 43, number 6, pp. 865–880.

Odink, J.G. en E. Van Imhoff (1984), *True versus measured Theil inequality*, Statistica Neerlandica 38, nr. 4, blz. 219–232.

Pen, J. (1971), Income distribution, Harmondsworth.

Theil, H. (1967), Economics and information Theory, Amsterdam.

Wolfson, M.C. (1997), *Divergent inequalities: theory and empirical results*, Review of Income and Wealth, series 43, number 4.

Tabel 1
Kengetallen ongelijkheid van het inkomen van personen, 2004

	Totale bevolking	Alleenstaanden	Leden van een meer-persoonshuishouden
Totaal (x 1 000)	15 954	2 368	13 586
Gemiddeld inkomen (x 1 000 euro)	20,8	16,9	21,5
Ongelijkheidsmaat			
Ratio 80/20	4,06	4,80	3,87
Relatieve IKA	0,57	0,57	0,55
Ginicoëfficiënt	0,27	0,29	0,26
Polarisatie-index	0,20	0,20	0,19
Theilcoëfficiënt	0,15	0,18	0,14

Bron: CBS (Inkomensstatistiek).