



Centraal Bureau voor de Statistiek

Centrum voor Beleidsstatistiek

**Documentatierapport Inkomenspanelonderzoek
(IPO) 1999V1**

VIDIV-nummer: 2007-428-MCB
Datum: 12 december 2007
Opgesteld door: Jan van der Worp

Bronvermelding

Publicatie van uitkomsten geschiedt door het onderzoeksbureau of de opdrachtgever op eigen titel. Verwijzing naar het CBS betreft uitsluitend het gebruik van de microdatabestanden. Dat wordt als volgt geformuleerd:

“Eigen berekening [*naam onderzoeksbureau, c.q. opdrachtgever*] op basis van bij het CBS beschikbaar gestelde microdatabestanden betreffende het Inkomenspanelonderzoek 1999, welke afkomstig zijn van het CBS.”

Versiegeschiedenis

Versie 1

071107 IPO 1999V1

071018 GBA-selectiebestand IPO 1999V1

De gebruiker dient rekening te houden met de volgende *do's & don'ts*:

- Uitsluitend de combinatie van RIN met SRTNUM 'R' of 'S' identificeert een persoon. Zie blz.20 voor een beschrijving.
- Omdat in het IPO al demografische informatie van alle personen (zowel kernpersonen als de overige personen) is opgenomen en de huishoudens binnen het IPO mogelijk op een andere wijze zijn samengesteld dan in het GBA-selectiebestand, wordt afgeraden van het GBA-selectiebestand gebruik te maken.
- Het IPO is een steekproefonderzoek. Daardoor hebben de uitkomsten statistische onnauwkeurigheden, waarmee bij het gebruik rekening gehouden moet worden. Zo kan de steekproefomvang te klein zijn om voor de te onderzoeken bevolkingsgroep betrouwbare gegevens op te leveren. Van even groot belang is echter dat de gebruiker weet op welke wijze de inkomens gemeten zijn en welke bewerkingen de basisgegevens hebben ondergaan in het verwerkingsproces naar het statistische eindbestand.
- Het CBS publiceert zelf alleen uitkomsten voor cellen met meer dan 20 duizend huishoudens of personen.
- Het bestand '071107 IPO 1999V1' bevat kernpersonen (de steekproefpersonen in engere zin) en de overige personen waarmee de kernpersonen samen een huishouden vormen. Longitudinale analyses kunnen alleen worden gebaseerd op de kernpersonen (AANDKP=1).
- Uitkomsten op basis van bestanden 1977-2000 wijken in een aantal opzichten af van uitkomsten op basis van de bestanden 2000 en verder. Dit is het gevolg van een herziening van de wijze van samenstelling van de Inkomensstatistiek. Ten opzichte van de oude reeks 1977-2000 zijn in de nieuwe reeks 2000 en verder onder meer de volgende verbeteringen aangebracht:
 - Voor het vaststellen van inkomsten uit vermogen zijn aanvullende bronnen gehanteerd;
 - De methode van ophogen is herzien, waardoor de aansluiting op de Huishoudensstatistiek is verbeterd;
 - Waar in eerdere publicaties werd ingedeeld naar kenmerken van het hoofd, is deze vervangen door de hoofdkostwinner;

- De indeling naar samenstelling van het huishouden en voornaamste inkomensbron van het huishouden is herzien.

Uitkomsten uit de oude reeks zijn hierdoor niet geheel vergelijkbaar met uitkomsten uit de nieuwe reeks. Voor het jaar 2000 zijn twee bestanden beschikbaar, één volgens de oude en één volgens de nieuwe methodologie. Met behulp hiervan kan de breuk die is ontstaan als gevolg van de herziening worden bepaald.

Inhoudsopgave

Bronvermelding.....	2
Versiegeschiedenis	3
1. Inleiding.....	8
2. Inhoud van de bestanden	9
2.1 Structuur van de bestanden	9
2.2 Overzicht van variabelen	9
2.3 Algemene toelichting.....	13
2.4 De basisgegevens	14
2.5 Steekproef, non-respons, weging	15
2.6 Item non-respons en imputatie.....	17
2.7 Compleetheid van de waarneming.....	18
2.8 Inkomenseenheid	19
2.9 De vergelijkbaarheid van de inkomensgegevens in de tijd.....	19
3. Beschrijving van de variabelen	21

1. Inleiding

Bij het Centrum voor Beleidsstatistiek (CvB) worden microdatabestanden van het Centraal Bureau voor de Statistiek (CBS) beschikbaar gemaakt voor gebruik door externe onderzoekers. Van alle personen die in het hier beschreven microdatabestand voorkomen, zijn gegevens uit de Gemeentelijke basisadministratie (GBA) geselecteerd. Deze selectie wordt het GBA-selectiebestand genoemd. Dit documentatierapport beschrijft de inhoud en structuur van het microdatabestand '071107 IPO 1999V1', dat onderdeel uitmaakt van de reeks 1977-2000. Het complete GBA-bestand wordt in een apart documentatierapport beschreven.

Hoofdstuk 2 beschrijft de structuur van het microdatabestand, geeft een opsomming van de variabelen en bevat een toelichting op het bestand met daarbij een indicatie van de kwaliteit. In hoofdstuk 3 wordt per variabele meer informatie gegeven. Voor de categoriale variabelen worden alle mogelijke scores en hun betekenis opgesomd.

Indien gewenst kan een bijlage geleverd worden met daarin de frequentieverdelingen van de categoriale variabelen: '071107 Bijlage 1 frequentietellingen IPO 1999V1'. In deze bijlage worden van de (semi-) continue variabelen de eerste en laatste vijf waarden gegeven.

2. Inhoud van de bestanden

2.1 Structuur van de bestanden

Microdatabestand

Een record in het microdatabestand bevat inkomensgegevens van een persoon uit de steekproef. Een persoon kan dus maar één record in het microdatabestand hebben. Het bestand bevat RIN en HHNUMMER als sleutel(s) waarmee gekoppeld kan worden met andere bestanden.

GBA-selectiebestand

Het bijbehorende GBA-selectiebestand bevat persoons- en huishoudensgegevens van alle personen die in het microdatabestand voorkomen en ook geregistreerd staan in de GBA. De persoonsgegevens bestrijken een periode van 1 januari 1995 tot 1 januari 2006. De huishoudensgegevens worden door het CBS berekend voor de stand op 1 januari. In het GBA-selectiebestand zijn huishoudensgegevens beschikbaar van 1998 tot en met 2006.

Zowel het microdatabestand als het GBA-selectiebestand is gesorteerd op respectievelijk SRTNUM en RIN. Met behulp van deze twee variabelen als koppelsleutel kan de gebruiker persoons- en huishoudensgegevens uit het GBA-selectiebestand toevoegen aan het microdatabestand, zodat deze gegevens in het onderzoek meegenomen kunnen worden.

Om RIN-nummers toe te kennen aan het microdatabestand is gebruik gemaakt van de meest recente versie van de GBA. Hierin is slechts een beperkt aantal variabelen aanwezig, daarom wordt dit bestand niet gebruikt voor het GBA-selectiebestand. Het GBA-selectiebestand maakt gebruik van gegevens tot 1 januari 2006. Hierdoor kunnen personen die na 1 januari 2006 naar Nederland zijn geïmmigreerd, niet worden teruggevonden in het GBA-selectiebestand.

2.2 Overzicht van variabelen

Onderstaand volgt een lijst van alle variabelen in het microdatabestand. Zie hoofdstuk 3 voor een nadere toelichting op de variabelen in het microdatabestand.

<i>Nummer</i>	<i>Variabele & Label</i>	<i>Format</i>	<i>Lengte</i>	<i>Decimalen</i>
1	SRTNUM Soort identificatiecode	Tekst	1	0
2	RIN Identificatienummer	Getal	9	0
3	HHNUMMER Encrypt hhnummer	Getal	9	0
4	HHLIDNR Huishoudlidnummer	Getal	2	0
5	PAARCODE Paarcode	Getal	1	0

<i>Nummer</i>	<i>Variabele & Label</i>	<i>Format</i>	<i>Lengte</i>	<i>Decimalen</i>
6	AANDKP Steekproefpersooncode	Getal	1	0
7	JAAR Onderzoeksjaar	Getal	4	0
8	FAKTOR Ophoogfaktor	Getal	9	0
9	EQUI Equivalentiefactor	Getal	3	0
10	SRTADR Soort adres	Getal	1	0
11	INKHH Inkomen huishouden bekend	Getal	1	0
12	INKHL Inkomen persoon bekend	Getal	1	0
13	PROVC Provinciecode	Getal	2	0
14	STEDGEM Stedelijkheid	Getal	1	0
15	URB71 Urbanisatiegraad	Getal	2	0
16	LFT Leeftijd 31-12	Getal	3	0
17	GBS Geslacht, burgerlijke staat	Getal	1	0
18	GEB_LAND Geboorteland	Getal	4	0
19	GEB_L_MO Geboorteland moeder	Getal	4	0
20	GEB_L_VA Geboorteland vader	Getal	4	0
21	ALLOCHT Allochtooncode	Getal	2	0
22	SECCOAL Sociaal-economische categorie	Getal	2	0
23	POSHH Positie in het huishouden	Getal	1	0
24	WEKEN Aantal weken inkomen	Getal	2	0
25	SAMHH Samenstelling huishouden	Getal	2	0
26	AHL Aantal huishoudensleden	Getal	2	0
27	BBIHALG Belangrijkste bron van inkomen huishouden	Getal	1	0
28	BESTINK Besteedbaar inkomen	Getal	9	0
29	BESTINKH Besteedbaar inkomen huishouden	Getal	9	0
30	BELIE Belastbaar inkomen excl. eenmalig inkomen	Getal	9	0
31	HC_AOW Huishoudenstype AOW-huishouden	Getal	2	0
32	HC_ETV Huishoudenstype een-en tweeverdieners	Getal	1	0
33	PPOS_ETV	Getal	1	0

<i>Nummer</i>	<i>Variabele & Label</i>	<i>Format</i>	<i>Lengte</i>	<i>Decimalen</i>
	Positie in het huishouden een- en tweeverdieners			
34	HHT_LI	Getal	2	0
	Huishoudenstype Laag inkomen			
35	HC_T	Getal	1	0
	Code laag inkomen actuele jaar t			
36	HC_T_1	Getal	1	0
	Code laag inkomen jaar t-1			
37	HC_T_2	Getal	1	0
	Code laag inkomen jaar t-2			
38	HC_T_3	Getal	1	0
	Code laag inkomen jaar t-3			
39	NI_T	Getal	9	0
	Netto inkomen actuele jaar t			
40	NI_T_1	Getal	9	0
	Netto inkomen jaar t-1			
41	NI_T_2	Getal	9	0
	Netto inkomen jaar t-2			
42	NI_T_3	Getal	9	0
	Netto inkomen jaar t-3			
43	BIBM_T	Getal	6	0
	Besteedbaar inkomen in % van beleidsmatig minimum actuele jaar t			
44	BIBM_T_1	Getal	6	0
	Besteedbaar inkomen in % van beleidsmatig minimum actuele jaar t-1			
45	BIBM_T_2	Getal	6	0
	Besteedbaar inkomen in % van beleidsmatig minimum actuele jaar t-2			
46	BIBM_T_3	Getal	6	0
	Besteedbaar inkomen in % van beleidsmatig minimum actuele jaar t-3			
47	YBLN	Getal	9	0
	Bruto loon of salaris			
48	YDIN	Getal	9	0
	Loon directeuren NV/BV			
49	YWUO	Getal	9	0
	Winst uit onderneming			
50	YFRL	Getal	9	0
	Freelance inkomen en dergelijke			
51	YECH	Getal	9	0
	Economische huurwaarde eigen woning			
52	AHYP	Getal	9	0
	Betaalde hypotheekrente			
53	APEW	Getal	9	0
	Betaalde pacht e.d. i.v.m. eigen woning			
54	YREW	Getal	9	0
	Restant inkomsten eigen woning			
55	YREN	Getal	9	0
	Ontvangen rente			
56	ABRE	Getal	9	0
	Betaalde rente			
57	YDIV	Getal	9	0
	Inkomsten uit effecten			
58	YONG	Getal	9	0
	Inkomsten uit onroerend goed			

<i>Nummer</i>	<i>Variabele & Label</i>	<i>Format</i>	<i>Lengte</i>	<i>Decimalen</i>
59	YOVI Andere vermogensinkomsten	Getal	9	0
60	YWVU Werkloosheidsuitkeringen	Getal	9	0
61	YWAC Wachtgelduitkeringen	Getal	9	0
62	YZWU Ziektewetuitkeringen	Getal	9	0
63	YAOU Arbeidsongeschiktheidsuitkeringen	Getal	9	0
64	YPLU Pensioen- en lijfrenteuitkeringen (excl AOW)	Getal	9	0
65	YAOW AOW	Getal	9	0
66	YAWW AWW/Anw	Getal	9	0
67	YABI Abw, Bijstand, IOAW, IOAZ	Getal	9	0
68	YASU Andere uitkeringen sociale verzekeringen	Getal	9	0
69	YALE Alimentatie van echtgenoot	Getal	9	0
70	YSTB Studie: (aanvullende) beurs	Getal	9	0
71	YTSK Tegemoetkoming studiekosten	Getal	9	0
72	YKB_ Kinderbijslag	Getal	9	0
73	YHRS Individuele huursubsidie	Getal	9	0
74	YPEW Premie eigen woningbezit	Getal	9	0
75	YGBB Geen inkomensbestanddelen bekend	Getal	9	0
76	YGBO Gedeelte van de inkomensbestanddelen onbekend	Getal	9	0
77	PWVG Werkgeverspremies ZW, WW, WAO	Getal	9	0
78	PZFG Werkgeverspremies ZFW, IZA, IZR, DGVP	Getal	9	0
79	PIPB Werknemerspremies pensioen, VUT en dergelijke	Getal	9	0
80	PWVV Werknemerspremies ZW, WW, WAO	Getal	9	0
81	PZWF Werknemerspremies ZW, IZA, IZR, DGVP	Getal	9	0
82	ALIB Loon- en inkomstenbelasting	Getal	9	0
83	APVV Premies sociale verzekeringen	Getal	9	0
84	AVB_ Vermogensbelasting	Getal	9	0
85	PPAR	Getal	9	0

Nummer	Variabele & Label	Format	Lengte	Decimalen
	Particuliere ziektekostenpremies			
86	PZFN Nominale premies ZFW	Getal	9	0
87	PAWN Nominale premies AWBZ	Getal	9	0
88	PPVZ Vrijwillige verzekering tegen ziekte, invaliditeit e.d.	Getal	9	0
89	PPEL Vrijwillige pensioen of lijfrentepremies	Getal	9	0
90	ABAL Alimentatie aan ex-echtgenoot	Getal	9	0
91	AVBL Verhaalde bijstand	Getal	9	0
92	ABUL Buitengewone lasten	Getal	9	0
93	AGIF Aftrekbare giften	Getal	9	0
94	AZLF Zelfstandigenaftrek	Getal	9	0
95	YSPL Spaarloon	Getal	9	0
96	YHWF Huurwaardeforfait	Getal	9	0
97	AMEE Meewerkaftrek	Getal	9	0
98	YAUT Auto werkgever	Getal	9	0
99	YFWI Fiscale winst	Getal	9	0

2.3 Algemene toelichting

Informatievoorziening Inkomenspanelonderzoek

Inkomensgegevens worden veelvuldig gebruikt op een groot terrein van beleidsaandachtspunten, zoals de inkomensongelijkheid, de omvang en samenstelling van de groep met lage inkomens en andere doelgroepen en de verschillen in bestedingsmogelijkheden van groepen. Het Inkomenspanelonderzoek (IPO) van het CBS heeft als doel te voorzien in deze informatie. De resultaten worden via Statline in tabelvorm beschikbaar gesteld en analyses worden gepubliceerd in een aantal jaarboeken van het CBS en de Sociaaleconomische Trends (voorheen: Sociaal-economische maandstatistiek). Daarnaast worden tabellen op aanvraag samengesteld.

Via deze publicatiemedia wordt een groot aantal gebruikers bediend. Voor een aantal gebruikers is dit echter niet voldoende. Voor deze gebruikers is de mogelijkheid gecreëerd onder bepaalde voorwaarden met de microdata van het IPO te werken.

Onderzoeksjaren

Voor on-site werken zijn standaard IPO-bestanden beschikbaar van de onderzoeksjaren 1977, 1981, 1985 en vanaf 1989 jaarlijks. Omdat het IPO vanaf 1989 is opgezet als een panelonderzoek zijn de bestanden 1989 en later geschikt voor longitudinaal onderzoek. Voor de onderzoeksjaren 1977, 1981 en 1985 is dit niet mogelijk. Wel zijn er verbanden tussen de onderzoeken 1981 - 1985 - 1989 omdat voor de onderzoeken 1981 en 1985 gewerkt is met een vaste adressensteekproef, terwijl 1989 grotendeels getrokken is uit het bestand 1985. Het onderzoek 1977 is een aparte steekproef die onafhankelijk is van de vervolgonderzoeken.

2.4 De basisgegevens

Bronnen

Het Inkomenspanelonderzoek maakt uitsluitend gebruik van gegevens die ontleend zijn aan administraties. De belangrijkste dataleverancier is de belastingdienst. Deze verstrekt voor alle steekproefpersonen:

- Demografische gegevens (geboorteland vanaf 1995)
- De basisgegevens voor de heffing van de inkomstenbelasting (aangiftebiljet, vastgesteld belastbaar inkomen enz.)
- De basisgegevens van de loonbelastingkaarten die door inhoudingsplichtigen naar de fiscus gestuurd zijn (veelal elektronisch)
- Gegevens over ontvangen rente (vanaf 1995)

Daarnaast worden gegevens verzameld uit de registraties van de huursubsidie en de studiefinanciering (TS18- en TS18+, WSF). Resterende ontbrekende inkomensbestanddelen zijn berekend (kinderbijslag en premies) of worden niet opgenomen (bijvoorbeeld een deel van de inkomensoverdrachten tussen huishoudens, zwarte inkomens en dergelijke).

Inkomensvariabelen

Op persoons- en huishoudensniveau zijn het besteedbaar en het belastbaar inkomen in de on-site basisbestanden opgenomen en is een vijftigtal inkomensbestanddelen onderscheiden. Met behulp hiervan kunnen andere inkomensbegrippen worden afgeleid.

Overige gegevens van personen en huishoudens

Per persoon zijn –naast de inkomensgegevens- persoons- en huishoudenskenmerken en regiovariabelen opgenomen; daarnaast een aantal afgeleide variabelen en enkele technische variabelen (bijvoorbeeld ophoogfactor).

2.5 Steekproef, non-respons, weging

Populatie

De steekproef is getrokken uit de bevolking die op 31 december van het onderzoeksjaar behoort tot de bevolking van Nederland. Mensen die in de loop van het jaar zijn overleden of geëmigreerd zijn buiten beschouwing gebleven. Ook mensen die niet in het bevolkingsregister waren ingeschreven (zoals NAVO-militairen en diplomatieke vertegenwoordigers, illegalen en vreemdelingen) zijn niet in het inkomensonderzoek opgenomen. Voor de jaren 1977, 1981 en 1985 ontbreekt verder de zogenaamde 'varende en rijdende' bevolking. Het gaat hier om mensen zonder vast woonadres of verblijfplaats.

Steekproeftrekking 1977 en 1981

Voor 1977 en 1981 is de steekproef getrokken uit een adressenbestand van het CBS waarvan de huishoudenssamenstelling bekend was op basis van de Arbeidskrachtentellingen 1977 en 1981. Dit bestand is aangevuld met de adressen van instellingen en tehuizen. Voor de beide jaren zijn andere adressenbestanden gebruikt. Bij de ophoging van de adressen is rekening gehouden met nieuwbouw.

De steekproeffractie voor beide jaren is 3,3%.

In deze jaren is de gehuwde vrouw in het onderzoek opgenomen, indien de echtgenoot op 31 december op het steekproefadres woonde. In de loop van het jaar overleden gehuwde vrouwen zijn in de steekproef opgenomen.

Steekproeftrekking en weging 1985.

Na het onderzoek 1981 zijn alle mutaties bijgehouden die op de steekproefadressen van 1981 plaats hebben gevonden. De personen die ultimo 1985 op een van deze adressen woonden, zijn in de steekproef 1985 opgenomen. Daarnaast heeft een aanvulling plaatsgevonden voor de nieuwe adressen tussen 1981 en 1985. Vanaf 1985 wordt bij de weging extra opgehoogd naar de bevolking per Corop-gebied en geslacht, burgerlijke staat en leeftijd (zie hierna bij weging vanaf 1985). Dit is achteraf ook toegepast op het onderzoek 1981.

Vanaf 1985 zijn de in de loop van het jaar overleden gehuwde vrouwen niet meer in het onderzoek opgenomen.

Steekproeftrekking 1989

Vanaf 1984 is de opzet van het inkomensonderzoek van een groot structuuronderzoek veranderd in een panelonderzoek, het Inkomenspanelonderzoek (IPO). Ieder jaar worden gegevens verzameld van in beginsel dezelfde personen. Over de jaren 1984-1986 bevatte de IPO-steekproef ongeveer 5 duizend personen. Met ingang van 1987 is de steekproef in fasen opgebouwd tot ongeveer 75 duizend personen in 1989 en volgende jaren.

Voor 1989 zijn de steekproefpersonen getrokken uit de personen van 10 jaar en ouder van het inkomensonderzoek 1985. De steekproef is alleen gecorrigeerd voor uitstroom in deze periode. Voor 1989 waren nog geen voorzieningen getroffen om de steekproef aan te vullen voor instroom via immigratie.

Steekproeftrekking 1990 t/m 1992

Vanaf 1990 is het IPO voortgezet, maar wordt er gecorrigeerd voor de instroom via immigratie; de steekproef is in 1990 aangevuld voor de instroom vanaf 1985. Voor 1991 tot 1993 werd de steekproef steeds aangevuld met een bijtrekking uit de personen van 15 jaar en ouder en de immigranten.

De bijtrekking van immigranten heeft ten opzichte van 1989 een lichte daling van het gemiddelde inkomen in 1990 tot gevolg gehad (circa 100 gulden), omdat immigranten een relatief laag inkomen hebben. Ook is er een effect geweest op de inkomensongelijkheid. De waargenomen trend 1989-1990 is echter wel correct.

Steekproeftrekking 1993 tot heden

Vanaf 1993 is de steekproef voor het panel verlegd van personen van 15 jaar en ouder naar personen ongeacht de leeftijd. Voor de bijtrekking van de personen van jonger dan 15 jaar is uit personen jonger dan 15 jaar uit de onderzoeken 1989 tot en met 1992 achteraf een aantal als steekproefpersoon aangewezen. Sinds 1993 wordt elk jaar bijgetrokken uit de instroom in de bevolking (geboorten en immigratie) en gecorrigeerd voor uitstroom (overlijden en emigratie).

Steekproefpersonen en overige personen in het onderzoek

De inkomensgegevens worden niet alleen verzameld voor de steekproefpersonen maar ook voor de personen die in het betreffende onderzoeksjaar behoren tot hetzelfde huishouden als de steekproefpersoon. Op deze wijze worden het huishoudensinkomen en kenmerken van het huishouden waartoe de steekproefpersoon behoort, afgeleid.

Weging vanaf 1985

De steekproef is vanaf 1985 opgehoogd door elk huishouden een gewicht te geven omgekeerd evenredig aan de trekkingskans van het huishouden. De personen krijgen het gewicht van het huishouden waartoe zij behoren. Om de nauwkeurigheid van de uitkomsten verder te verbeteren en aansluiting bij de bevolkingsstatistiek te krijgen, is bovendien poststratificatie toegepast naar leeftijd, geslacht, burgerlijke staat (tot 1990) en regio. Voor de jaren 1990-2000 wordt op vergelijkbare wijze aansluiting gezocht bij de Jaarlijkse huishoudensstatistiek van het CBS, met stratificatie naar huishoudensgrootte en leeftijd van het hoofd. Bij deze aansluiting zijn de categorieën een- en tweepersoonshuishoudens waarvan het hoofd jonger dan 35 jaar is, samengevoegd. Het IPO heeft namelijk een te hoog aantal eenpersoonshuishoudens omdat alle ongehuwde studenten die niet bij (een van) hun ouders wonen als eenpersoonshuishoudens worden aangemerkt. Voor de Jaarlijkse huishoudensstatistiek, die gebaseerd is op enquêtes, geeft deze groep zelf aan of ze al of niet met anderen een huishouden vormen. Herwegen zonder rekening te houden met dit definitieverschil leidt tot slechte uitkomsten en is daarom achterwege gebleven.

Non-respons

De beschreven populatie wordt door het IPO in principe volledig waargenomen. In een aantal gevallen wordt het inkomen echter in het geheel niet of voor een deel

waargenomen. In de bestanden zijn deze huishoudens terug te vinden met de variabele *inkomen huishouden bekend* (code 0 geeft aan dat het waargenomen inkomen ontbreekt of onvolledig is). Daarnaast kan via de variabele *aantal weken inkomen* nagegaan worden op welk deel van het jaar het inkomen betrekking heeft. Huishoudens waarvan geen enkel lid een volledig jaarinkomen heeft gehad, worden in bepaalde cijferopstellingen als aparte groep zichtbaar gemaakt. Hiervoor is niet gecorrigeerd. Nader onderzoek op dit punt is nog lopende.

2.6 Item non-respons en imputatie

De bronnen waarmee de inkomensgegevens zijn samengesteld leveren niet direct alle afzonderlijke inkomensbestanddelen. Door middel van aanvullende berekeningen is hiervoor zoveel mogelijk gecorrigeerd. In een aantal gevallen is dit niet volledig mogelijk, waardoor het besteedbare inkomen onderschat is.

Aanvullende berekeningen voor ontbrekende bestanddelen

Voor ontbrekende bestanddelen zoals kinderbijslag en de premies voor de werknemersverzekeringen en de ziektekostenverzekering zijn aanvullende berekeningen gemaakt. Ook de inkomsten uit eigen woning zijn via een rekenmodel geschat (economische huurwaarde). Bij de beschrijving van de afzonderlijke inkomensbestanddelen is dit aangegeven.

Ontbrekende inkomensbestanddelen en onzuiverheid van bestanddelen

Niet in alle gevallen kon een acceptabele berekeningsmethode worden ontwikkeld. Dit geldt bijvoorbeeld voor pensioenpremies. Deze bestanddelen zijn daarom niet geïmputeerd.

Een ander punt is dat bepaalde inkomensvormen waargenomen zijn samen met een ander bestanddeel, waardoor de juiste inkomensbron niet of onvolledig is te onderkennen. Dit geldt bijvoorbeeld voor ZW- en WAO-uitkeringen die -samen met loon- via de werkgever uitbetaald worden. Ook inkomens via de toeslagenwet worden niet afzonderlijk waargenomen, maar in combinatie met de uitkering zelf.

Door deze tekortkomingen wordt het totaal besteedbaar inkomen niet vertekend, maar wel de afzonderlijke componenten.

Niet of onvolledig waargenomen inkomens

Naast onzuiverheden op bestanddeleniveau worden ook bestanddelen niet waargenomen die wel deel uitmaken van het gecoördineerde besteedbaar inkomen. Voorbeelden hiervan zijn kinderalimentatie en andere inkomensoverdrachten tussen huishoudens. Het inkomen van betalende huishoudens wordt hierdoor overschat, dat van ontvangende huishoudens onderschat. Een tweede onderschatting ontstaat doordat voor de personen die geen aanslag inkomstenbelasting ontvangen alleen het inkomen is waargenomen dat door werkgevers of uitkeringsinstanties is uitbetaald. Vermogensinkomsten bij deze groep zijn niet bijgeschat. Aannemelijk is dat het hier in het algemeen om geringe bedragen zal gaan: voor grotere bedragen is immers aangifte verplicht.

Opmerking: correctie berekening particuliere ziektekostenpremies in 1981

Bij het samenstellen van de bestanden 1981 bleek er een onvolkomenheid te zitten bij de ziektekostenpremies voor bijstandontvangers. In een aantal gevallen was ten onrechte particuliere ziektekostenpremie berekend in plaats van de (lagere) premie ziekenfondswet. Hiervoor is in de bestanden gecorrigeerd.

2.7 Compleetheid van de waarneming

In hoofdstuk 2 is aangegeven hoe de steekproef is getrokken en hoe met (item) non-respons is omgegaan. Hier worden de consequenties hiervan voor de volledigheid van de waarneming samengevat.

Volledigheid waarneming bevolking

Het IPO beschrijft de bevolking en huishoudenssituatie aan het einde van het onderzoeksjaar. Ambassadepersoneel en andere personen niet opgenomen in het bevolkingsregister blijven buiten beschouwing. Voor 1977 – 1985 ontbreekt daarnaast de varende en rijdende bevolking. De afstemming op de bevolkings- en huishoudensstatistiek is vanaf 1989 als extra element in de weging ingebouwd.

Huishoudens zonder (waargenomen) inkomen

Voor een aantal huishoudens is de inkomenswaarneming niet volledig geweest of zijn geen inkomens waargenomen. Ook huishoudens met uitsluitend inkomens in de vorm van kinderbijslag, huursubsidie of tegemoetkoming studiekosten worden hiertoe gerekend. Deze huishoudens, die met de variabele *inkomen huishouden bekend* zijn te onderkennen, zijn volledigheidshalve opgenomen. In de gepubliceerde CBS-cijfers blijft deze groep steeds buiten beschouwing. Bij deze huishoudens zijn alle voorkomende bestanddelen onderdrukt.

Opgemerkt wordt dat voor het bepalen van deze variabele een hulpinkomen wordt afgeleid. Indien dit nul is, wordt aan dat huishouden de code *onbekend inkomen* toegekend. In een beperkt aantal gevallen kan het voorkomen dat door het plussen en minnen het saldo op 0 uitkomt, terwijl er toch bestanddelen zijn die deel uitmaken van een geldig inkomen. In het verleden is hiervoor een andere werkwijze gehanteerd; er werd toen gekeken naar het voorkomen van bepaalde bestanddelen. In de bestanden die nu beschikbaar zijn, is de huidige werkwijze toegepast. Hierdoor zijn kleine verschillen ontstaan met eerder gepubliceerde cijfers.

Inkomensbestanddeel onbekend

Door tekortkomingen in het basismateriaal kan van bepaalde inkomensbronnen niet de aard aangegeven worden zodat ze niet ingedeeld kunnen worden naar bestanddeel. Bovendien kunnen er verschillen zijn in geteld belastbaar inkomen vanuit de relevante inkomensbestanddelen en het vastgestelde belastbaar inkomen. In dat geval is uitgegaan van het vastgestelde belastbaar inkomen. Het verschil is bij *inkomensbestanddeel onbekend* opgenomen.

Onvolledige inkomensbestanddelen

Een aantal inkomensbestanddelen kan in de huidige opzet alleen waargenomen worden via het aangiftebiljet voor de inkomstenbelasting. Het gaat hier bijvoorbeeld om rente-inkomsten en de inkomsten uit eigen woning. De inkomensgegevens van de personen waarvoor uitsluitend gegevens zijn waargenomen via de loonbelastingkaart blijven beperkt tot de inkomens die via deze weg bekend zijn. Voor zover bij deze personen nog andere inkomsten voorkomen, is er sprake van enige onderschatting bij deze andere inkomsten.

2.8 Inkomenseenheid

Binnen de Inkomensstatistiek worden zowel personen als huishoudens als inkomenseenheden gebruikt. Voor bepaalde cijferopstellingen worden alle personen afzonderlijk als inkomenseenheid beschouwd met uitzondering van personen die samen een paar vormen; leden van een paar worden dan samen als één inkomenseenheid behandeld. Binnen de bestanden is aan alle leden van het huishouden dezelfde sleutelvariabele toegekend, te weten het CBS-nummer van de steekproefpersoon. Leden van een paar (of paren) binnen een huishouden zijn te onderkennen aan de paarcodes.

2.9 De vergelijkbaarheid van de inkomensgegevens in de tijd.

De definities van inkomen en de afleidingen van de inkomensbegrippen konden binnen de inkomensstatistieken over de periode 1977-2000 niet steeds gelijk gehouden worden. De belangrijkste oorzaken hiervan zijn de veranderingen in de regelgeving en de basisgegevens vanaf 1977. Voorbeelden hiervan zijn de veranderingen in uitbetaling van de AOW in 1985, de invoering van het stelsel van studiefinanciering in 1986 en de belastingherziening in de Oort-wetgeving van 1991. De basisgegevens worden nu vrijwel volledig elektronisch aangeleverd, terwijl dit in de jaren '70 en '80 vrijwel volledig in papiervorm plaats vond. Daarnaast spelen de gewijzigde onderzoeksopzet als gevolg van de verlegging van structuur- naar panelonderzoek en de inhoudelijke verbeteringen om te komen tot gecoördineerde inkomensbegrippen binnen het CBS een rol.

Omdat het doel van de Inkomensstatistiek het beschrijven van de ontwikkelingen in de inkomensverdeling is, is steeds veel aandacht besteed aan het corrigeren voor deze verschillen en het samenstellen van tijdreeksen. Bij het samenstellen van de bestanden voor on-site analyses zijn de uitkomsten van deze aanvullende bewerkingen volledig meegenomen. Op een aantal kleine onderdelen konden nog verdere verbeteringen aangebracht worden in de bestanden 1977-1985 (wijzigingen in het afleiden van al of niet huishoudensinkomen bekend en de berekening van particuliere ziektekostenpremies).

De gegevens van de basisbestanden zijn om deze reden voor een groot aantal inkomensanalyses in de tijd geschikt. Voor een aantal wijzigingen is niet gecorrigeerd, bijvoorbeeld voor de individualisering van de AOW in 1985. Voor bepaalde specifieke analyses is het daardoor onvermijdelijk eerst nog verdere correcties op de basisgegevens toe te passen. Dit wordt aan de gebruiker zelf overgelaten.

Kwaliteit

Het microdatabestand is geen publicatiebestand. Dit betekent dat voor veel variabelen door het CBS weinig of geen controles en correcties hebben plaatsgevonden. Op de variabelen die wel door het CBS bekeken zijn, zijn op grond van de beschikbare informatie correcties uitgevoerd. Deze correcties zijn als zodanig in het microdatabestand overgenomen.

Het microdatabestand over het jaar 1999 is samengesteld op basis van het jaarbestand van het IPO. Dat jaarbestand is bij de verantwoordelijke sector van het CBS gecontroleerd en gecorrigeerd. Er is bij het voorbereiden van het microdatabestand gezorgd voor een consistente relatie tussen het databestand en de codeomschrijving. Codes waarvan de betekenis niet meer achterhaald kon worden, zijn op een aparte code 'Ongeldige score' gezet.

Omdat het geen officieel vastgesteld publicatiebestand betreft, heeft dit tot gevolg dat de bestanden niet noodzakelijk exact overeenkomen met Statline, de elektronische databank van het CBS (statline.cbs.nl).

Een goede stelregel om in eerste instantie te kunnen inschatten of een variabele in het microdatabestand van goede kwaliteit is, is te kijken naar Statline. Als een kenmerk is opgenomen in Statline, is dat over het algemeen van goede kwaliteit. Als een kenmerk niet is opgenomen, dan is de kans groot dat deze variabele van matige of slechte kwaliteit is.

De gegevens op Statline kunt u (voor de periode 1990-2000) vinden onder Inkomens en bestedingen → Archief.

3. Beschrijving van de variabelen

In dit hoofdstuk wordt per variabele meer informatie gegeven. Voor een beschrijving van de rest van de variabelen wordt verwezen naar '071107 Bijlage 2 Beschrijving variabelen IPO 1990V1'.

1	SRTNUM
	Soort identificatiecode
R	RIN-persoon (persoon is gevonden in GBA)
S	Versleuteld sofinummer
	<p>Bronbestanden die aan het CBS geleverd worden, bevatten doorgaans als persoonsidentificator het sofinummer. Het sofinummer wordt als zeer identificerend beschouwd en daarom vervangt het CBS dit nummer door een intern persoonsidentificatienummer, het betekenisloze RIN-nummer.</p> <p>Het gedocumenteerde bestand is voorzien van de door het CBS toegevoegde variabelen soort identificatiecode (SRTNUM) en identificatienummer (RIN), die samen het sofinummer vervangen. Records met een sofinummer dat is teruggevonden in de GBA, zijn voorzien van een geldig RIN-nummer. Deze records hebben de waarde 'R' op SRTNUM gekregen.</p> <p>De overige records bevatten een sofinummer dat niet is teruggevonden in de GBA. Afhankelijk van de situatie hebben zij de waarde 'F' of 'S' op SRTNUM gekregen.</p> <p>Indien het sofinummer geldig is maar niet wordt teruggevonden in de GBA, dan wordt het sofinummer versleuteld via een vaste formule en krijgt SRTNUM de waarde 'S'.</p> <p>De combinatie van SRTNUM 'R' en RIN identificeert een uniek persoon. Evenzo kan worden aangenomen dat personen met SRTNUM 'S' een unieke waarde op RIN hebben gekregen. De meest voorkomende reden dat zo'n persoon wel een geldig sofinummer heeft maar niet in de GBA staat ingeschreven, is namelijk dat hij of zij in het buitenland woont. Door immigratie kan een persoon het ene jaar niet en het volgende jaar wel in de GBA staan ingeschreven.</p> <p>Als een bestand meer records per persoon bevat, dan kan voor SRTNUM 'R' en SRTNUM 'S' worden aangenomen dat records met eenzelfde RIN-nummer dezelfde persoon aanduiden. Binnen één bestand is het geen probleem om records met SRTNUM 'R' of 'S' te aggregeren op persoonsniveau. Records met SRTNUM 'R' of 'S' kunnen ook worden gekoppeld aan records in een ander bestand met dezelfde waarden op SRTNUM en RIN: deze records hebben betrekking op dezelfde persoon. De gebruiker kan met behulp van SRTNUM en RIN ook gegevens uit het GBA-selectiebestand toevoegen aan het bestand.</p> <p>Er moet echter een kanttekening worden gemaakt bij het koppelen van verschillende bestanden. Een persoon kan namelijk in het ene bestand de waarde 'R' op SRTNUM hebben en in het andere bestand de waarde 'S', en vice versa. Deze situatie kan zich voordoen als een persoon ten tijde van het ene bestand wél in de GBA stond ingeschreven en ten tijde van het andere bestand niet. Omdat de records in dat geval ook een ander RIN-nummer hebben gekregen, zijn zij niet aan elkaar te koppelen.</p>
2	RIN
	Identificatienummer

Persoonsidentificatienummer, ter vervanging van het sofinummer. Zie toelichting bij SRTNUM.