



Paper

Monitor loonverschillen mannen en vrouwen, 2016

Technische toelichting

Linda Muller
Jeanine Floris
Katja Chkalova
Ben Dankmeyer
Maarten Bloem

November 2018

Inhoud

Inleiding 3

1. **Populatie 3**
2. **Onderzoeksmethode 3**
3. **Operationalisering 6**
4. **Bepaling gecorrigeerde beloningsverschillen 7**
5. **Oaxaca-Blinder decompositie methode 10**
6. **Ophogen van steekproeftotalen 11**
7. **Kwaliteit van de uitkomsten 12**
8. **Vergelijking met cijfers uit eerder onderzoek 13**

Bijlage 1 14

Bijlage 2 18

Inleiding

In het onderzoek 'Monitor loonverschillen mannen en vrouwen' staan beloningsverschillen tussen mannen en vrouwen bij de overheid en in het bedrijfsleven centraal. Deze technische toelichting bij het onderzoek bevat een toelichting op de gebruikte onderzoeksmethode en de gebruikte bronbestanden. Eerst wordt besproken hoe het onderzoeksbestand voor peiljaar 2016 is samengesteld en hoe de verschillende begrippen in dit onderzoek zijn geoperationaliseerd. Vervolgens volgt een toelichting op de regressieanalyse en de Oaxaca-Blinder decompositie methode, waarna zal worden ingegaan op de kwaliteit van de uitkomsten.

In dit onderzoek worden de resultaten over 2016 vergeleken met die van eerdere jaren vanaf 2008. Er is hierbij zoveel mogelijk aangesloten bij het eerdere onderzoek over beloningsverschillen van De Mooij et al. (2010), Geerdinck et al. (2012), Verschuren et al. (2014) en Muller et al. (2016). In bijlage 1 zijn de belangrijkste uitkomsten uit de regressiemodellen samengevat. Tot slot bevat bijlage 2 een overzicht van gebruikte bronnen, begrippen, afkortingen en literatuur.

1. Populatie

De populatie van dit onderzoek bestaat uit alle banen van vier uur of meer per maand, van werknemers woonachtig in Nederland en tussen de 15 tot en met 64 jaar. Het peilmoment is de laatste vrijdag van september 2016. Een persoon kan meer dan één baan hebben en telt dan meer dan één keer mee in de onderzoekspopulatie. De populatie is opgesplitst in banen bij de overheid en banen in het bedrijfsleven. Deze twee subpopulaties zijn apart onderzocht.

2. Onderzoeksmethode

Samenstellen onderzoeksbestand

Voor de bepaling van het aantal banen en de berekening van uurlonen en beloningsverschillen is een onderzoeksbestand samengesteld met als basis de baaninformatie uit het Stelsel van Sociaal-statistische Bestanden (SSB) en drie jaargangen van de Enquête Beroepsbevolking (EBB). De baaninformatie is gebaseerd op de loonaangiftes die de Belastingdienst ontvangt van werkgevers. Het onderzoeksbestand bevat alleen de banen van werknemers uit de loonaangifte die gekoppeld konden worden aan een persoon uit een van deze drie EBB-jaargangen.

Het onderzoeksbestand is verrijkt met gegevens over de banen, informatie over de werknemers zelf en over de bedrijven waar zij werken. Deze gegevens zijn afkomstig uit verschillende bronnen. Demografische kenmerken, zoals geslacht, leeftijd en herkomstgroepering en inkomensgegevens van de partner zijn afkomstig uit het SSB. Onderwijsniveau, beroepsniveau en handicap of chronische ziekte zijn afkomstig uit de EBB. Informatie over de winstgevendheid van ondernemingen (waar personen werken), is met behulp van het Algemeen Bedrijven Register (ABR) toegevoegd uit de Statistiek Financiën

2.1 Kerncijfers onderzoekspopulaties 2008 tot en met 2016

	Bedrijfsleven									
	2008		2010		2012		2014		2016	
	man	vrouw	man	vrouw	man	vrouw	man	vrouw	man	vrouw
	x 1 000									
Aantal banen	3 584	3 085	3 468	3 113	3 452	3 156	3 379	3 099	3 473	3 156
waarvan										
Voltijd	2 573	775	2 474	727	2 375	678	2 290	641	2 334	639
Deeltijd, 12 uur of meer	610	1 646	588	1 691	664	1 758	668	1 736	728	1 820
Deeltijd, minder dan 12 uur	401	664	406	695	413	720	421	723	412	696
Leeftijd										
15 tot 23 jaar	515	539	501	552	472	524	448	502	485	526
23 tot 35 jaar	908	820	895	819	877	816	869	804	895	827
35 tot 45 jaar	914	776	843	737	791	704	729	647	702	615
45 tot 55 jaar	791	663	777	684	809	732	804	729	820	736
55 tot 65 jaar	456	287	452	322	503	380	529	417	572	452
Onderwijsniveau										
Laag	1 162	950	1 112	947	1 035	891	870	760	888	743
Midden	1 546	1 448	1 460	1 430	1 474	1 465	1 485	1 428	1 502	1 447
Hoog	859	678	885	731	924	788	974	874	1 026	917
Onbekend	17	9	12	6	18	13	49	36	58	49
Bedrijfstak										
Landbouw en visserij	77	33	62	28	61	27	60	26	60	28
Delfstoffenwinning	6	1	7	0	7	1	8	2	8	2
Industrie	643	186	572	162	562	158	549	159	549	158
Energie- en waterleidingbedrijven	19	6	38	10	43	12	44	11	42	12
Bouwnijverheid	342	37	327	36	305	35	266	32	260	31
Handel	671	598	658	593	669	597	655	593	674	597
Horeca	149	170	156	169	164	179	168	189	181	208
Vervoer, opslag en communicatie	310	114	432	143	433	150	423	151	428	158
Financiële instellingen	143	121	141	114	141	112	135	102	158	106
Zakelijke dienstverlening	868	618	720	564	720	556	740	569	784	620
Openbaar bestuur	13	14	22	21	17	15	20	14	19	16
Gesubsidieerd onderwijs	25	22	26	25	26	24	21	25	23	25
Gezondheids- en welzijnzorg	174	971	188	1 054	188	1 102	181	1 047	178	1 021
Cultuur en overige dienstverlening	137	167	112	162	108	152	103	149	103	147
Particulier hh met personeel en extra-territoriale organisaties	7	26	6	28	8	34	7	28	8	28
Onbekend	-	-	1	2	-	-	-	-	-	-
Sector										
Rijksoverheid										
Onderwijs										
Defensie										
Politie										
Rechterlijke macht										
Gemeenten										
Provincie										
Waterschappen										

2.1 Kerncijfers onderzoekspopulaties 2008 tot en met 2016 (slot)

	Overheid									
	2008		2010		2012		2014		2016	
	man	vrouw	man	vrouw	man	vrouw	man	vrouw	man	vrouw
	x 1 000									
Aantal banen	494	511	495	534	477	536	471	531	455	542
waarvan										
Voltijd	392	172	391	171	359	156	360	157	343	165
Deeltijd, 12 uur of meer	84	304	85	329	97	344	91	338	90	344
Deeltijd, minder dan 12 uur	18	35	19	34	21	36	20	36	23	33
Leeftijd										
15 tot 23 jaar	16	18	16	17	16	14	16	14	15	13
23 tot 35 jaar	87	138	95	152	85	142	91	141	93	140
35 tot 45 jaar	110	137	105	139	95	135	92	130	87	135
45 tot 55 jaar	164	149	155	148	144	149	132	142	123	144
55 tot 65 jaar	117	69	124	79	137	96	141	103	138	111
Onderwijsniveau										
Laag	53	41	51	38	53	36	45	34	41	30
Midden	167	151	164	155	156	147	158	134	148	128
Hoog	273	318	279	340	266	351	266	361	262	381
Onbekend	1	1	1	1	2	1	3	3	3	3
Bedrijfstak										
Landbouw en visserij										
Delfstoffenwinning										
Industrie										
Energie- en waterleidingbedrijven										
Bouwnijverheid										
Handel										
Horeca										
Vervoer, opslag en communicatie										
Financiële instellingen										
Zakelijke dienstverlening										
Openbaar bestuur										
Gesubsidieerd onderwijs										
Gezondheids- en welzijnszorg										
Cultuur en overige dienstverlening										
Particulier hh met personeel en extra-territoriale organisaties										
Onbekend										
Sector										
Rijksoverheid	76	53	72	51	73	54	68	53	65	57
Onderwijs	188	325	192	341	180	343	180	345	178	353
Defensie	55	9	52	8	47	6	54	10	46	8
Politie	40	22	41	23	40	22	37	18	38	19
Rechterlijke macht	1	2	2	3	2	3	2	4	2	4
Gemeenten	117	91	118	98	118	99	113	91	109	91
Provincie	8	6	10	7	8	5	7	5	7	5
Waterschappen	8	3	9	3	9	3	9	4	10	4

Bron: CBS.

van niet-financiële ondernemingen (NFO). Deze informatie is op ondernemingsniveau gekoppeld aan de baangegevens. In bijlage 2 worden de bronbestanden kort beschreven.

Ophogen steekproef

Om uitspraken te kunnen doen over alle banen bij de overheid en het bedrijfsleven in Nederland, zijn de banen uit deze 'gestapelde steekproef' opgehoogd naar de totale populatie van 1 miljoen banen van werknemers bij de overheid en 6,6 miljoen banen in

het bedrijfsleven. In paragraaf 6 wordt het weegmodel verder toegelicht. In tabel 2.1 zijn enkele kerncijfers behorend bij de onderzoekspopulaties (banen) voor de peilmomenten 2008, 2010, 2012, 2014 en 2016 samengevat.

3. Operationalisering

In deze paragraaf wordt toegelicht hoe de belangrijkste begrippen uit het onderzoek zijn geoperationaliseerd.

Uurloon

Het individuele uurloon is gelijk aan het basisloon per regulier gewerkt uur. Het basisloon is gelijk aan het (fiscaal) jaarloon, exclusief bijzondere beloning en overwerkloon, maar inclusief de fiscale waarde van niet in geld uitgekeerde belaste vergoedingen. Het aantal reguliere uren is gelijk aan het totale aantal verloonde uren, exclusief overwerkuren en verlofuren in verband met vakantie, adv en algemeen erkende feestdagen. Van alle banen op 30 september 2016 is het gemiddelde uurloon berekend voor het hele jaar 2016.

Gemiddeld uurloon

Het uurloon is berekend door voor elke baan het jaarloon te nemen en deze te delen door het aantal regulier gewerkte uren per jaar. Vervolgens worden de uurloonen per baan gemiddeld. In formule ziet dit er als volgt uit:

$$\text{Gemiddeld uurloon} = \frac{\sum_{b=1}^n \text{loon}_b / \text{uur}_b}{n}$$

waarbij loon_b het jaarloon bij baan b weergeeft en uur_b het aantal gewerkte uren per jaar in baan b . Tot slot staat n voor het aantal banen.

Ongecorrigeerd beloningsverschil

Het ongecorrigeerde beloningsverschil is het procentuele verschil in gemiddelde uurloonen tussen categorieën werknemers.

Inkomen partner

Van alle personen in het bestand is gekeken of zij een partner hebben. De inkomensgegevens van de partner zijn vervolgens via het SSB gekoppeld. Het persoonlijk inkomen van de partner omvat inkomen uit arbeid, inkomen uit eigen bedrijfsvoering, uitkeringen inkomensverzekeringen, uitkeringen sociale voorzieningen (met uitzondering van kinderbijslag) en ontvangen inkomensoverdrachten (alimentatie en dergelijke) verminderd met de betaalde premies voor inkomensverzekeringen.

Winst bedrijf per werkzame persoon

Onder winst wordt het bedrijfsresultaat per ondernemingengroep verstaan. Het bedrijfsresultaat is de netto omzet min de lonen, afschrijvingen en kosten van de omzet plus de baten uit hoofde van investeringspremies, subsidies en dergelijke. De ondernemingengroep bestaat uit ondernemingen die in financieel, organisatorisch en economisch opzicht zodanig verweven zijn dat ze voor de belasting als één onderneming worden aangemerkt. Vervolgens is de winst per werkzame persoon berekend door de winst te delen door het aantal werkzame personen bij de ondernemingengroep.

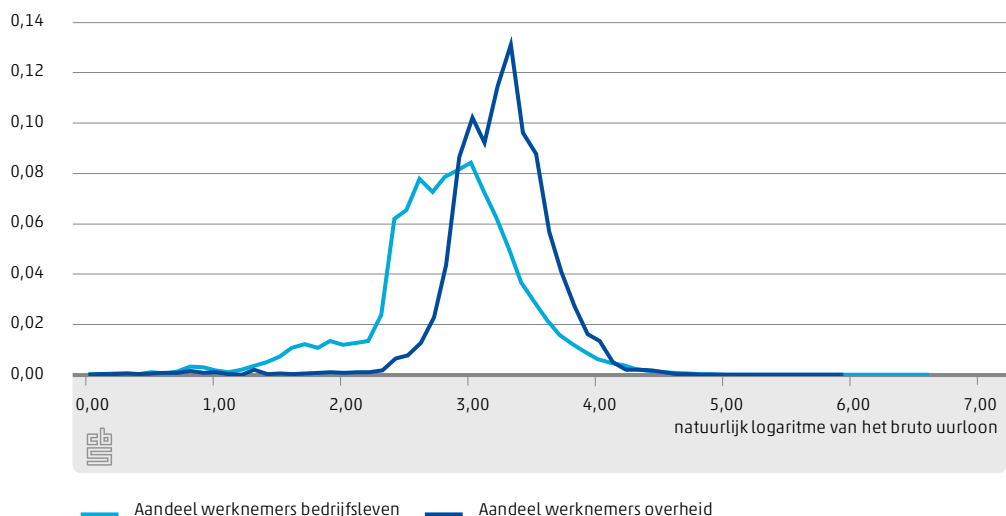
Percentage vrouwelijke collega's

Het percentage vrouwelijke collega's is het aantal vrouwelijke collega's ten opzichte van het totaal aantal werknemers in een bedrijf, exclusief de werknemer zelf.

4. Bepaling gecorrigeerde beloningsverschillen

Het gecorrigeerde beloningsverschil is het verschil in (meetkundig) gemiddeld uurloon tussen categorieën werknemers, dat overblijft na correctie voor de achtergrondkenmerken. Feitelijk wordt bekeken hoe groot het verschil in uurloon is tussen bijvoorbeeld mannen en vrouwen, als zij op alle andere achtergrondkenmerken gelijk zijn. Dit noemen we het 'corrigeren voor achtergrondkenmerken' en resulteert in een gecorrigeerd beloningsverschil. Dit gecorrigeerde beloningsverschil is bepaald met behulp van een multi-pele regressieanalyse. Met deze techniek wordt nagegaan in hoeverre de afhankelijke variabele (in dit geval uurloon) kan worden verklaard met behulp van verschillende onafhankelijke (of verklarende) variabelen.

4.1 Verdeling van de natuurlijke logaritme van het uurloon voor overheid en bedrijfsleven, september 2016



Een van de voorwaarden om een regressieanalyse te kunnen uitvoeren, is dat de afhankelijke variabele normaal verdeeld is. Het uurloon is aan de onderkant begrensd, maar aan de bovenkant van de uurloonverdeling kunnen grote uitschieters voorkomen. Dat betekent dat het uurloon niet normaal verdeeld is. In de regressieanalyse is ervoor gekozen om in plaats van het uurloon de natuurlijke logaritme van het uurloon te gebruiken als afhankelijke variabele. Dit is bij benadering namelijk wel normaal verdeeld, zie figuur 4.1 voor de uurlonen van 2016.

De beloningsfunctie van het geschatte model ziet er als volgt uit:

$$\ln(Y) = \alpha + \sum_j \beta_j X_{ij} + \epsilon_i$$

waarbij:

$\ln(Y)$	de afhankelijke variabele: de natuurlijke logaritme van het uurloon Y , $\ln(\text{uurloon})$
α	het intercept
β_j	de regressiecoëfficiënten, behorend bij variabele X_j
X_{ij}	de score van een individu i op de variabele X_j
ϵ_i	de storingsterm, ofwel het residu, behorend bij individu i

In dit model geeft de regressiecoëfficiënt β_j de verandering aan van de afhankelijke variabele $\ln(\text{uurloon})$ als gevolg van een verandering van de verklarende variabelen X_j . Het intercept geeft het gemiddelde $\ln(\text{uurloon})$ van iemand die voor alle variabelen in het model in de referentiecategorie valt. De referentiecategorieën zijn overgenomen uit Verschuren et al. (2014). De storingsterm ϵ_i is het verschil tussen de modelvoorspelling van het $\ln(\text{uurloon})$ van individu i , en het werkelijke $\ln(\text{uurloon})$ van individu i .

Het regressiemodel bevat vrijwel uitsluitend categoriale variabelen, met als enige uitzondering de variabele 'percentage vrouwelijke collega's'. De categoriale variabelen zijn als dummyvariabelen in het model opgenomen. Een dummyvariabele is een variabele die de waarde 0 of de waarde 1 kan aannemen. Zo heeft de dummyvariabele van de sector onderwijs twee waarden: niet werkzaam in het onderwijs (0) en wel werkzaam in het onderwijs (1). Per dummyvariabele wordt een regressiecoëfficiënt geschat die de afwijking weergeeft van het uurloon ten opzichte van de referentiecategorie, wanneer de overige variabelen gelijk blijven.

De reden voor het gebruik van dummyvariabelen is dat de meeste variabelen geen continue verdeling hebben. Enkele variabelen die wel een continu verloop kennen, bijvoorbeeld leeftijd, zijn ingedeeld in categorieën die niet even groot zijn. Daarom is ervoor gekozen om ook deze variabelen als dummy's in het regressiemodel op te nemen.

In bijlage 1 van deze methodologische toelichting worden de uitkomsten van de regressieanalyses weergegeven door middel van de regressiecoëfficiënten. Regressiecoëfficiënt β_j geeft de invloed van categorie j van de verklarende variabele X aan op de afhankelijke variabele $\ln(\text{uurloon})$. Wanneer, bijvoorbeeld, de regressiecoëfficiënt voor het hebben afgerond van een universitaire opleiding een hogere waarde heeft dan die voor het hebben afgerond van een hbo-opleiding, dan betekent dit dat een universitaire opleiding tot een hoger uurloon leidt dan een hbo-opleiding. Om het

geschatte beloningsverschil in euro's tussen een categorie j en de bijbehorende referentiecategorie uit te drukken in een percentage, volstaat het de exponent van β_j te nemen en deze te percenteren: $(e^{\beta_j} - 1) \cdot 100\%$.

Naast de regressiecoëfficiënten worden ook de standaardfouten SE gepubliceerd. Hiermee is het mogelijk om een betrouwbaarheidsinterval bij de beloningsverschillen te berekenen. De grenzen voor het 95% betrouwbaarheidsinterval van regressiecoëfficiënt β_j zijn $\beta_j \pm (1,96 \cdot SE)$. De betekenis is dat we bij herhaling van de procedure, met steeds nieuwe (aselecte) steekproeven uit dezelfde populatie, mogen verwachten dat 95% van de zo berekende intervallen dezelfde parameter zullen bevatten.

Daarnaast zijn t -waarden opgenomen in de tabellen met schattingsresultaten. De t -waarde of overschrijdingskans (van een gegeven steekproefuitkomst) is de kans dat in de verdeling gegeven door de nulhypothese de waarde van de toetsingsgrootte wordt behaald of overschreden (links, rechts dan wel tweezijdig). De t -waarde is groter naarmate de kans kleiner is dat de regressiecoëfficiënt toevallig afwijkt van 0. De t -waarde vat als het ware de bewijskracht van de steekproefuitkomst in gestandaardiseerde vorm samen.

Coëfficiënten die significant verschillen van de referentiegroep met een betrouwbaarheid van 99 procent, zijn in de tabellen gemarkeerd met een kardinaalteken (#). Als een regressiecoëfficiënt niet significant afwijkt van de referentiecategorie, dan is niet aangetoond dat er in de totale populatie een verschil is tussen het gemiddelde uurloon van personen in de betreffende categorie en de referentiegroep.

Een maat die iets zegt over de verklaaringskracht van het model is de proportie verklaarde variantie R^2 . Deze maat geeft aan welk deel van de verschillen in uurlonen kan worden verklaard door de achtergrondkenmerken in de beloningsfunctie. Een lage R^2 kan erop duiden dat niet alle variabelen die van invloed zijn op de hoogte van het uurloon, in het model zijn opgenomen.

4.2 Verklaarde variantie en interpretatiekracht

R^2	Verklaarde variantie	Interpretatie kracht model
<0,1	< 10%	Zeer zwak
0,1-0,25	10-25%	Zwak
0,25-0,5	25-50%	Matig
0,5-0,75	50-75%	Sterk
0,75-0,9	75-90%	Zeer sterk
> 0,9	> 90%	Uitzonderlijk sterk

Op deze manier zijn er twee regressiemodellen opgesteld, waarin een breed scala aan achtergrondfactoren is opgenomen. Met het eerste model worden verschillen in uurloon bij de overheid verklaard. Het tweede model verklaart verschillen in uurloon in het bedrijfsleven. In bijlage 1 is een overzicht te vinden van de verklarende variabelen die in de twee modellen zijn opgenomen en de bijbehorende uitkomsten. Omdat het

gaat om twee verschillende modellen waar verschillende verklarende variabelen in zijn opgenomen, moeten de uitkomsten uit beide modellen los van elkaar beoordeeld worden. Zo kan het effect van leeftijd in het model voor de overheid niet vergeleken worden met het effect van leeftijd in het model voor het bedrijfsleven.

5. Oaxaca-Blinder decompositie methode

Naast het bestaande regressiemodel, dat zorgt voor longitudinaal vergelijkbare cijfers over 'gecorrigeerde loonverschillen', is tevens de Oaxaca-Blinder methode uitgevoerd. De Oaxaca-Blinder decompositie analyse is een variant van de meervoudige regressiemethode. Deze methode wordt doorgaans gebruikt om een kloof in gemiddelde uitkomsten tussen twee groepen te bekijken. Dit wordt gedaan door middel van een aparte lineaire regressieanalyse per groep. Deze methode is geschikt wanneer het verschil in het gemiddelde uurloon tussen mannen en vrouwen nader wordt onderzocht (Konings, 2005).

Met de Oaxaca-Blinder methode kunnen loonverschillen tussen mannen en vrouwen uitgesplitst worden in een 'verklaarbaar' deel en een residu of 'onverklaarbaar' deel (Jann, 2008). Het verklaarbare deel geeft weer welk gedeelte van het loonverschil kan worden toegeschreven aan verschillen in objectieve factoren, zoals onderwijs of werkervaring, tussen mannen en vrouwen. Daarentegen weerspiegelt het onverklaarbare deel het gedeelte van de beloningsverschillen tussen mannen en vrouwen dat overblijft na correctie voor de verschillen in de achtergrondvariabelen die worden meegenomen in het model. De Oaxaca-Blinder decompositie analyse maakt dit onderscheid op het niveau van achtergrondvariabelen. Dit is mogelijk omdat er voor iedere groep een aparte vergelijking geschat wordt, en daarom kan er bijvoorbeeld bekeken worden of een man met een extra jaar ervaring hoger beloond wordt dan een vrouw met een extra jaar ervaring.

Bij toepassing van de Oaxaca-Blinder decompositie wordt voor ieder geslacht de volgende loonvergelijking geschat:

$$\ln(Y_i) = \beta_i X_i + \epsilon_i$$

Waarbij:

$\ln(Y_i)$	de afhankelijke variabele (de natuurlijke logaritme van het uurloon Y , $\ln(\text{uurloon})$) voor elke groep i : mannen (m) en vrouwen (v)
β_i	is een vector van de te schatten regressie-coëfficiënten, behorend bij variabele X_j
X_i	vector van de factoren die bijdragen aan het verklaarde deel zoals onderwijs en ervaring
ϵ_i	de storingsterm, ofwel het residu

De gemiddelde lonen voor elke groep i kan dan weergegeven worden door:

$$\ln(\bar{Y}_m) = \beta_m \bar{X}_m + \epsilon_m \text{ en } \ln(\bar{Y}_v) = \beta_v \bar{X}_v + \epsilon_v$$

Het verschil tussen de gemiddelden van deze twee regressies geeft het loonverschil weer:

$$\ln(\bar{Y}_m) - \ln(\bar{Y}_v) = \beta_m \bar{X}_m - \beta_v \bar{X}_v$$

Het loonverschil tussen beide geslachten wordt vervolgens uitgesplitst naar het verklaarbare en het onverklaarbare deel:

$$\ln(\bar{Y}_m) - \ln(\bar{Y}_v) = \beta_m (\bar{X}_m - \bar{X}_v) + \bar{X}_v (B_m - B_v)$$

$\beta_m (\bar{X}_m - \bar{X}_v)$ is dan het effect op het uurloon ten gevolge van verschillen in de gemiddelde kenmerken tussen mannen en vrouwen, dus het verklaarbare deel. De term $\bar{X}_v (B_m - B_v)$ weerspiegelt het onverklaarbare deel of residu.

In dit onderzoek is de Oaxaca-Blinder techniek toegepast op het model van het bedrijfsleven. Voor deze analysetechniek is het model van de overheid op dit moment niet geschikt om tot betrouwbare uitkomsten komen. Dit heeft te maken met de onderliggende data die in dit onderzoek mede is gebaseerd op de Enquête beroepsbevolking (EBB). De EBB is een enquête waarvoor een steekproef wordt getrokken. Deze steekproef dient een heel ander doel dan het in kaart te brengen van het beloningsverschil tussen mannen en vrouwen. Dit in combinatie met het feit dat de afbakening van de overheid een kleiner aantal waarnemingen omvat dan het bedrijfsleven maakt dat de resultaten van de Oaxaca-Blinder decompositie analyse op dit moment niet voldoende betrouwbaar zijn. In het rapport worden zodanig ook alleen de resultaten voor het bedrijfsleven besproken.

6. Ophogen van steekproef-totalen

De banen van werknemers zijn gekoppeld met drie jaren EBB (2015–2017). Omdat de EBB een steekproefonderzoek is onder personen, moet worden opgehoogd naar alle banen van werknemers op 30 september 2016. De weging is in twee stappen gedaan.

De eerste stap bestaat uit het ophogen van drie jaar EBB naar het aantal personen van 15 jaar en ouder in Nederland in 2016. Het EBB-gewicht is hierbij gecorrigeerd voor het feit dat de steekproef voor de drie EBB-jaren niet even groot is. In deze stap is herwogen naar de volgende kruisingen: geslacht en burgerlijke staat; geslacht en leeftijdsklasse; geslacht en landsdeel; herkomst en generatie.

De tweede stap bestaat uit het wegen naar het aantal banen op 30 september 2016. Hierbij zijn de gewichten uit de vorige stap eerst herschaald naar het totale aantal banen op 30 september 2016 volgens de loonaangifte. In deze stap is herwogen naar de volgende kruisingen: geslacht en leeftijd; sbi; geslacht en dienstverband; geslacht en grootteklasse; herkomst en dienstverband; type baan.

In de uitvoering van de regressieanalyses wordt eveneens gebruik gemaakt van ophooggewichten. Deze zorgen ervoor dat er in de berekeningen beter rekening wordt gehouden met de onderlinge verhoudingen tussen groepen werknemers.

7. Kwaliteit van de uitkomsten

Onnauwkeurigheid kleine aantallen

Zoals in ieder steekproefonderzoek hebben opgehoogde cijfers, in dit geval over uurlonen en de beloningsverschillen, een onnauwkeurigheidsmarge. Naarmate de aantallen kleiner zijn, gaan zij gepaard met hogere relatieve onnauwkeurigheidsmarges. Het samenvoegen en middelen van gegevens uit drie verschillende EBB-jaargangen vergroot de omvang van het onderzoeksbestand en beperkt deze marges. Op deze manier kan op meer gedetailleerd niveau uitspraken gedaan worden over de resultaten. Aantallen in de steekproef die gebaseerd zijn op minder dan 25 waarnemingen worden niet gepubliceerd.

Doordat in dit onderzoek verschillende bronnen in combinatie met de EBB zijn gebruikt, kunnen sommige uitkomsten verschillen van andere door het CBS gepubliceerde cijfers en vorige edities van dit onderzoek. In de rapportage 'Monitor loonverschillen mannen en vrouwen, 2016' zijn absolute aantallen afgerond op duizendtallen en zijn percentages afgerond op hele procenten.

Gebruik minder recente en meest recente gegevens EBB

Gegevens over onderwijs en beroep zijn in het onderzoeksbestand afkomstig uit drie jaargangen van de EBB, namelijk 2015, 2016 en 2017. Dat betekent dat gegevens over een baan op 30 september 2016 kunnen zijn verrijkt met gegevens over het onderwijs- en beroepsniveau van bijna twee jaar daarvoor. Vooral bij banen van jongeren kan dit leiden tot een vertekening van de uitkomsten, omdat hun situatie snel kan veranderen. Dit is bijvoorbeeld het geval wanneer een student met alleen een vwo-diploma en een bijbaantje begin 2016 is geënquêteerd en vervolgens eind 2016 is afgestudeerd en een baan is gaan uitoefenen op wetenschappelijk niveau. Dit leidt tot een onderschatting van het onderwijs- en beroepsniveau van jongeren.

8. Vergelijking met cijfers uit eerder onderzoek

Om de vergelijkbaarheid van de onderzoeksresultaten uit vorige edities van dit onderzoek te waarborgen, is ervoor gekozen de onderzoeksopzet, het regressiemodel en het weegmodel zoveel mogelijk gelijk te houden aan het eerder uitgevoerd onderzoek. Er zijn geen noemenswaardige veranderingen ten opzichte van het vorige onderzoek te vermelden.

Bijlage 1

1. Regressieresultaten overheid

	Beta-coëfficiënt		Standaardfout (SE)	t-waarde	Aantal waarnemingen
Geslacht					
Man		referentiecategorie			12 898
Vrouw	-0,046	#	0,006	-8,147	14 905
Leeftijd					
15-17 jaar	-1,183	#	0,050	-23,905	87
18 jaar	-0,611	#	0,048	-12,754	74
19 jaar	-0,597	#	0,041	-14,729	97
20 jaar	-0,283	#	0,038	-7,462	99
21 jaar	-0,016		0,029	-0,547	162
22 jaar	-0,171	#	0,027	-6,254	200
23 tot 35 jaar		referentiecategorie			5 297
35 tot 45 jaar	0,197	#	0,009	23,104	5 386
45 tot 55 jaar	0,249	#	0,010	24,795	7 941
55 tot 65 jaar	0,285	#	0,010	27,330	8 460
Herkomstgroepering					
Persoon met een Nederlandse achtergrond		referentiecategorie			24 356
Persoon westerse migratieachtergrond eerste generatie	-0,004		0,013	-0,343	646
Persoon westerse migratieachtergrond tweede generatie	0,019		0,010	1,822	1 410
Persoon niet-westerse migratieachtergrond eerste generatie	-0,029		0,012	-2,518	864
Persoon niet-westerse migratieachtergrond tweede generatie	0,026		0,015	1,760	527
Arbeidshandicap of chronische ziekte					
Wel arbeidshandicap of chronische ziekte	-0,013		0,010	-1,290	1 662
Geen arbeidshandicap of chronische ziekte		referentiecategorie			11 718
Onbekend	0,055		0,031	1,775	14 423
Onderwijsniveau					
Basisonderwijs	-0,195	#	0,021	-9,245	365
Vmbo	-0,162	#	0,011	-14,984	1 589
Mbo		referentiecategorie			7 513
Hbo	0,175	#	0,007	24,377	10 733
Wo	0,322	#	0,008	40,241	7 416
Onbekend	0,056		0,031	1,799	187
Onderwijsrichting					
Algemeen onderwijs	0,176	#	0,011	15,961	3 134
Leraar		referentiecategorie			6 052
Taalwetenschappen, geschiedenis, kunst	0,080	#	0,011	7,100	1 561
Sociale wetenschappen, bedrijfskunde	0,091	#	0,008	11,134	6 898
Natuurwetenschappen, informatica	0,049	#	0,012	4,246	1 588
Techniek, industrie, bouwkunde	0,028	#	0,011	2,639	2 385
Landbouw, diergeneeskunde	0,069	#	0,019	3,623	479
Gezondheidszorg, welzijn	0,080	#	0,009	8,794	2 892
Persoonlijke dienstverlening, vervoer	0,040	#	0,012	3,433	1 783
Onbekend	0,053	#	0,015	3,492	1 031
Inkomen partner					
Inkomen partner onbekend	-0,009		0,054	-0,163	35
Inkomen partner onder WML	0,013		0,008	1,666	4 468
Inkomen partner WML tot modaal		referentiecategorie			5 485
Inkomen partner modaal tot 2x modaal	0,008		0,007	1,177	8 527
Inkomen partner 2x modaal of meer	0,034	#	0,010	3,574	2 775
Geen partner	-0,016		0,007	-2,328	6 513

1. Regressieresultaten overheid (slot)

	Beta-coëfficiënt		Standaardfout (SE)	t-waarde	Aantal waarnemingen
Arbeidsduur					
Volttijd					13 663
Deeltijd >12 uur	-0,004		0,006	-0,755	12 633
Deeltijd <12 uur	-0,054 #		0,011	-5,024	1 507
Beroepsniveau					
Beroepsniveau 1	-0,014		0,016	-0,857	577
Beroepsniveau 2					5 418
Beroepsniveau 3	0,117 #		0,008	14,552	3 788
Beroepsniveau 4	0,231 #		0,007	31,603	16 854
Onbekend	0,136 #		0,018	7,638	1 166
Beroepsrichting management					
Ja	0,031		0,013	2,450	1 220
Nee					26 583
Werkervaring					
Aantal jaar gewerkt vanaf 15 - 0-1 jaar	-0,100 #		0,021	-4,839	409
Aantal jaar gewerkt vanaf 15 - 1-4 jaar	-0,061 #		0,018	-3,472	582
Aantal jaar gewerkt vanaf 15 - 5-9 jaar					1 579
Aantal jaar gewerkt vanaf 15 - 10-19 jaar	0,071 #		0,010	6,864	6 014
Aantal jaar gewerkt vanaf 15 - meer dan 20 jaar	0,099 #		0,013	7,883	18 603
Onbekend	0,136		0,018	2,023	616
Soort werknemer					
Reguliere werknemer					27 318
Stagair	-1,856 #		0,026	-70,515	335
WSW	-0,110		0,115	-0,956	9
Oproepkracht	-0,091 #		0,031	-2,968	141
Contractsoort					
Onbepaalde tijd	0,081 #		0,007	11,871	22 648
Bepaalde tijd					5 155
Leidinggevende functie					
Ja	0,119 #		0,006	19,642	5 891
Nee					21 912
Sector					
Overheid					3 297
Onderwijs	-0,025 #		0,009	-2,883	15 174
Defensie	0,007		0,014	0,470	1 389
Politie	0,150 #		0,012	12,487	1 545
Rechterlijke macht	0,163 #		0,030	5,458	161
Gemeenten	-0,103 #		0,009	-12,048	5 503
Provincies	0,095 #		0,021	4,563	365
Waterschappen	0,059 #		0,021	2,849	369
Grootteklasse					
Meer dan 500 werknemers	0,047 #		0,006	8,267	18 327
Minder dan 500 werknemers					9 476
Regio					
Werkgemeente Noord-Nederland	-0,003		0,008	-0,343	2 874
Werkgemeente Oost-Nederland	-0,023 #		0,006	-3,798	5 587
Werkgemeente West-Nederland					13 340
Werkgemeente Zuid-Nederland	-0,016		0,006	-2,462	4 940
Werkgemeente onbekend	-0,064 #		0,006	-2,462	1 062
Percentage vrouwelijke collega's	0,001 #		0,000	6,939	27 803
Constante	2,455 #		0,020	121,635	27 803
R ²	0,589				

De uitkomsten verschillen met een betrouwbaarheid van ten minste 99 procent van referentiegroep.

2. Regressieresultaten bedrijfsleven

	Beta-coëfficiënt	Standaardfout (SE)	t-waarde	Aantal waarnemingen
Geslacht				
Man	referentiecategorie			81 651
Vrouw	-0,075 #	0,002	-32,349	75 949
Leeftijd				
15 jaar	-0,876 #	0,014	-62,177	835
16 jaar	-0,758 #	0,009	-80,678	2 677
17 jaar	-0,611 #	0,008	-75,006	3 610
18 jaar	-0,485 #	0,007	-65,277	3 780
19 jaar	-0,356 #	0,007	-50,185	3 559
20 jaar	-0,248 #	0,007	-36,471	3 476
21 jaar	-0,157 #	0,007	-24,240	3 386
22 jaar	-0,076 #	0,006	-12,333	3 461
23 tot 35 jaar	referentiecategorie			34 861
35 tot 45 jaar	0,122 #	0,003	36,727	28 255
45 tot 55 jaar	0,161 #	0,004	41,981	40 002
55 tot 65 jaar	0,190 #	0,004	46,113	29 698
Herkomstgroepering				
Persoon met een Nederlandse achtergrond	referentiecategorie			133 035
Persoon westerse migratieachtergrond eerste generatie	-0,041 #	0,004	-9,869	4 588
Persoon westerse migratieachtergrond tweede generatie	-0,003	0,004	-0,845	7 259
Persoon niet-westerse migratieachtergrond eerste generatie	-0,060 #	0,004	-16,623	7 673
Persoon niet-westerse migratieachtergrond tweede generatie	-0,002	0,004	-0,528	5 045
Arbeidshandicap of chronische ziekte				
Wel arbeidshandicap of chronische ziekte	-0,046 #	0,004	-11,339	7 911
Geen arbeidshandicap of chronische ziekte	referentiecategorie			61 346
Onbekend	-0,009 #	0,002	-5,601	88 343
Onderwijsniveau				
Basisonderwijs	-0,090 #	0,004	-20,415	8 694
Vmbo	-0,057 #	0,003	-22,182	30 157
Mbo	referentiecategorie			68 311
Hbo	0,125 #	0,003	46,523	31 180
Wo	0,302 #	0,004	84,101	16 873
Onbekend	0,000	0,008	0,022	2 385
Onderwijsrichting				
Algemeen onderwijs	0,065 #	0,006	10,237	36 917
Leraar	referentiecategorie			3 849
Taalwetenschappen, geschiedenis, kunst	-0,017	0,008	-2,320	4 677
Sociale wetenschappen, bedrijfskunde	0,060 #	0,006	9,904	34 339
Natuurwetenschappen, informatica	0,059 #	0,008	7,881	5 312
Techniek, industrie, bouwkunde	0,065 #	0,006	10,150	24 779
Landbouw, diergeneeskunde	0,036 #	0,009	4,234	3 183
Gezondheidszorg, welzijn	0,090 #	0,006	14,425	23 925
Persoonlijke dienstverlening, vervoer	0,041 #	0,007	6,295	13 092
			onbekend:	7 527
Huishoudenspositie				
Thuiswonend kind	-0,006	0,004	-1,574	29 411
Partner in paar zonder kinderen	-0,050 #	0,003	-20,174	36 060
Partner in paar met kinderen1				63 420
Ouder in eenouderhuishouden	0,044 #	0,005	8,871	6 238
Overig (lid van een) huishouden	-0,004	0,008	-0,451	1 497
Alleenstaand	referentiecategorie			20 974
Inkomen partner				
Inkomen partner onbekend	0,080 #	0,019	4,164	225
Inkomen partner onder WML	0,020 #	0,003	6,761	27 837
Inkomen partner WML tot modaal	referentiecategorie			29 085
Inkomen partner modaal tot 2x modaal	0,007 #	0,003	2,578	33 751
Inkomen partner 2x modaal of meer	0,064 #	0,005	13,962	8 582
Geen partner	-0,078 #	0,003	-24,618	58 120

2. Regressieresultaten bedrijfsleven (vervolg)

	Beta-coëfficiënt	Standaardfout (SE)	t-waarde	Aantal waarnemingen
Arbeidsduur				
Voltijd	referentiecategorie			68 289
Deeltijd >12 uur	-0,045 #	0,002	-19,221	62 236
Deeltijd <12 uur	-0,050 #	0,003	-15,661	27 075
Beroepsniveau				
Beroepsniveau 1	-0,067 #	0,003	-20,668	14 313
Beroepsniveau 2	referentiecategorie			65 008
Beroepsniveau 3	0,123 #	0,003	45,732	26 181
Beroepsniveau 4	0,218 #	0,003	75,175	38 328
Onbekend	-0,015 #	0,004	-3,618	13 770
Beroepsrichting management				
Ja	0,139 #	0,004	31,882	8 895
Nee	referentiecategorie			148 705
Werkervaring				
Aantal jaar gewerkt vanaf 15 - 0-1 jaar	-0,020 #	0,005	-3,897	6 333
Aantal jaar gewerkt vanaf 15 - 1-4 jaar	-0,009	0,005	-2,029	10 956
Aantal jaar gewerkt vanaf 15 - 5-9 jaar	referentiecategorie			15 165
Aantal jaar gewerkt vanaf 15 - 10-19 jaar	0,062 #	0,004	17,662	32 226
Aantal jaar gewerkt vanaf 15 - meer dan 20 jaar	0,112 #	0,004	25,679	79 294
Onbekend	0,037 #	0,005	7,613	13 626
Soort werknemer				
Reguliere werknemer	referentiecategorie			127 823
Stagair	-1,426 #	0,008	-184,771	2 140
WSW	-0,294 #	0,009	-32,596	1 488
Uitzendkracht	-0,044 #	0,005	-9,783	8 821
Oproepkracht	-0,010 #	0,004	-2,921	12 596
DGA	0,278 #	0,006	45,529	4 732
Contractsoort				
Onbepaalde tijd	0,102 #	0,002	46,367	98 387
Bepaalde tijd	referentiecategorie			59 213
Leidinggevende functie				
Ja	0,099 #	0,002	40,894	30 372
Nee	referentiecategorie			127 228
Bedrijfstak				
Landbouw en visserij	referentiecategorie			2 081
Delfstoffenwinning	0,263 #	0,024	10,921	231
Industrie	0,091 #	0,008	11,107	16 685
Energie- en waterleidingbedrijven	0,115 #	0,012	9,349	1 308
Bouwnijverheid	0,139 #	0,009	15,797	6 507
Handel	-0,003	0,008	-0,424	29 527
Horeca	-0,038 #	0,009	-4,431	8 243
Vervoer	0,057 #	0,008	6,935	13 782
Financiële instellingen	0,217 #	0,009	23,755	6 596
Zakelijke dienstverlening	0,061 #	0,008	7,534	31 904
Openbaar bestuur	0,102 #	0,014	7,137	950
Gesubsidieerd onderwijs	0,089 #	0,013	6,972	1 282
Gezondheids- en welzijnzorg	0,171 #	0,009	20,248	31 623
Cultuur en overige dienstverlening	0,062 #	0,009	6,950	6 073
Particuliere huishoudens met personeel en extra-territoriale organisaties	0,388 #	0,014	27,453	808
Grootteklasse				
Kleinbedrijf (minder dan 10 werknemers)	-0,055 #	0,003	-18,978	26 841
Middenbedrijf (10 tot 100 werknemers)	referentiecategorie			40 770
Grootbedrijf (100 of meer werknemers)	0,039 #	0,002	17,917	89 989

2. Regressieresultaten bedrijfsleven (slot)

	Beta-coëfficiënt	Standaardfout (SE)	t-waarde	Aantal waarnemingen
Winst bedrijf per werkzaam persoon				
Minder dan 10 euro per jaar	0,025 #	0,003	7,573	17 609
10 tot 2 000 euro per jaar	-0,017 #	0,003	-5,374	22 629
2 000 tot 10 000 euro per jaar	referentiecategorie			33 178
10 000 tot 32 000 euro per jaar	0,029 #	0,003	9,632	21 437
Meer dan 32 000 euro per jaar	0,128 #	0,004	33,699	11 991
Onbekend	0,013 #	0,003	4,967	50 756
Regio				
Werkgemeente Noord-Nederland	-0,065 #	0,003	-19,353	13 356
Werkgemeente Oost-Nederland	-0,053 #	0,003	-21,553	28 948
Werkgemeente West-Nederland	referentiecategorie			62 962
Werkgemeente Zuid-Nederland	-0,045 #	0,002	-18,673	32 856
Werkgemeente onbekend	-0,020 #	0,004	-5,601	19 478
Percentage vrouwelijke collega's	-0,001 #	0,000	-20,329	157 600
Constante	2,539 #	0,011	230,286	157 600
R ²	0,700			

De uitkomsten verschillen met een betrouwbaarheid van ten minste 99 procent van referentiegroep.

Bijlage 2

Bronnen

Enquête Beroepsbevolking (EBB)

De EBB is een doorlopende enquête onder personen van 15 jaar en ouder die in Nederland wonen, met uitzonderingen van personen in inrichtingen, instellingen en tehuizen (de institutionele bevolking). Het doel van de EBB is om inzicht te krijgen van de relatie tussen mens en arbeidsmarkt. Gegevens worden vastgesteld op het moment van enquêteren.

De EBB is een steekproef waarop elk jaar ongeveer 90 duizend personen responderen. Een deel van deze 90 duizend personen valt buiten het bereik van dit onderzoek omdat het geen baan heeft. Om voldoende massa te krijgen voor het onderzoek zijn drie jaargangen EBB (2015-2017) gebruikt. Het kan voorkomen dat een persoon meerdere banen heeft. In dat geval zijn de gegevens van een persoon uit de EBB gekoppeld aan meerdere banen.

De ophooggewichten die beschikbaar zijn in de EBB, hogen op naar de populatie personen van 15 jaar en ouder in Nederland in het betreffende jaar. Om de drie EBB-bestanden per jaargang gezamenlijk op te kunnen hogen naar de populatie banen van werknemers in 2016 zijn de EBB-gewichten aangepast (zie paragraaf 6). Hiervoor is gebruik gemaakt van dezelfde herwegingsmethode als in het eerder beloningsverschillen onderzoek (zie De Mooij et al. (2010), Geerdinck et al. (2012), Verschuren et al. (2014) en Muller et al. (2016)).

De volgende persoonskenmerken uit de EBB zijn gebruikt voor dit onderzoek: handicap of chronische ziekte, onderwijsrichting, onderwijsniveau, beroepsniveau, beroepsrichting, regio van de werkgemeente, het aantal jaar dat iemand gewerkt heeft vanaf de 15e verjaardag en of iemand een leidinggevende functie heeft.

Stelsel van Sociaal-statistische Bestanden (SSB)

Het SSB is een stelsel van registers en enquêtes, die op persoonsniveau aan elkaar zijn gekoppeld. Per jaargang worden meer dan 50 registers gebruikt. Deze registers hebben betrekking op verschillende sociaaleconomische onderwerpen, zoals banen, uitkeringen, woningen en onderwijs.

De doelpopulatie van het SSB bestaat uit alle personen die in Nederland wonen, en personen die niet in Nederland wonen maar in Nederland werken of een uitkering dan wel pensioen vanuit Nederland ontvangen.

Voor dit onderzoek zijn gegevens over de volgende onderwerpen uit het SSB gebruikt:

- Banen (2016): De baangegevens zijn ontleend aan de Polisadministratie van UWV.
- Demografische gegevens gebaseerd op de Basisregistratie Personen (BRP) waaronder: geslacht, leeftijd herkomstgroepering en generatie, huishoudenspositie, het hebben van kinderen en of iemand een partner heeft. Het vaststellen van persoonsgegevens sluit aan bij het peilmoment van de banen (30 september 2016).
- Persoonlijk inkomen van personen. Hieruit wordt het inkomen van de partner bepaald.

Algemeen Bedrijven Register (ABR)

In het Algemeen Bedrijven Register (ABR) worden bedrijven en instellingen met hun identificatie- en structuurgegevens geregistreerd. Op basis van dit bestand wordt vastgesteld welke bedrijven tot één onderneming behoren, om uiteindelijk informatie over de winstgevendheid te verkrijgen.

Statistiek Financiën van niet-financiële ondernemingen (NFO)

De statistiek Financiën van niet-financiële ondernemingen (NFO) bevat informatie over de jaarrekening van alle niet-financiële rechtspersoonlijkheid bezittende ondernemingen in Nederland, die vennootschapsbelastingplichtig zijn. De NFO wordt in dit onderzoek gebruikt om de winst per werkzame persoon vast te stellen.

Begrippen

Arbeidsduur – Dit is de indeling naar voltijd- en deeltijdbanen. Vaak bestaat een voltijdbaan uit 36 tot 40 uur, maar dit kan verschillen per bedrijf. In principe wordt de meest waargenomen wekelijkse arbeidsduur gebruikt om het aantal uren van een voltijdbaan te bepalen. Wanneer dit niet geloofwaardig is, bijvoorbeeld omdat er weinig voltijd wordt gewerkt, is de wekelijkse arbeidsduur voor voltijdbanen uit cao-gegevens afgeleid. Een persoon heeft een voltijdbaan wanneer hij/zij per week een aantal uren werkt (exclusief overwerkuren) dat minimaal 95 procent bedraagt van de gebruikelijke wekelijkse voltijdsarbeidsduur in het bedrijf of de bedrijfstak. Een deeltijder is iemand die per week minder dan 95 procent van de gebruikelijke wekelijkse arbeidsduur werkt.

Baan – Een expliciete of impliciete arbeidsovereenkomst tussen een persoon en een economische eenheid waarin is vastgelegd dat arbeid zal worden verricht waartegen een (financiële) beloning staat. Een baan wordt in de gegevensbestanden geoperationaliseerd door een unieke combinatie van een persoon en bedrijf. Baanwisselingen binnen een bedrijf worden hierdoor niet waargenomen.

Basisloon – Het basisloon is een variabele die gebruikt wordt om een zo constant mogelijk loonbegrip te hanteren. Het wordt benaderd door uit te gaan van het fiscaal loon en daar diverse componenten van af te halen, zoals bijzondere beloningen.

Bedrijfsleven – Het bedrijfsleven omvat in dit onderzoek zowel particuliere bedrijven als gesubsidieerde instellingen. Voorbeelden van gesubsidieerde instellingen zijn de gezondheids- en welzijnzorg, de uitvoeringsorganen voor de sociale verzekeringen en de sociale werkplaatsen.

Beroepsniveau – De indeling naar beroep is overeenkomstig de International Standard Classification of Occupations 2008 (ISCO 2008). In dit onderzoek worden de volgende categorieën van beroepsniveaus onderscheiden:

- Beroepsniveau 1 (ISCO 2008): onder beroepsniveau 1 vallen beroepen met eenvoudig en routinematig lichamelijk en handmatig werk met gebruik van handgereedschap zoals een spade, of eenvoudige elektrische apparaten als een stofzuiger.
- Beroepsniveau 2 (ISCO 2008): onder beroepsniveau 2 vallen beroepen met taken als het bedienen van machines en elektronische apparaten, voertuigen besturen, onderhouden en repareren van elektrische en mechanische apparaten en het bewerken, ordenen en opslaan van gegevens.
- Beroepsniveau 3 (ISCO 2008): onder beroepsniveau 3 vallen beroepen met taken als het uitvoeren van complexe technische en praktische taken die een uitgebreide feitenkennis, technische kennis en kennis van procedures op een specifiek gebied vereisen.
- Beroepsniveau 4 (ISCO 2008): onder beroepsniveau 4 vallen beroepen met taken als het uitvoeren het oplossen van ingewikkelde problemen en nemen van beslissingen die zijn gebaseerd op een uitgebreide theoretische en praktische kennis op een gespecialiseerd gebied.

Beroepsrichting – De Beroepenindeling ROA CBS (BRC) 2014 is een van de ISCO 2008 afgeleide indeling van beroepen bedoeld voor toepassing in analyses en statistieken op nationaal niveau. De BRC 2014 is afgeleid van de 3 en 4 digits beroepencodes uit de ISCO 2008.

In dit onderzoek worden de volgende categorieën van beroepsrichting onderscheiden:

- *Pedagogische beroepen*: deze beroepsklasse omvat de docenten, sportinstructeurs, leidsters kinderopvang en onderwijsassistenten.
- *Creatieve en taalkundige beroepen*: deze beroepsklasse omvat de auteurs en kunstenaars, vakspecialisten op artistiek en cultureel gebied.
- *Commerciële beroepen*: deze beroepsklasse omvat de adviseurs marketing, public relations en sales, vertegenwoordigers, inkopers en verkopers.
- *Bedrijfseconomische en administratieve beroepen*: deze beroepsklasse omvat de specialisten bedrijfsbeheer en administratie, specialisten personeels- en loopbaanontwikkeling en administratief personeel.

- *Managers*: deze beroepsklasse omvat de managers die leidinggeven aan het gehele bedrijf of aan een afdeling binnen het bedrijf.
- *Openbaar bestuur, veiligheid en juridische beroepen*: deze beroepsklasse omvat de overheidsambtenaren en -bestuurders, juristen, beveiligingspersoneel en militairen.
- *Technische beroepen*: deze beroepsklasse omvat de ingenieurs en onderzoekers wis-, natuur- en technische wetenschappen, technici en toezichthouders bouw en industrie, procesoperators, bouw- en metaalarbeiders en voedselverwerkende beroepen.
- *ICT beroepen*: deze beroepsklasse omvat de ICT-specialisten, gebruikersondersteuning ICT, radio- en televisie technici.
- *Agrarische beroepen*: deze beroepsklasse omvat de tuinders, akkerbouwers, veetelers en hulpkrachten landbouw.
- *Zorg en welzijn beroepen*: deze beroepsklasse omvat de artsen, therapeuten en verpleegkundigen, specialisten op maatschappelijk gebied, vakspecialisten gezondheidszorg, sociaal werkers, groeps- en woonbegeleiders en verzorgenden.
- *Dienstverlenende beroepen*: deze beroepsklasse omvat de medewerkers persoonlijke dienstverlening, horecapersoneel, schoonmakers en keukenhulpen.
- *Transport en logistiek beroepen*: deze beroepsklasse omvat de bestuurders voertuigen en bedieners mobiele installaties, en hulpkrachten transport en logistiek.
- *Overig*: deze beroepsklasse omvat werkzame personen waarvan het beroep onbekend of niet in te delen is.

Deeltijd – Zie Arbeidsduur.

Economische activiteit (SBI) – De verzameling van werkzaamheden, gericht op de productie van goederen en diensten. Het gaat hierbij niet alleen om activiteiten van het bedrijfsleven, maar ook om de activiteiten van niet op winst gerichte instellingen en de overheid.

Gecorrigeerd beloningsverschil – Het gecorrigeerde beloningsverschil is het verschil in (meetkundig) gemiddeld uurloon, tussen twee groepen werknemers, dat overblijft na correctie voor verschillen in achtergrondkenmerken. In dit onderzoek staat het gecorrigeerd beloningsverschil tussen mannen en vrouwen centraal. Zie paragraaf 4 voor meer informatie over gecorrigeerde beloningsverschillen.

Herkomstgroepering – Voor de indeling van personen naar etnische achtergrond is de CBS-indeling naar herkomstgroepering gebruikt. De herkomstgroepering van een persoon wordt vastgesteld aan de hand van diens geboorteland en dat van zijn ouders.

Leeftijd – De leeftijd van een persoon wordt bepaald op 30 september 2016.

Modaal inkomen – In dit onderzoek is voor het modaal inkomen een bedrag van 33 000 euro gebruikt. Het modaal inkomen is het bruto inkomen net onder de maximum premiegrens van de zorgverzekeringswet. Dit is niet gelijk aan het statistisch modaal (het meest voorkomende) inkomen.

Ongecorrigeerd beloningsverschil – Het procentuele verschil tussen het (rekenkundig) gemiddelde uurloon van categorieën werknemers.

Onderwijsrichting – De onderwijsrichting is de richting van de behaalde studie. De richtingen zijn ingedeeld naar onderwijsrichting volgens de International Standard Classification of Education (ISCED 2008). Het CBS heeft de gedetailleerde ISCED-gebieden nog verder uitgesplitst met behulp van de Standaard Onderwijsindeling (SOI) 2016. Bij alle indelingen gaat het steeds om een bundeling van opleidingen die qua richting in het onderwijs nauw aan elkaar verwant zijn.

In dit onderzoek worden de volgende categorieën van onderwijsrichting onderscheiden:

- *Onderwijs*: Dit omvat onderwijsrichtingen rondom het onderwijs.
- *Taalwetenschappen, geschiedenis, kunst*: Dit omvat onderwijsrichtingen rondom taalwetenschappen, geschiedenis en kunst.
- *Sociale wetenschappen, bedrijfskunde*: Dit omvat onderwijsrichtingen rondom sociale wetenschappen en bedrijfskunde.
- *Natuurwetenschappen, informatica*: Dit omvat onderwijsrichtingen rondom natuurwetenschappen en informatica.
- *Techniek, industrie, bouwkunde*: Dit omvat onderwijsrichtingen rondom techniek, industrie en bouwkunde.
- *Landbouw, diergeneeskunde*: Dit omvat onderwijsrichtingen rondom landbouw en diergeneeskunde.
- *Gezondheidszorg, welzijn*: Dit omvat onderwijsrichtingen rondom gezondheidszorg en welzijn.
- *Persoonlijke dienstverlening, vervoer*: Dit omvat onderwijsrichtingen rondom persoonlijke dienstverlening en vervoer.

Onderwijsniveau – Het behaalde onderwijsniveau is het niveau van de hoogste met succes afgerond onderwijs. De opleidingen zijn ingedeeld naar onderwijsniveau volgens de Standaard onderwijsindeling (SOI 2016). De SOI 2016 is een classificatie van opleidingen naar niveau en richting.

De SOI 2016 omvat een basisindeling naar niveau en een publicatieindeling naar niveau. De basisindeling naar niveau van de SOI 2016 is een indeling in twee aggregatieniveaus. Op het hoogste niveau worden 7 niveaus onderscheiden, en binnen niveau 3, 4 en 5 is een uitsplitsing naar de 3 subniveaus laag, midden en hoog gemaakt.

De publicatieindeling naar niveau van de SOI 2016 is een hiërarchische indeling in 4 aggregatieniveaus. Op het hoogste niveau worden 3 niveaus onderscheiden, laag, middelbaar en hoog onderwijsniveau. Het 2e aggregatieniveau bevat 5 categorieën en het 3e aggregatieniveau 8 categorieën. Het 4e aggregatieniveau is een uitbreiding ten opzichte van de SOI 2006 en is ontwikkeld om de toepassingsmogelijkheden van de publicatie-indeling verder te verbeteren. Dit is een indeling in 18 categorieën.

Overheid – De cao-sector overheid omvat alle publiekrechtelijke bedrijven en is onderverdeeld in acht deelsectoren: rijksoverheid, onderwijs, defensie, politie, rechterlijke macht, gemeenten, provincies en waterschappen.

Persoon van de eerste generatie – Persoon die in het buitenland is geboren met ten minste één in het buitenland geboren ouder.

Persoon van de tweede generatie – Persoon die in Nederland is geboren en van wie ten minste één ouder in het buitenland is geboren.

Persoon met een migratieachtergrond – Persoon van wie ten minste één ouder in het buitenland is geboren.

Persoon met een Nederlandse achtergrond – Persoon van wie de beide ouders in Nederland zijn geboren.

Persoon met een niet-westerse migratieachtergrond – Persoon met als herkomstgroepering een van de landen in de werelddelen Afrika, Latijns-Amerika en Azië (exclusief Indonesië en Japan) of Turkije.

Persoon met een westerse migratieachtergrond – Persoon met als herkomstgroepering een van de landen in de werelddelen Europa (exclusief Turkije), Noord-Amerika en Oceanië of Indonesië of Japan.

Reguliere uren – Reguliere uren zijn de basisuren minus het aantal uren feestdagen en uren van algemene plus leeftijdsspecifieke verlofdagen. Basisuren zijn de uren waar het normale loon van de werknemer tegenover staat. Als er extra wordt gewerkt tegen een hogere vergoeding, dan zijn dat overwerkuren.

Soort werknemer – Soort baan dat een werknemer heeft, onderverdeeld naar regulier, stagiair, Wet sociale werkvoorziening (WSW), uitzendkracht, oproepkracht en directeur-groootaandeelhouder (DGA).

Uurloon – Het basisloon van een baan per regulier gewerkt uur. Het basisloon is gelijk aan het (fiscaal) jaarloon, exclusief bijzondere beloning en overwerkloon, maar inclusief de fiscale waarde van niet in geld uitgekeerde belaste vergoedingen. Het aantal reguliere uren is gelijk aan het totale aantal verloonde uren in het jaar, exclusief overwerkuren en verlofuren in verband met vakantie, adv en algemeen erkende feestdagen.

Voltijd – Zie Arbeidsduur.

Afkortingen

ABR – Algemeen Bedrijven Register
adv – Arbeidsduurverkortung
BRC – Beroepenindeling ROA CBS
BRP – Basisregistratie Personen
CBS – Centraal Bureau voor de Statistiek
DGA – Directeur-grotaandeelhouder
EBB – Enquête Beroepsbevolking
Havo – Hoger algemeen vormend onderwijs
ISCED – International Standard Classification of Education
ISCO – International Standard Classification of Occupations
Mbo – Middelbaar beroepsonderwijs
NFO – Statistiek Financiën van niet-financiële ondernemingen
ROA – Research Centre for Education and the Labour Market
SBI 2008 – Standaard Bedrijfsindeling 2008
SOI – Standaard onderwijsindeling
SFO – Statistiek Financiën van Ondernemingen
SOI – Standaard onderwijsindeling
SZW – Ministerie van Sociale Zaken en Werkgelegenheid
UWV – Uitvoeringsinstituut Werknemersverzekeringen
Vmbo – Voorbereidend middelbaar beroepsonderwijs
Vwo – Voorbereidend wetenschappelijk onderwijs
WML – Wettelijk minimumloon
Wo – Wetenschappelijk onderwijs
WSW – Wet sociale werkvoorziening

Literatuur

De Mooij, M., Bottelberghs, I., Goedhuys, M., Van den Tillaart, J. en Wagner, C. (2010). *Gelijk loon voor gelijk werk? Banen en lonen bij overheid en bedrijfsleven, 2008*, CBS, Den Haag.

Geerdinck, M., Geijtenbeek, L., Graham, J., Sluiter, N. en Wagner, C. (2012). *Gelijk loon voor gelijk werk? Banen en lonen bij overheid en bedrijfsleven, 2010*, CBS, Den Haag.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4), 453-479.

Konings, J. (2005). Worden vrouwen gediscrimineerd op de arbeidsmarkt? Een micro-econometrische analyse voor België. Katholieke Universiteit Leuven, Departement Economie.

Muller, L., Floris, J., Verhallen-Schumacher, D., Dankmeyer, B., Chkalova, K. en Verschuren, L. (2016). *Gelijk loon voor gelijk werk? Banen en lonen bij overheid en bedrijfsleven, 2014*, CBS, Den Haag.

Verschuren, L., Bakker, K., Dankmeyer, B., de Mooij, M., Schulte, M. en Van den Tillaart, J. (2014). *Gelijk loon voor gelijk werk? Banen en lonen bij overheid en bedrijfsleven, 2012*, CBS, Den Haag.

Verklaring van tekens

Niets (blanco)	Een cijfer kan op logische gronden niet voorkomen
.	Het cijfer is onbekend, onvoldoende betrouwbaar of geheim
*	Voorlopige cijfers
**	Nader voorlopige cijfers
2017-2018	2017 tot en met 2018
2017/2018	Het gemiddelde over de jaren 2017 tot en met 2018
2017/'18	Oogstjaar, boekjaar, schooljaar enz., beginnend in 2017 en eindigend in 2018
2015/'16-2017/'18	Oogstjaar, boekjaar, enz., 2015/'16 tot en met 2017/'18

In geval van afronding kan het voorkomen dat het weergegeven totaal niet overeenstemt met de som van de getallen.

Colofon

Uitgever

Centraal Bureau voor de Statistiek
Henri Faasdreef 312, 2492 JP Den Haag
www.cbs.nl

Prepress

Centraal Bureau voor de Statistiek

Ontwerp

Edenspiekermann

Inlichtingen

Tel. 088 570 70 70
Via contactformulier: www.cbs.nl/infoservice

© Centraal Bureau voor de Statistiek, Den Haag/Heerlen/Bonaire, 2018.
Verveelvoudigen is toegestaan, mits het CBS als bron wordt vermeld.