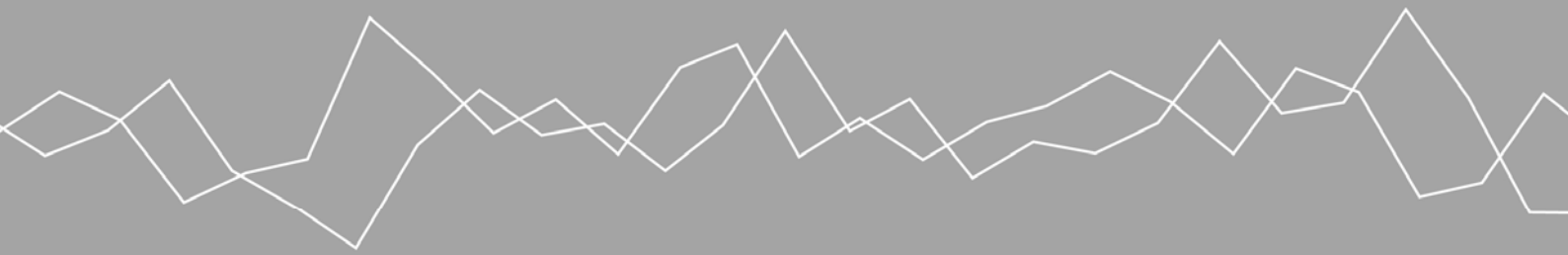


Verdiepende analyse van loonverschillen



Amsterdam, oktober 2008
In opdracht van het Ministerie van SZW

Verdiepende analyse van loonverschillen

De loonachterstand van vrouwen verder uitgediept

Ernest Berkhout (SEO)
Chris van Klaveren (SEO)
Kea Tijdens (AIAS)
Wiemer Salverda (AIAS)



seo economisch onderzoek

“De wetenschap dat het goed is”

SEO Economisch Onderzoek doet onafhankelijk toegepast onderzoek in opdracht van overheid en bedrijfsleven. Ons onderzoek helpt onze opdrachtgevers bij het nemen van beslissingen. SEO Economisch Onderzoek is gelieerd aan de Universiteit van Amsterdam. Dat geeft ons zicht op de nieuwste wetenschappelijke methoden. We hebben geen winstoogmerk en investeren continu in het intellectueel kapitaal van de medewerkers via promotietrajecten, het uitbrengen van wetenschappelijke publicaties, kennisnetwerken en congresbezoek.

SEO-rapport nr. 2008-7

ISBN 978-90-6733-459-4

Copyright © 2008 SEO Economisch Onderzoek Amsterdam. Alle rechten voorbehouden. Het is geoorloofd gegevens uit dit rapport te gebruiken in artikelen en dergelijke, mits daarbij de bron duidelijk en nauwkeurig wordt vermeld.

Inhoudsopgave

Voorwoord	i
Samenvatting en conclusies	iii
Summary	ix
1 Inleiding.....	1
2 Onderzoeksopzet & herhaalbaarheid	3
2.1 Onderzoeksopzet.....	3
2.1.1 Betrouwbaarheidsintervallen puntschattingen	3
2.1.2 Alternatieve modelspecificatie	3
2.1.3 Alternatieve schattingsmethode	4
2.1.4 Alternatieve verklaringen.....	4
2.2 Herhaalbaarheid.....	4
3 Uitbreiding bestaande analyse binnen traditionele schattingsmethode	7
3.1 Basisreplicatie	7
3.2 Betrouwbaarheid van de puntschatter	11
3.2.1 Technische verantwoording.....	11
3.2.2 Betrouwbaarheidsgrafieken loonkloof AVO-onderzoek	13
3.3 Verbetering modelspecificatie	20
3.3.1 Modeluitbreiding en herdefiniëring variabelen	21
3.3.2 Niet-lineaire specificaties en interactietermen: technische leeswijzer	24
3.3.3 Niet-lineaire specificaties en interactietermen: resultaten	26
3.4 Conclusie uitbreiding modelspecificatie.....	31
4 Alternatieve schattingsmethode: ‘matching’	33
4.1 Kenmerken van matching.....	33
4.1.1 Methoden van matchen	34
4.2 Matching resultaten basisloonschatting	38
4.3 Matching resultaten modeluitbreiding.....	40
4.4 Conclusie alternatieve schattingsmethode.....	41
5 Alternatieve verklaringen: literatuurstudie	43

5.1	Definitie en geschiedenis van de loonkloof	43
5.1.1	Definities en begrippen.....	43
5.1.2	Wetgeving en overheidsinitiatieven	44
5.2	Verklarende factoren.....	45
5.2.1	Economische factoren: menselijk kapitaal, sector en arbeidsduur.....	45
5.2.2	Psychologische factoren: een ‘fair’ loon, taakbeschrijving en glazen plafond	47
5.2.3	Sociologische factoren: de impact van gezinsopbouw en segregatie naar sekse	48
5.3	Onderzoek naar loonverschillen in Nederland.....	49
5.3.1	Inleiding.....	49
5.3.2	Economische factoren	50
5.3.3	Psychologische factoren	54
5.3.4	Sociologische factoren	55
5.4	Conclusie onderzoek naar de loonkloof in Nederland.....	58
6	Alternatieve verklaringen: een empirische verkenning.....	63
6.1	Data Loonwijzer	63
6.2	Replicatie basisloonschatting.....	64
6.3	Aanvullende verklaringen voor verschillen in loon	69
6.3.1	Wordt het loonverschil man-vrouw beter verklaard?.....	70
6.3.2	Welke variabelen spelen een rol?.....	70
6.4	Conclusie alternatieve verklaringen	70
7	Kanttekeningen & aanbevelingen	70
7.1	Kanttekeningen bij de onderzoeksresultaten	70
7.2	Aanbevelingen voor vervolgonderzoek	70
Bijlage A	Definities AVO dataset.....	70
Bijlage B	Beschrijvende statistieken AVO	70
Bijlage C	Betrouwbaarheidsintervallen	70
Bijlage D	Loonschattingen voor bedrijfsleven en overheidssector	70
Bijlage E	Definities Loonwijzer-dataset	70
Bijlage F	Aanvullende variabelen Loonwijzer	70

Bijlage G	Nederlandse onderzoeken, opgenomen in het overzicht van studies naar de loonverschillen man-vrouw.....	70
------------------	---	-----------

Voorwoord

In dit rapport doen we verslag van een onderzoek naar aanvullende verklaringen voor de loonkloof tussen mannen en vrouwen. Het onderzoek is uitgevoerd in opdracht van het Ministerie van Sociale Zaken en Werkgelegenheid. Het is het resultaat van een bijzonder vruchtbare samenwerking tussen SEO Economisch Onderzoek (SEO) en het Amsterdams Instituut voor ArbeidsStudies (AIAS) van de Universiteit van Amsterdam.

Het onderzoek borduurt verder op eerder onderzoek naar beloningsverschillen in opdracht van het Ministerie van Sociale Zaken en Werkgelegenheid, uitgevoerd door de Arbeidsinspectie. Wij maken dan ook gebruik van dezelfde gegevensbron, het Arbeidsvoorwaardenonderzoek 2004 (AVO). Wij danken de Arbeidsinspectie, in het bijzonder Judith Hoeben, voor het ter beschikking stellen van deze data en de toelichting daarbij.

Daarnaast zijn we veel dank verschuldigd aan de immer positief geformuleerde op- en aanmerkingen van de begeleidingscommissie, bestaande uit Siebrand Bisschop (Commissie Gelijke Behandeling), Özcan Erdem (Arbeidsinspectie), Karin Hagoort (Centraal Bureau voor de Statistiek), Harm Jonkhart (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties), Charles de Vries (Ministerie van Onderwijs, Cultuur en Wetenschap) en Camiel Jansen, Angela de Keizer en Jana van Kuringen (allen namens het Ministerie van Sociale Zaken en Werkgelegenheid). Ondanks deze inspirerende begeleiding zijn uiteraard de auteurs volledig verantwoordelijk voor de inhoud van het rapport.

Ernest Berkhout
Chris van Klaveren
Kea Tijdens
Wiemer Salverda.

Samenvatting en conclusies

In alle landen en op allerlei niveaus is sprake van loonverschillen. Waar sommige verschillen verklaard kunnen worden omdat ze samenhangen met verschillen in productiviteit, is dat bij beloningsverschillen tussen mannen en vrouwen nauwelijks van toepassing. Wanneer alleen gekeken wordt naar directe loonverschillen, zonder rekening te houden met verklarende factoren als opleiding, ervaring, deeltijdfactor, setcor etc. blijken vrouwen veel minder te verdienen dan mannen. Onderzoeken van onder andere SEO Economisch Onderzoek¹ en de Arbeidsinspectie² laten zien dat ook na correctie voor allerlei relevante factoren er telkens een (veel kleiner) onverklaard loonverschil overblijft. Vrouwen blijken dan nog altijd zo'n 3 tot 7% minder loon te ontvangen dan mannen. Dit resterende verschil wordt door sommigen gezien als beloningsdiscriminatie, maar dat is niet juist. Er zijn meerdere verklaringen mogelijk voor dit verschil: beloningsdiscriminatie is er daar slechts één van, maar tegelijkertijd kan ook in het verklaarde deel discriminatie besloten zitten. Een deel van het onverklaarde beloningsverschil zal altijd blijven bestaan vanwege beperkingen aan databronnen en methoden, terwijl men bij de interpretatie van de primaire loonverschillen ook rekening moet houden met secundaire beloning, subjectieve factoren en zelfselectie van vrouwen in relatief minder betalende banen. Het Ministerie van Sociale Zaken en Werkgelegenheid wil meer inzicht verkrijgen in andere mogelijke verklaringen van het resterende 'gecorrigeerde beloningsverschil'. Door uitbreiding van de onderzoeksmethoden en het gebruik van alternatieve databronnen kan inderdaad een meer gedetailleerde beeld geschetst worden, zo blijkt uit de voorliggende rapportage.

Vertrekpunt is bovengenoemde analyse die in 2006 is uitgevoerd door de Arbeidsinspectie op basis van het tweejaarlijkse Arbeidsvoorwaardenonderzoek. In ons onderzoek kijken we welke factoren niet onderzocht worden in bovengenoemde rapportage maar wel relevant zijn voor de verklaring van loonverschillen m-v. Vanwege de verschillen in de omvang van de belangrijkste databron kunnen de resultaten van het onderzoek slechts vergeleken worden binnen de overheid apart en binnen de marktsector apart.³ Het onderzoek leent zich daarom niet voor interpretatie van (verschillen in) loonverschillen *tussen* overheid en marktsector.

Uitbreiding bestaande analyse met traditionele schattingsmethoden

We beginnen het onderzoek in Hoofdstuk 3 met een inspectie van de onzekerheidsmarges, behorend bij de basisreplicatie van het eerdere Arbeidsinspectie-onderzoek. Deze marges zijn groter voor het bedrijfsleven dan voor de overheid, het loonverschil tussen mannen en vrouwen wordt derhalve in de marktsector minder nauwkeurig gemeten. De grotere variatie komt waarschijnlijk doordat de lonen in de marktsector nu eenmaal veel minder uniform zijn dan in de overheidssector, waar vaker met vaste functiewaarderingsystemen gewerkt wordt. Bovendien

¹ E.E. Berkhout, A. Heyma & W. Salverda (2006), *Beloningsverschillen tussen de marktsector en de collectieve sector in 2004*, SEO rapport nr. 889.

² Ö. Erdem, M.A. Bos & J. Hoeben (2006), *De arbeidsmarktpositie van werknemers in 2004. Verschillen in beloning en mobiliteit*.

³ Voor de overheid zijn veel minder verklarende variabelen beschikbaar dan voor de marktsector. Daarnaast is ook niet voor alle sectoren van de overheid informatie beschikbaar: zo ontbreekt bijvoorbeeld informatie over werknemers bij gemeente, provincie en politie.

bevat het bedrijfsleven in deze definitie ook de zorgsector, een sector die relatief weinig karakteristieken gemeen heeft met de andere subsectoren in het bedrijfsleven.

Bedrijfsleven

Uit de basisreplicatie voor het bedrijfsleven resulteert het reeds bekende gecorrigeerde loonverschil van 7,4%. In de volgende stap breiden we de basisreplicatie uit met nieuwe variabelen. Daaruit blijkt dat vrouwen in het bedrijfsleven minder vaak een winstuitkering ontvangen, maar als ze die krijgen is hij wel net zo groot als die van mannen. Ook de meeste andere toegevoegde variabelen dragen bij in de nauwkeuriger ontleding van loonverschillen. Desondanks verandert het onverklaarde loonverschil tussen man-vrouw met deze uitbreiding nauwelijks in grootte. Wel is duidelijk dat uitzendwerk verschilt van een reguliere baan; uitzendkrachten verdienen gemiddeld ook veel minder. Dit bevestigt het vermoeden dat in de basisreplicatie uitzendkrachten onterecht gerekend werden tot de personen met een reguliere baan. Aparte categorieën onderscheiden is voor toekomstig onderzoek wenselijk.

Vervolgens breiden we de analyse nog verder uit (door de toevoeging van interactie- en hogere ordetermen) en wint de verklaring van loonverschillen nog meer aan dimensie. In de subsectoren bouw, handel en transport (typische ‘mannensectoren’) blijkt het beloningsverschil tussen mannen en vrouwen groter te zijn dan in de zorgsector (een typische ‘vrouwensector’). Gemiddeld resulteert in de meest uitgebreide analyse een loonverschil van 6,6%, minder dus dan in de basisreplicatie, maar nog steeds substantieel.

De gebruikte AVO-dataset bevat voor de overheidssector veel minder geschikte variabelen dan voor de marktsector, en heeft ook niet alle subsectoren van de overheid vertegenwoordigd: gemeenten, provincies, waterschappen, militairen en politie ontbreken. Bij de interpretatie van de resultaten is het belangrijk om dit voor ogen te houden. Alle hierna vermelde resultaten kunnen niet zonder meer vergeleken worden *tussen* beide sectoren, en waar we spreken van de ‘overheid’ bedoelen we in feite alleen de in het onderzoek vertegenwoordigde sectoren.

Tabel 1.1 **Overzicht van loonverschillen volgens de traditionele methode**

	<i>Bedrijfsleven</i>
Basisloonschatting / replicatie	7,4
+ uitbreiding nieuwe variabelen	7,6
+ interactie- & hogere orde effecten	6,6
	<i>Overheid</i>
Basisloonschatting / replicatie	4,0
+ uitbreiding nieuwe variabelen	3,8
+ interactie- & hogere orde effecten	2,7

Overheid

Bij de overheidssector geeft de basisreplicatie het bekende loonverschil van 4,0%. Niemand ontvangt een winstgerelateerde uitkering en nagenoeg iedereen een niet-winstgerelateerde uitkering (bijvoorbeeld een eindejaarsuitkering). In het meest uitgebreide model (met extra variabelen, interactietermen en hogere ordetermen) bedraagt het onverklaarde loonverschil gemiddeld 2,7%. We concluderen daarom dat ook voor de overheid de gemaakte

modelaanpassingen zorgen voor een lagere waarde van het resulterende loonverschil tussen mannen en vrouwen, en dat er dus meer te verklaren valt dan tot nog toe bekend was. Ook hier zijn er echter belangrijke verschillen tussen de subsectoren, met het grootste loonverschil in de subsector Onderwijs. Verdere uitsplitsing binnen deze omvangrijke subsector was uiteraard wenselijk geweest, het betreft namelijk een heterogene groep variërend van primair- tot universitair onderwijs, die in totaal 6 verschillende CAO's omvat. Helaas kan dit relevante onderscheid in de AVO-data niet gemaakt worden.

Technische kanttekening: alternatieve schattingsmethoden

Een nadeel van de traditionele schattingsmethodes (zoals OLS) is de beperking van resultaten door de gekozen functionele vorm. Theoretisch is het mogelijk om in een dergelijke analyse bijvoorbeeld jonge vrouwen met oude mannen te vergelijken, ook als ze eigenlijk niet goed vergelijkbaar zijn. Om dit nadeel te overwinnen nemen we in Hoofdstuk 4 onze toevlucht tot een modernere schattingsmethode: *propensity score matching*. Vanwege het specialistische karakter van deze exercitie verwijzen we de geïnteresseerde lezer naar de conclusies aan het einde van Hoofdstuk 4 en naar Hoofdstuk 7 waarin de resulterende kanttekeningen & aanbevelingen zijn opgetekend. In deze algemene samenvatting vermelden we slechts de belangrijkste conclusies: in de traditionele methode worden –noodzakelijkerwijs– soms vrouwen en mannen vergeleken die eigenlijk slecht vergelijkbaar zijn. Propensity score matching is een zinvolle methode voor dit soort onderzoek. Beide schattingsmethoden geven in ieder geval aan dat er uit de AVO-dataset nog meer informatie te halen valt dan tot nu toe bekend was.

Literatuuronderzoek naar de loonkloof in Nederland

Hoofdstuk 5 geeft een overzicht van de relevante wetenschappelijke literatuur, op het gebied van economische, psychologische en sociologische oorzaken van de loonkloof. De Nederlandse onderzoeken overziend moeten we constateren dat de economische effecten goed in kaart gebracht zijn, behalve het sectoreffect. Beloningsverschillen tussen sectoren vinden in Nederland hun institutionele grondslag in cao's en de achterliggende functiewaarderingen, maar de relatie tussen de institutionele beloningsstructuur en de loonverschillen m-v zijn tot nu toe eigenlijk niet onderzocht. Een beter inzicht in de seksecompositie van sectoren en de invloed daarvan op het loonverschil (*'verdien je meer in een sector waar meer mannen werken?'*) zal het onverklaarde deel daarvan zeker doen verminderen.

Kijken we naar de psychologische effecten, dan zijn de effecten van mogelijke sekseverschillen in individuele loononderhandelingen tussen werkgever en werknemer nog nauwelijks onderzocht (*'Onderhandelen mannen efficiënter, en/of meer op primaire beloning gericht?' 'Is de onderhandelingsruimte van mannen anders?'*). Een probleem hierbij is de meting van de onderhandelingsruimte, vooral bij secundaire arbeidsvoorwaarden. Een correct ontwerp voor een dergelijk onderzoek vereist inzicht in personeelsbestanden van bedrijven en personeelsdossiers. Op voorhand kan geen inschatting gemaakt worden van het mogelijke effect hiervan op de verkleining van het onverklaarde deel van de loonkloof.

Wat de sociologische effecten betreft, zijn de effecten van beroepensegregatie goed onderzocht, maar die van seksecompositie en hiërarchische structuur van organisaties veel minder. Ook zou de relatie tussen gezinsvorming en arbeidsaanbod (*stoppen met werken als zij kinderen krijgt gebeurt vaker door laag opgeleide vrouwen dan hoogopgeleide vrouwen*) en het effect daarvan op het loonverschil m-v nog uitgebreider onderzocht kunnen worden (*'hoe groot is het nadeel van enkele jaren buiten het arbeidsproces staan?'*). De loonkloof is bovendien groter onder oudere werknemers. Een beter begrip van het loonverschil m-v in de oudere leeftijdsgroepen zou daarom kunnen bijdragen aan een verkleining van het onverklaarde deel.

Alternatieve verklaringen: een empirische verkenning

In Hoofdstuk 6 bekijken we in hoeverre het loonverschil nog verder ontleed kan worden met behulp van aanvullende, vaak subjectieve, variabelen. In hoeverre is er een uitruil tussen salaris en andere arbeidsvoorwaarden? Nemen bijvoorbeeld vrouwen in functies met afwisselend werk en weinig reistijd genoeg met minder salaris? Is de compenserende waarde van het afwisselende werk of de geringe reistijd voor hen belangrijker? Dergelijke vragen zijn niet te beantwoorden met het AVO-bestand, zodat we gebruik maken van de Loonwijzer dataset. De focus is daardoor niet meer op het absolute percentage loonverschil man-vrouw, maar op het gedeelte onverklaard loonverschil dat met nieuwe variabelen alsnog verklaard kan worden: welke nieuwe variabelen spelen hierin een rol?

Het blijkt dat de toevoeging van huishoudkenmerken (*is er een partner, een eigen huis, jonge kinderen?*), subjectieve baankenmerken (*is er afwisselend werk, leuke collega's*), het percentage vrouwen per functie en enkele overige kenmerken een extra verklaring geeft voor het loonverschil tussen mannen en vrouwen. Al deze variabelen samen verklaren grofweg 20 procent van het aanvankelijke gecorrigeerde loonverschil uit de basisreplicatie. Er spelen dus meer factoren een rol bij man-vrouw loonverschillen dan de standaard variabelen die normaal gesproken beschikbaar zijn in datasets gebaseerd op registraties. Dit zegt overigens niets over de wenselijkheid of redelijkheid van deze loonverschillen.

Hoewel het onderzoek niet is opgezet om allerlei afgeleide hypothesen te toetsen, biedt de empirische verkenning in dit hoofdstuk wel degelijk aanknopingspunten bij enkele theorieën uit de wetenschappelijke literatuur over loonverschillen. Tevredenheid met reistijd blijkt inderdaad een looncompenserende factor te zijn (zij het beperkt in omvang) bij zowel bedrijfsleven als overheid; een goede combinatie werk-vrije tijd is dit alleen bij de overheid en de combinatie familie-werk is dit alleen in het bedrijfsleven. Daarentegen worden negatieve baankenmerken als saai, gevaarlijk en ongezond werk niet gecompenseerd met een hoger loon, maar blijken het indicatoren van een slechte arbeidsmarktpositie.

Onder de huishoudensvariabelen blijken enkele selectie-effecten zichtbaar, die deels van invloed zijn op het gemiddelde loon maar waarschijnlijk minstens zozeer op de participatiebeslissing. Onze studie is niet toereikend om op basis van enkel deze huishoudensvariabelen concrete beleidsaanbevelingen te kunnen formuleren. Zij biedt echter wel enkele handvaten voor verdiepend onderzoek dat meer specifiek zou moeten kijken naar de effecten van beleid op het reserveringsloon, en de invloed daarvan op de participatiebeslissing.

Het interessantst is wellicht dat een in Hoofdstuk 3 geformuleerde conclusie hier wederom bijval krijgt. Hoe meer mannen in een bepaalde functie werken, hoe hoger het loon. Er blijken typische mannen- en typische vrouwenberoepen te zijn en juist die vrouwenberoepen zijn vaak de lagerbetaalde beroepen.

Summary

Wage differentials exist in all countries and at all levels. While some differentials are attributable to differences in productivity, this is hardly true in the case of differentials between men and women. If wage differentials are viewed in isolation, disregarding explanatory factors such as education, experience, part-time ratio, sector etc., women are found to earn much less than men. Surveys conducted by SEO Economic Research⁴ and the Labour Inspectorate (Arbeidsinspectie)⁵ show that even after adjusting for all kinds of relevant factors a (much smaller) unexplained wage gap always remains. The surveys show that women still receive some 3-7% less pay than men. This gap is regarded by some as pay discrimination, but that is not correct. Various explanations can be given for this difference: pay discrimination is only one of them, but discrimination may also exist in the explained part. Part of the wage gap will always remain unexplained owing to the limitations of data sources and methods, and when interpreting wage differentials account must also be taken of fringe benefits, subjective factors and self-selection of women in jobs that pay relatively less well. The Ministry of Social Affairs and Employment wishes to have more information about other possible explanations of the residual ‘adjusted wage gap’. The present report shows that indeed a more detailed picture can be obtained by expanding the research methods and using alternative data sources.

Starting point is the above-mentioned analysis carried out by the Labour Inspectorate in 2006 on the biennial survey ‘Arbeidsvoorwaardenonderzoek’ (AVO). In our analysis we examine which factors were not taken into account but are relevant to an understanding of the gender pay gap. Owing to the differences in the scope of the main data source the results of the analysis allow comparison only within the public sector separately and within the private sector separately.⁶ The results are not suitable for interpreting wage differentials (and differences in wage differentials) *between* the public and private sectors.

Expansion of existing analysis to include traditional methods of estimation

We start our study in Chapter 3 by considering the uncertainty margins relating to the replication of the previous Labour Inspectorate analysis. As these margins are greater for the private sector than the public sector, the pay differential between men and women is measured less accurately in the private sector. The higher variance is probably because pay in the private sector is less uniform than in the public sector, where standard job evaluation systems are used more frequently. In addition the private sector includes, for the purposes of this definition, the health sector which has relatively little in common with other parts of the private sector.

⁴ E.E. Berkhout, A. Heyma & W. Salverda (2006), *Beloningsverschillen tussen de marktsector en de collectieve sector in 2004 (Wage differentials between the private and public sectors in 2004)*, SEO Report No. 889.

⁵ Ö. Erdem, M.A. Bos & J. Hoeben (2006), *De arbeidsmarktpositie van werknemers in 2004. Verschillen in beloning en mobiliteit (The Labour market position of employees in 2004. Differences in pay and mobility)*.

⁶ Fewer explanatory factors are available for the public sector than for the private sector. In addition, information is not available for all parts of the public sector; for example, no information is available about staff employed by municipal and provincial authorities and the police.

Private sector

The replication for the private sector produces the known adjusted pay differential of 7.4%. In the next step we expand the basic replication to include new variables. This shows that women in the private sector receive profit-based bonuses less frequently than men, but if they do receive such a bonus it is just as large as that of their male counterparts. Most other added variables also contribute to a more accurate analysis of wage differentials. Nonetheless, the unexplained gender wage differential barely changes. It is clear, however, that there is a marked difference between temporary agency job and a standard job: on average, agency workers earn much less. This confirms the suspicion that in the basic replication temporary staff were wrongly classified as persons with an standard job. For future analysis, it would be desirable to distinguish between these separate categories.

We then extend the analysis still further by the addition of interaction and higher order terms, with the result that the explanation of wage differentials acquires a further dimension. The gender wage differential is larger in the construction, trade and transport subsectors (typically ‘male dominated sectors’) than in the health sector (a typically ‘female dominated sector’). In the most extensive analysis, this results in a pay differential of 6.6% on average, which is less than in the basic replication but still substantial.

The data set used for this purpose contains many fewer variables for the public sector than for the private sector and also does not cover all parts of the public sector: municipalities, provinces, water control corporations, the armed forces and the police are all missing. It is important to remember this when interpreting the results. None of the results referred to below can therefore be automatically compared *between* the two sectors, and where we refer to the ‘public sector’ we in fact mean only the subsectors represented in the survey.

Overview of wage differentials, using traditional estimation method

	<i>Private sector</i>
Basic wage estimate / replication	7.4
+ addition of new variables	7.6
+ interaction & higher order effects	6.6
	<i>Public sector</i>
Basic wage estimate / replication	4.0
+ addition of new variables	3.8
+ interaction & higher order effects	2.7

Public sector

The basic replication in the case of the public sector gives the known adjusted wage differential of 4.0%. No one receives profit-based bonuses and almost everyone receives non-profit related payments (e.g. a year-end bonus). In our most detailed model (with extra variables, interaction terms and higher order terms) the unexplained wage differential averages 2.7%. We therefore conclude that also in the public sector our model adjustments result in a lower estimate of the gender wage gap, so there is more to be explained than hitherto known. However, there are major differences between the subsectors, the largest pay differential being in education. A further breakdown of this large subsector would naturally have been desirable, since it consists of

a heterogeneous group varying from primary to university education and comprises a total of six different collective labour agreements. However relevant this distinction may be, it can not be made in our data.

Technical remark: alternative estimation methods

A disadvantage of the traditional estimation methods (such as OLS) is the limitation of results by the chosen functional form. In such an analysis it is theoretically possible to compare, for example, young women with old men, even if they are in fact incomparable. To overcome this disadvantage, we have resorted to a more modern estimation method in Chapter 4, namely propensity score matching. Owing to the specialist character of this analysis, we refer interested readers to the conclusions at the end of Chapter 4 and to Chapter 7 in which the resulting remarks and recommendations are noted. In this general summary we only mention the main conclusions: although some men and women are incomparable, the traditional method sometimes does compare them. Propensity score matching is a useful additional method for estimating wage differentials. Both estimation methods indicate that more information can be derived from the AVO data set than known hitherto.

Literature study of the gender pay gap in the Netherlands

Chapter 5 gives an overview of the relevant academic literature on the economic, psychological and sociological causes of the pay gap. After examining the Dutch surveys we conclude that they provide good information about the economic effects, apart from the sector effect. Wage differentials between sectors are institutionalised in the Netherlands through collective agreements, but the relationship between the institutional pay structure and wage differentials between men and women has not yet really been studied. More insight in the gender composition of sectors and how this influences wage differentials (*‘Do you earn more in a male dominated sector?’*) will certainly reduce the unexplained part of the differentials to some extent.

As regards the psychological effects, little if any study has yet been carried out into the effects of possible gender-related differences in individual pay negotiations between employer and employee (*‘Do men negotiate more efficiently and/or are they focused more on basic pay?’ ‘Do men have different scope for negotiation?’*). Measuring the scope for negotiation, especially as regards fringe benefits, is a problem in this connection. To design such a study properly it is necessary to have information about company workforces and access to personnel files. It is not possible to gauge in advance what effect this may have on reducing the unexplained part of the gender pay gap.

As far as the sociological aspects are concerned, the effects of occupational segregation have been well documented, but those of the gender composition and hierarchical structure of organisations to a much lesser extent. Much more extensive research could also be done into the relationship between family formation and labour supply (*‘lower educated women are much more likely than their higher educated counterparts to stop working when they have children’*) and the effect of this on the gender pay gap (*‘how great is the disadvantage of not being part of the employment process for some years?’*). The gender pay gap is also larger among older employees. Better understanding of the gender pay

gap in the older age groups could therefore help to reduce the unexplained part of the differential.

Alternative explanations: empirical research

In Chapter 6 we examine to what extent the gender wage gap can be analysed in even greater detail using additional, often subjective variables. To what extent is there a trade-off between salary, fringe benefits and other job conditions? For example, are women in jobs that involve varied work and little commuting time willing to accept less salary? Is the compensatory value of varied work or little commuting time more important to them? As such questions cannot be answered with the AVO data set, we have used the Wage Indicator data set. As a result, the focus is no longer on the absolute percentage of the gender wage gap but on the hitherto unexplained part of the gap that can now be explained using new variables. Which new variables are important?

Apparently the addition of household characteristics (*does the woman have a partner, an owner-occupied home, young children?*), subjective job characteristics (*does she have varied work, pleasant colleagues?*), the percentage of women per type of job and a few other characteristics provides an extra explanation of the gender pay gap. All these variables together explain roughly 20 percent of the initially adjusted wage gap from the basic replication. So there are other (subjective) factors that play a role in explaining the gender pay gap, besides those that are available in data sets based on registrations.

Although the survey was not set up to test all kinds of hypotheses based on the data, the empirical research in this chapter does provide a basis for assessing some theories from the academic literature on wage differentials. Satisfaction with travelling time does indeed prove to be a compensating factor for lower pay (albeit of limited scope) in both the private and the public sector; a good work-life balance is a compensating factor only in the public sector and a good work-family balance is a compensating factor only in the private sector. By contrast, negative job characteristics such as boring, dangerous and unhealthy work are not offset by higher pay, but prove to be indicators of a weak position in the labour market.

The household variables include a few selection effects that partly influence average pay, but probably have just as much effect on the decision to participate in the labour market. Our study is not fitted to formulate specific policy recommendations on the basis of these household variables alone. It does, however, provide a few directions for more in-depth research. Such research should be aimed more specifically at the effects of public policies on the reservation wage and how that influences the participation decision.

Most interesting to note is that a conclusion already formulated in Chapter 3 once again receives support in Chapter 6. The higher the number of men who work in a given job, the higher the pay. Apparently are typically male dominated occupations and typically female dominated occupations, and precisely the female occupations are often the lowest paid.

1 Inleiding

In alle landen en sectoren en op allerlei niveaus is sprake van loonverschillen. Tussen ouderen en jongeren, hoge functies en lage functies, vrouwen en mannen. Zouden de eerste twee nog gerechtvaardigd kunnen worden door verschillen in productiviteit, bij de verschillen naar geslacht is dat nauwelijks van toepassing. Wanneer alleen gekeken wordt naar directe loonverschillen, zonder rekening te houden met verklarende factoren als opleiding, ervaring, sector etc. blijken vrouwen aanzienlijk minder te verdienen dan mannen. In Nederland in 2004 zo'n 14% bij de overheid en zelfs 21% in het bedrijfsleven (Arbeidsinspectie, 2006). Onderzoeken van bijvoorbeeld SEO Economisch Onderzoek en de Arbeidsinspectie laten zien dat ook na correctie voor allerlei relevante factoren er telkens een onverklaard loonverschil overblijft. Vrouwen blijken dan nog altijd zo'n 3 tot 7% minder loon te ontvangen dan mannen. Dit resterende verschil wordt door sommigen gezien als beloningsdiscriminatie, maar dat is niet geheel juist. Er zijn meerdere verklaringen mogelijk voor dit verschil. Beloningsdiscriminatie is er daar slechts één van.

Bij het Ministerie van Sociale Zaken en Werkgelegenheid bestaat de wens om meer inzicht te krijgen in mogelijke verklaringen van het resterende verschil. Daarom heeft zij aan SEO Economisch Onderzoek gevraagd om, in samenwerking met het Amsterdams Instituut voor Arbeidsstudies (AIAS), een verdiepend onderzoek uit te voeren naar het (gecorrigeerde) beloningsverschil tussen mannen en vrouwen. In dit rapport doen wij verslag van onze bevindingen. Hierbij is uitdrukkelijk niet gekozen voor een onderzoek naar belonings*discriminatie* (hiervoor zou een aanvullend onderzoek met behulp van de vignettenmethode nodig zijn) maar ligt de nadruk op mogelijke nog niet eerder empirisch onderzochte aanvullende verklaringen voor belonings*verschillen* tussen mannen en vrouwen. Vertrekpunt is het onderzoek naar beloningsverschillen dat in 2006 is uitgevoerd door de Arbeidsinspectie. Vanwege de verschillen tussen sectoren bij de totstandkoming van de belangrijkste databron kunnen de resultaten van het onderzoek slechts vergeleken worden binnen de overheid apart en binnen de marktsector apart. Het onderzoek leent zich niet voor interpretatie van (verschillen in) loonverschillen *tussen* overheid en marktsector.

In Hoofdstuk 2 gaan we nader in op de opzet van de analyses, en geven we de mogelijkheden en beperkingen aan van toekomstige herhaling van de methode op andere databronnen. Hoofdstuk 3 beschrijft de resultaten van onze replicatie van het bovengenoemde eerdere onderzoek en wat er met dezelfde databron nog meer mogelijk is: de bijbehorende betrouwbaarheidsintervallen, enkele aanpassingen in de modelspecificatie alsmede enkele uitbreidingen met informatie die niet eerder is gebruikt. Vervolgens rapporteren we in Hoofdstuk 4 de resultaten van toepassing van een alternatieve schattingsmethode ('matching') op dezelfde databron. Daarmee hebben we de eerdere databron uitgeput en gaan we in Hoofdstuk 5 op zoek naar wat er in de literatuur nog meer aan verklaringen voor de loonkloof man-vrouw genoemd wordt. Tot slot willen we enkele van deze alternatieve verklaringen empirisch onderzoeken met behulp van een nieuwe databron, de 'Loonwijzer'. In hoeverre deze alternatieve verklaringen relevant zijn voor de verklaring van de loonkloof rapporteren we in Hoofdstuk 6. Tot slot verzamelen we in Hoofdstuk 7 nog enkele kanttekeningen en formuleren we aanbevelingen voor vervolgonderzoek naar (seksegerelateerde) beloningsverschillen.

2 Onderzoeksopzet & herhaalbaarheid

2.1 Onderzoeksopzet

De onderzoeksbehoefte laat zich volgens ons vertalen in de volgende onderzoeksvraag:

Welke aanpak/analyse kan leiden tot een uitgebreidere verklaring van de loonkloof tussen mannen en vrouwen, aanvullend op het onderzoek van de Arbeidsinspectie?

Het antwoord op deze onderzoeksvraag dient, met het oog op toekomstig beloningsonderzoek, gericht te zijn op herhaalbare methodieken. Het onderzoek van de Arbeidsinspectie zoals gerapporteerd in ‘De arbeidsmarktpositie van werknemers in 2004’ is het startpunt van onze analyse. De onderzoeksopzet valt uiteen in meerdere onderdelen:

- Bepalen van de betrouwbaarheidsintervallen van de schattingen uit het eerdere onderzoek. Hoe scherp zijn de gerapporteerde verschillen eigenlijk?
- Optimaliseren van de gebruikte modelspecificatie van regressies op de dataset van het Arbeidsvoorwaardenonderzoek (AVO). Bijvoorbeeld het toestaan van interactie-effecten en hogere orde-effecten. Zie Bijlage A en Bijlage B voor een beschrijving van de AVO-dataset.
- Toepassen van een alternatieve, in theorie efficiëntere, schattingsmethode. Leidt ‘propensity score matching’ tot dezelfde uitkomsten?
- Tenslotte het toevoegen van uitgebreidere verklaringen met behulp van informatie die niet in de AVO-dataset aanwezig is maar wel uit andere onderzoeken en bestaande gegevens gehaald kan worden. In hoeverre kunnen we aantonen dat deze alternatieve verklaringen een (significante) rol spelen?

2.1.1 Betrouwbaarheidsintervallen puntschattingen

Een eerste belangrijk item is het betrouwbaarheidsinterval. In de bijlagen van het Arbeidsinspectie-onderzoek is helaas niet te achterhalen hoe groot het betrouwbaarheidsinterval over het onverklaarde loonverschil is. Er worden geen standaarddeviaties aangegeven. De conclusie dat het gecorrigeerde loonverschil tussen mannen en vrouwen in het bedrijfsleven 7% bedraagt, is een puntschatting. Het is bekend dat dit een significant van nul afwijkende waarde betreft, dus het 95%-betrouwbaarheidsinterval is niet zo groot dat het ook de waarde nul kan bevatten. Maar het is niet uitgesloten dat het loonverschil in de populatie in werkelijkheid 6% of 8% kan zijn. Voor een volledige interpretatie moet dit uitgezocht worden, de uitkomsten staan vermeld in paragraaf 3.2 en Bijlage C.

2.1.2 Alternatieve modelspecificatie

Bij de specificatie van het te schatten model is sprake van een afweging tussen detaillering en betrouwbaarheid. Vallen de betrouwbaarheidsintervallen op efficiënte wijze verder te minimaliseren door bijvoorbeeld interactie-effecten op te nemen, of tweede-orde effecten? Als de

werkelijke effecten niet geheel gevangen kunnen worden in de huidige onafhankelijke lineaire relaties, is het goed mogelijk dat een iets complexere modelspecificatie zorgt voor meer efficiënte schattingen en kleinere onbetrouwbaarheidsmarges. Resultaten van deze onderzoeksstap zijn vermeld in paragraaf 3.3.

2.1.3 Alternatieve schattingsmethode

Bij het meten van loonverschillen wordt bijna altijd gebruik gemaakt van Mincer-vergelijkingen: een multivariate regressie van loonbepalende factoren op de logaritme van het uurloon. Met deze methode wordt de variatie in het loon toegeschreven aan de variatie in allerlei kenmerken. Personen worden nooit direct met elkaar vergeleken. In theorie zou met behulp van de alternatieve matching-methode nauwkeuriger loonvergelijkingen uitgevoerd moeten kunnen worden. Bij matching wordt het loon van koppels van zoveel mogelijk op elkaar lijkende mannen en vrouwen vergeleken. Bij een vrouw met bepaalde karakteristieken wordt een man met zoveel mogelijk dezelfde karakteristieken gezocht. Het gemiddelde van het beloningsverschil over al deze koppels van mannen en vrouwen geeft een nog betere kijk op het bestaan van structurele beloningsverschillen tussen verder vrijwel identieke mannen en vrouwen. Een nadeel van de matching methode is dat het vinden van goed vergelijkbare mannen en vrouwen niet op voorhand is verzekerd. Dit zal daarom onderdeel zijn van het onderzoek. De resultaten worden gerapporteerd in Hoofdstuk 4.

2.1.4 Alternatieve verklaringen

Een uitgebreide verklaring van factoren die bijdragen aan de loonkloof tussen mannen en vrouwen strekt natuurlijk veel verder dan wat toevallig beschikbaar is aan kwantitatieve gegevens in één enkele registratie. Hiervoor zijn meerdere gegevens van belang, waaronder ook subjectieve en kwalitatieve factoren. Enerzijds kunnen die gevonden worden in bestaande literatuur. Anderzijds zijn dit soort factoren bijvoorbeeld aanwezig in de Loonwijzer dataset. We zullen daarom eerst in Hoofdstuk 5 de resultaten van een multidisciplinaire literatuurstudie uitvoeren, om te achterhalen welke alternatieve verklaringen zoal genoemd worden. Tenslotte zullen we in Hoofdstuk 6 wederom een empirisch onderzoek uitvoeren, maar nu inclusief subjectieve variabelen (voor zover beschikbaar en relevant gebleken). Hierbij zullen we zoveel mogelijk dezelfde achtergrondvariabelen gebruiken als in de eerdere onderzoeksstappen, maar nu toegepast op de Loonwijzer-dataset. Een beschrijving van deze dataset is opgenomen in 0. Hierbij is niet het doel opnieuw een loonverschil man-vrouw te bepalen (daarvoor heeft de Loonwijzer dataset teveel nadelen) maar vooral het relatieve belang van enkele 'softe factoren' te toetsen.

2.2 Herhaalbaarheid

Tot op heden werd voor het Ministerie van Sociale Zaken & Werkgelegenheid tweemaal een onderzoek naar beloningsverschillen uitgevoerd door de Arbeidsinspectie, op basis van gegevens uit het Arbeidsvoorwaardenonderzoek (AVO-bestand). In de toekomst is het de bedoeling dat dit tweemaaljaarlijkse loononderzoek niet meer uitgevoerd gaat worden door de Arbeidsinspectie, maar door het CBS, op een ander nog niet openbaar gegevensbestand. Mede hierom zal ons onderzoek grotendeels gericht zijn op het zoeken naar verbeteringen in de onderzoekstechnieken, zodanig dat de resultaten hiervan in principe herhaalbaar zijn op andere databronnen. In de praktijk zal de

door het CBS te gebruiken databron uiteraard afwijken van de tot nu toe gebruikte AVO-data. Waarschijnlijk zullen bepaalde variabelen niet langer beschikbaar zijn, wellicht zullen een aantal nieuwe variabelen beschikbaar komen die dat momenteel nog niet zijn, en zonder meer zal de gewijzigde systematiek van dataverzameling consequenties hebben voor de kwaliteit en betrouwbaarheid van de gegevens. De kwaliteit en betrouwbaarheid van een dergelijk onderzoek zal hier sterk mee samenhangen. In principe zijn de door ons voorgestelde en empirisch verder te specificeren wetenschappelijke technieken herhaalbaar, maar dat betekent niet dat er simpelweg sprake kan zijn van een 'herhalingsonderzoek' wanneer die empirische specificaties worden toegepast op een andere databron. Elk toegepast onderzoek is immers maatwerk, en simpelweg toepassen van de voor dataset A gespecificeerde onderzoekstechnieken op dataset B is vaak niet optimaal.

Belangrijkste punt van aandacht is het verschil tussen survey-gegevens en registratie-gegevens. Survey gegevens worden verzameld met een specifiek onderzoeksdoel, meestal expliciet gericht op het doen van wetenschappelijke analyses. Het nadeel van survey-gegevens is de hoge kosten van dataverzameling, zodat altijd slechts een steekproef van de totale populatie ondervraagd zal worden. Mits deze steekproef groot genoeg is kan dat zonder groot verlies van nauwkeurigheid. Registratiebestanden zijn altijd opgezet vanwege andere doeleinden dan het doen van wetenschappelijk onderzoek en het leggen van oorzakelijke verbanden. Hierdoor valt op voorhand niet te zeggen in hoeverre het CBS er in slaagt een *vergelijkbaar* bestand te creëren voor het doen van beloningsonderzoeken. Nadeel van registratiebestanden is het ontbreken van subjectieve variabelen ('hoe waardeert u uw inkomen?'). Ook is momenteel nog geen integrale registratie beschikbaar van het behaalde onderwijsniveau van werknemers, essentieel voor beloningsonderzoeken.⁷

Het nieuwe databestand van het CBS met beloningsgegevens over 2006 zal daarom gebaseerd zijn op de koppeling van verschillende registraties én surveys, in het bijzonder gegevens uit de polisadministratie afkomstig van de Belastingdienst en het UWV, de Enquête Beroepsbevolking en de Gemeentelijke Basisadministratie. De verwachting is dat in de tweede helft van 2008 de eerste nieuwe loongegevens (over de jaren 2005 en 2006) beschikbaar zullen zijn. In principe beschikt het CBS dan over de loongegevens van iedereen die aangifte doet bij de Belastingdienst. Maar om beloningsonderzoek uit te kunnen voeren is meer nodig, met name betrouwbare gegevens over gevolgde opleiding en de huidige baan. Hiervoor zal het nieuwe loongegevensbestand gekoppeld moeten worden aan de EBB, zodat een nieuw bestand ontstaat vergelijkbaar met de Loonstructuuronderzoeken dat in omvang beperkt wordt door de doorsnede van beide bestanden (de EBB zal ongeveer 90 duizend observaties bevatten, de andere bestanden zijn in theorie integraal; koppeling van drie jaargangen EBB vergoot het aantal observaties tot ruim 150 duizend). Het moge duidelijk zijn dat een op een dergelijke complexe manier opgebouwd databestand andere karakteristieken zal hebben dan het uit één bron stammende AVO-bestand.

⁷ Het CBS beschikt weliswaar over gegevens uit het CRIHO, Examenresultatenregister VO, onderwijsnummerbestand BO & VO, etc. maar deze verre van integraal. In feite bevatten ze alleen relevante informatie voor de meest recente cohorten, aangezien de (koppelbare) registraties nog niet zo lang bestaan.

De geschiktheid van nog niet bestaande bestanden kan niet vooraf beoordeeld worden. In principe zal gebruik van de gegevens van de Belastingdienst zeer betrouwbare loongegevens opleveren, maar zullen enkele specifieke gegevens zoals type werknemer en toepassing salarisschalen verdwijnen. Belangrijk is ook het ontbreken van inzicht in prestatiegerelateerde looncomponenten: alleen de totale loonsom is straks bekend. Uiteindelijk zal een bestand resulteren waarmee kleinere foutmarges bereikt kunnen worden ten aanzien van de afhankelijke variabele, loon.. Maar tevens zijn straks veel minder verklarende variabelen beschikbaar, die het huidige AVO-bestand juist zo geschikt maken voort de verklaring van loonverschillen. De ongecorrigeerde loonverschillen kunnen dus exacter weergegeven worden, maar de gecorrigeerde loonverschillen zeker niet.

Ondanks bovenstaande voorbehoud m.b.t. herhalingsonderzoek, menen wij dat het wel degelijk zinvol is om te onderzoeken hoe er extra informatie verkregen kan worden uit bestaande data. Dat kan wellicht dienen als 'ideaalsituatie', om inzicht te geven in welke belangrijke factoren nog een rol spelen buiten de nieuwe modelschattingen om. Tevens wordt zo duidelijk welke lacunes er nog zijn in het toekomstige loonverschillenonderzoek, en wat er nog wenselijk zou zijn aan nieuwe dataverzameling.

3 Uitbreiding bestaande analyse binnen traditionele schattingsmethode

Loonschattingen: betrouwbaarheid, modelspecificatie en uitbreiding

In dit hoofdstuk baseren we ons op de gegevens die verzameld zijn ten behoeve van het eerdere onderzoek van over loonverschillen dat eerder is uitgevoerd door de Arbeidsinspectie, op basis van de AVO-dataset. In paragraaf 3.1 beschrijven we een replicatie van de loonschatting in dit onderzoek, die we de ‘basisreplicatie’ noemen. Met deze onderzoeksresultaten als uitgangspunt gaan we vervolgens (in paragraaf 3.2) dieper in op de betrouwbaarheid van de puntschattingen. In paragraaf 3.3 besteden we uitgebreid aandacht aan de verbeteringen en uitbreidingen die mogelijk zijn op deze basisloonschatting. Bijlage A bevat de definitie van alle variabelen in de AVO-dataset.

3.1 Basisreplicatie

In het AVO-onderzoek wordt een Mincer-achtige⁸ loonvergelijking geschat die er voor een persoon als volgt uitziet:

$$(1) \quad \ln w_s = \beta_{0,s} + \beta_{1,s} \text{Vrouw}_s + \mathbf{X}_s' \boldsymbol{\beta}_s + u_s \quad \text{voor } s = b, o$$

Subscript s geeft aan of een persoon werkzaam is voor een overheidsinstantie (o) of in het bedrijfsleven (b). In deze vergelijking stelt w_s het uurloon voor en deze wordt als volgt gedefinieerd:

$$w_s = \frac{\text{Bruto maandloon} + \text{som extra loon}}{\text{Aantal arbeidsuren per maand}}$$

Uit deze definitie volgt dat er verschillende inkomsten die niet direct gerekend kunnen worden tot het brutoloon worden meegerekend en we verwijzen hiernaar als de som van het extra loon. Onder het extra loon verstaan we het bedrag wat ontvangen is aan persoonlijke toelage, ploegentoeslag, tarief, provisie, inconveniënten, loon in natura en overige loon. De definitie van deze looncomponenten én de verklarende variabelen zijn opgenomen in Bijlage A.

Het bruto uurloon wordt gespecificeerd als de natuurlijke logaritme van het bruto uurloon. Daarmee wordt recht gedaan aan de scheve verdeling van uurlonen: de rechterstaart is langer en dikker dan de linker. De schattingsresultaten zijn door de exponentiële specificatie bovendien minder gevoelig voor eventuele grote uitschieters in het uurloon naar boven.

⁸ We gebruiken de term Mincer-achtig omdat Mincer in zijn baanbrekende artikel de loonvergelijking definieert als een functie waarbij de logaritme van het loon afhangt van het scholingsniveau in jaren, ervaringsjaren en ervaringsjaren in het kwadraat. Omdat we het aantal jaren scholing niet waarnemen kunnen we strict genomen niet spreken van een Mincer-vergelijking.

De variabelen die gebruikt worden om de variatie in het uurloon te verklaren staan aan de rechterkant van de vergelijking. Omdat we in deze studie in het bijzonder geïnteresseerd zijn in het loonverschil tussen mannen en vrouwen wordt in vergelijking (1) het effect van geslacht op uurloon apart vermeld. Dit effect wordt gemeten door een variabele op te nemen die 1 is als een persoon een vrouw is en anders 0. Zodoende wordt het loonverschil tussen mannen en vrouwen gerepresenteerd door de parameter $\beta_{1,s}$.

De overige variabelen waarvan we vermoeden dat ze de waargenomen variatie in het uurloon verklaren zitten in de matrix \mathbf{X}_s . Deze variabelen zijn van cruciaal belang. Als een belangrijke variabele wordt weggelaten dan kan de dummy ‘vrouw’ een deel van dit effect oppikken en resulteert een onzuivere meting van het loonverschil tussen mannen en vrouwen. De kenmerken die we meenemen voor zowel de overheidssector en het bedrijfsleven zijn:

- geslacht,
- leeftijd,
- arbeidsduur,
- anciënniteit,
- som van het extra loon,
- economische sector waarin de persoon werkzaam is,
- functieniveau.

Beschrijvende statistieken van de verklarende variabelen zijn opgenomen in Bijlage A. Voor het bedrijfsleven is in de dataset meer informatie beschikbaar dan voor de overheidssector. Door de loonvergelijking apart te schatten voor het bedrijfsleven en de overheidssector, kunnen we voor het bedrijfsleven ook nog de volgende kenmerken opnemen:

- etnische afkomst,
- type arbeidscontract,
- functiesoort,
- opleidingsniveau,
- bedrijfsregio,
- werknemerstype,
- of betaalde lonen gebaseerd zijn op een salarisschaal,
- grootteklasse.

Door deze kenmerken ook op te nemen in de loonvergelijking zal het loonverschil zuiverder gemeten worden. Echter, er is ook een keerzijde aan het apart schatten van de loonvergelijking. Het type baan waarin personen werkzaam zijn is het gevolg van een keuze. Deze baankeuze kan tot gevolg hebben dat sommige personen systematisch eerder kiezen voor een baan in de overheidssector terwijl andere personen systematisch eerder kiezen voor een baan in de bedrijfssector. Kortom, er vindt een zelfselectie plaats naar sector en door het apart schatten van de loonvergelijking zijn we niet in staat om te corrigeren voor deze zelfselectie. Dus door meer relevante variabelen op te nemen in de loonvergelijking meten we het loonverschil zuiverder maar tegelijkertijd ontstaat er een onzuiverheid doordat de zelfselectie naar sector genegeerd wordt.

Doordat we voor de overheidsector verschillende belangrijke variabelen niet meenemen in de loonvergelijking is het niet mogelijk om de loonverschillen die uiteindelijk gevonden worden in

de overheidssector en het bedrijfsleven met elkaar te vergelijken en ten opzichte van elkaar te evalueren. Dit komt omdat we in de loonschatting voor de overheidssector niet kunnen corrigeren voor enkele belangrijke samenstellingseffecten zoals opleiding en functiesoort terwijl dit wel gebeurt voor het bedrijfsleven. We benadrukken daarom dat we in dit onderzoek de overheidssector en het bedrijfsleven als twee aparte arbeidsmarkten beschouwen en dat we apart voor elke arbeidsmarkt bestuderen hoe dit geschatte loonverschil varieert over de verschillende gekozen specificaties.

In Tabel 3.1 presenteren we de geschatte loonverschillen tussen mannen en vrouwen. De loonverschillen worden verkregen op basis van schattingen voor de totale steekproef en voor verschillende subgroepen. De loonverschillen representeren de schattingsparameter, $\beta_{1,v}$, die behoort bij de variabele vrouw. Hieruit volgt direct dat een negatieve waarde aangeeft dat vrouwen gemiddeld minder verdienen dan mannen. Voor de volledigheid rapporteren we de schattingsresultaten voor de totale steekproef in Bijlage D.

De eerste rij toont de loonverschillen tussen mannen en vrouwen, apart voor de steekproef van het bedrijfsleven en de steekproef van de overheidssector. Omdat de kwaliteit en representativiteit van deze steekproeven nogal verschillen, kunnen de uitkomsten van de beide sectoren niet zomaar met elkaar worden vergeleken.

Om een indruk te krijgen over hoe de loonkloof m-v varieert tussen verschillende subgroepen is de loonvergelijking telkens geschat voor al deze subgroepen. Voor de overheidssector geldt dat het loonverschil toeneemt naarmate personen ouder zijn, personen minder dan 12 uur werken of als personen een baan hebben met een hoog functieniveau. Ook voor het bedrijfsleven vinden we een groter loonverschil naarmate personen ouder zijn. Echter voor personen ouder dan 55 jaar geldt dat het loonverschil geldt dat het loonverschil vergeleken kan worden met dat van personen in de leeftijdscategorie van 23 tot 35 jaar.

Hoe hoger het functieniveau hoe groter het loonverschil bij het bedrijfsleven. Deze gevonden relatie kan gerelateerd zijn aan het feit dat beter betaalde functies (nog) veelal bekleed worden door mannen. In tegenstelling tot de overheidssector waar het loonverschil het grootst is tussen mannen en vrouwen met een deeltijdbaan minder dan 12 uur geldt voor het bedrijfsleven dat de loonverschillen groter zijn tussen mannen en vrouwen met een voltijdbaan. In eerste instantie ligt het voor de hand om dit verschil toe te schrijven aan het feit dat we voor het bedrijfsleven arbeidscontract-kenmerken opnemen en voor de overheidssector niet. Echter, ook voor de arbeidscontract-variabelen zien we dat de loonverschillen groter zijn tussen mannen en vrouwen met een regulier- of uitzendbaan dan tussen mannen en vrouwen met een flexibel arbeidscontract wat juist het verschil onderstreept.

Verder is voor het bedrijfsleven het loonverschil tussen autochtone mannen en vrouwen groter dan het loonverschil tussen allochtone mannen en vrouwen. Ook is het loonverschil groter naarmate de bedrijfsgrootte afneemt. Naar verwachting zijn de loonverschillen tussen mannen en vrouwen groter naarmate de economische sector een sector betreft waarin mannen beter vertegenwoordigd zijn. Geordend van het grootste loonverschil naar het kleinste loonverschil gaat het hier om de volgende sectoren: bouw, transport, landbouw en visserij en handel.

Tabel 3.1 Geschatte loonkloof m-v 2004, bedrijfsleven en overheidssector, naar subpopulatie

Kenmerk	Subgroep	Loonverschil bedrijfsleven	Loonverschil overheidssector
Alle personen		-0,074	-0,040
Etnische afkomst	Autochtoon	-0,077	
	Allochtoon	-0,046	
Arbeidsduur	Voltijd	-0,093	-0,038
	Deeltijd \geq 12 uur	-0,047	-0,032
	Deeltijd $<$ 12 uur	-0,026	-0,087
Arbeidscontract	Regulier/uitzend bepaalde tijd	-0,085	
	Regulier/uitzend onbepaalde tijd	-0,061	
	Flexibel onbepaalde tijd	-0,033	
	Flexibel bepaalde tijd	0,010	
Leeftijd	Tot en met 22 jaar	-0,031	-0,002
	23 tot 35 jaar	-0,050	-0,019
	35 tot 45 jaar	-0,088	-0,030
	45 tot 55 jaar	-0,101	-0,045
	55 jaar en ouder	-0,053	-0,055
Grootteklasse	Kleinbedrijf (1-9)	-0,102	
	Middenbedrijf (10-99)	-0,078	
	Grootbedrijf (100 of meer)	-0,064	
Functieniveau	Functieniveau laag	-0,031	-0,022
	Functieniveau midden	-0,081	-0,011
	Functieniveau hoog	-0,105	-0,027
	Functieniveau onbekend		-0,132
Economische sector	Landbouw en visserij	-0,091	
	Industrie	-0,079	
	Bouw	-0,114	
	Handel	-0,086	
	Horeca	-0,047	
	Transport	-0,102	
	Zakelijke dienstverlening	-0,063	
	Gezondheids- en welzijnszorg	-0,060	
	Overige dienstverlening	-0,052	

3.2 Betrouwbaarheid van de puntschatter

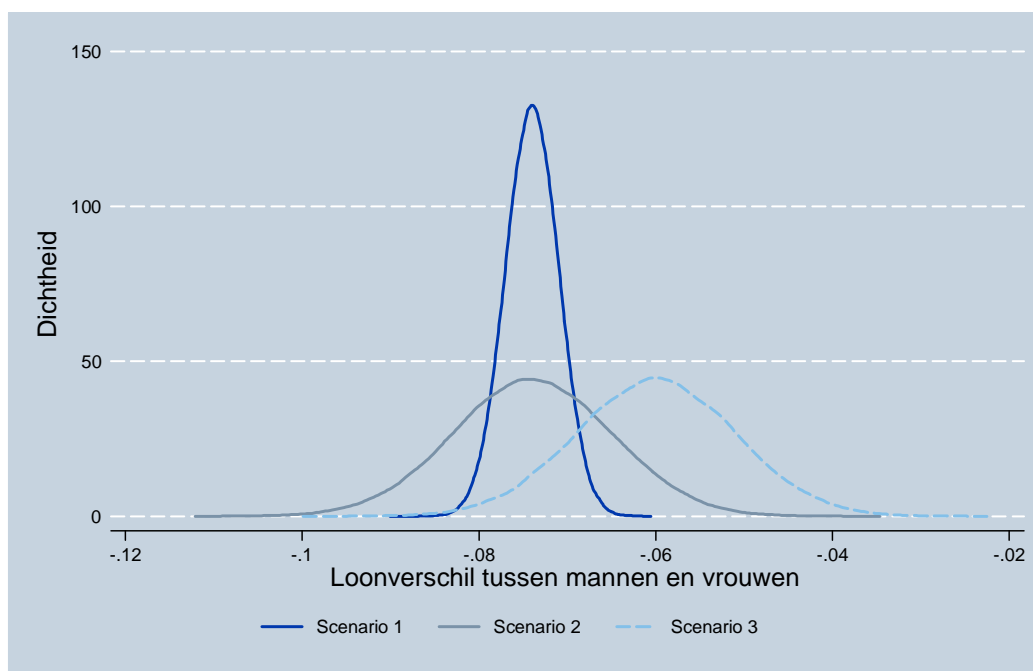
Loonverschillen verkregen op basis van een econometrische analyse worden veelal geëvalueerd op basis van hun grootte. Naar deze grootte wordt dan verwezen als de puntschatter. Een ander belangrijk aspect is de betrouwbaarheid van deze puntschatter maar deze wordt in de regel minder of niet belicht.

3.2.1 Technische verantwoording

Een indicatie over de betrouwbaarheid van een puntschatter wordt verkregen door gebruik te maken van een betrouwbaarheidsinterval. In de praktijk wordt een betrouwbaarheidsinterval vaak verkeerd geïnterpreteerd en daarom besteden we hieraan kort aandacht. Stel dat het loonverschil tussen mannen en vrouwen voor het bedrijfsleven geschat wordt op 7.4 procent, zoals in paragraaf 3.1. Op basis van de steekproef kunnen we dan een interval berekenen rondom de puntschatting waarin met grote waarschijnlijkheid het *werkelijke* loonverschil ligt. Naar dit interval wordt verwezen als het betrouwbaarheidsinterval en dit wordt aangeduid met een percentage, bijvoorbeeld het '95-procent betrouwbaarheidsinterval'. De interpretatie hiervan is dat wanneer we uit dezelfde populatie een groot aantal willekeurige steekproeven zouden trekken en hiervoor het betrouwbaarheidsinterval berekenen, in ongeveer 95 procent van de gevallen het daadwerkelijke populatiegemiddelde binnen het berekende interval ligt.

In hoeverre een betrouwbaarheidsinterval informatief is wordt verduidelijkt aan de hand van een voorbeeld in Figuur 3.1. Daarin tonen we drie hypothetische verdelingen van het loonverschil tussen mannen en vrouwen en naar deze verdelingen verwijzen we als scenario 1, 2 en 3. Voor elke verdeling geeft de horizontale as de loonverschillen weer en de verticale as de dichtheid.

Figuur 3.1 Simulaties loonverschillen tussen mannen en vrouwen



De verdeling in scenario 1 is de verdeling van het loonverschil tussen mannen en vrouwen op basis van de loonschatting voor het bedrijfsleven. Deze verdeling heeft een gemiddeld loonverschil tussen mannen en vrouwen van -7.4 procent. In scenario 2 is een verdeling gesimuleerd waarbij het gemiddelde gelijk is aan scenario 1 maar waarbij de variatie van de schatter van het loonverschil groter is (3 keer zo groot om precies te zijn). In scenario 3 is de variatie van de schatter van het loonverschil gelijk aan dat van scenario 2 maar is de gemiddelde schatting van het loonverschil - 6 procent. We merken op dat het negatieve loonverschil aangeeft dat de vrouw minder verdient dan de man.

We richten ons eerst op verdelingen in scenario 1 en 2. Voor beide verdelingen kunnen we met vrijwel 100 procent zekerheid stellen dat het loonverschil tussen mannen en vrouwen zal liggen tussen de -0.12 en de -0.03. Dit is echter niet erg informatief als we willen weten welke schatter betrouwbaarder is. Voor scenario 1 kunnen we met vrijwel 100 procent zekerheid stellen dat het loonverschil in de populatie zal liggen tussen de -0.9 en -0.6 maar dit geldt niet voor scenario 2. In de regel definieert men een 95 procent betrouwbaarheidsinterval. Op basis van de gebruikte steekproef wordt dan minimaal en maximaal loonverschil berekend zodat we met 95 procent zekerheid kunnen stellen dat het gedefinieerde betrouwbaarheidsinterval het werkelijke loonverschil bevat. Als we dit interval bepalen voor scenario 1 en 2 dan zorgt de kleinere (standaard) variatie in de schatter voor scenario 1 ervoor dat het betrouwbaarheidsinterval kleiner zal zijn dan dat van scenario 2. Zodoende wordt scenario 1 gezien als de meer betrouwbare schatter op basis van de verdeling van de schatter. Bestudering van de verdelingen van de puntschatter in scenario 1 en 2 is wat overbodig omdat beide puntschatters gelijk zijn. In dit geval kan direct aan de standaardfout gezien worden welke puntschatter meer betrouwbaar is.

Tevens vergelijken we de verdeling van scenario 1 met die van scenario 3. Stel dat de situatie in scenario 3 het geschatte loon representeert dat gevonden wordt als er een extra relevante verklarende variabele wordt meegenomen in het model. Als er puur geëvalueerd wordt op de grootte van de puntschatter dan concluderen we dat het loonverschil minder groot wordt na de toevoeging van deze relevante verklarende variabele. Deze conclusie is echter voorbarig. Doordat de standaarddeviatie van de puntschatter in scenario 3 groter is dan die van scenario 1 is er een substantiële overlap tussen beide verdelingen. Wanneer we een 95 procent betrouwbaarheidsinterval zouden bepalen voor beide verdelingen dan is het niet onwaarschijnlijk dat de echte populatieparameter een waarde heeft die in beide intervallen voorkomt. Het betrouwbaarheidsinterval behorende bij scenario 3 sluit daarom niet uit dat de populatie parameter ligt in het betrouwbaarheidsinterval bepaald voor scenario 1. Bovendien zorgt de grotere standaarddeviatie voor meer onzekerheid over de werkelijke waarde van de populatieparameter.

We kiezen er in de volgende paragraaf niet voor om de 95 procent betrouwbaarheidsintervallen te bespreken. In plaats daarvan tonen we de verdelingen van de verkregen puntschatters zoals in Figuur 3.1. Deze verdelingen worden gesimuleerd op basis van het geschatte loonverschil en de bijbehorende variatie van de puntschatter.⁹ Dit heeft als voordeel dat de interpretatie gelijk is aan het voorbeeld van Figuur 3.1. Bovendien wordt op deze manier duidelijker in hoeverre de verdelingen van de geschatte puntschatters elkaar overlappen. Voor de volledigheid worden de 95

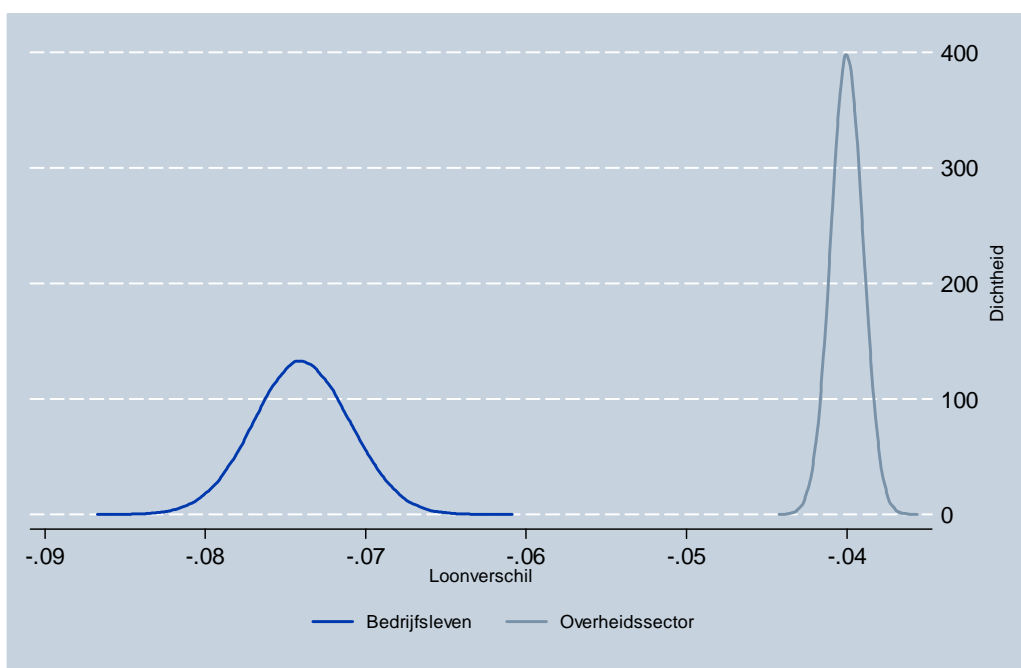
⁹ Op basis van het geschatte loonverschil a en de bijbehorende standaardfout sd simuleren we de verdeling van het loonverschil op basis van 10000 trekkingen waarbij de normale verdelingsfunctie er als volgt uitziet: $a + sd \cdot N^{-1}(U(R))$. In deze verdeling staat R voor een willekeurig getal getrokken uit een uniforme verdeling (U) en staat N^{-1} voor de inverse normale verdeling.

procent betrouwbaarheidsintervallen van de hieronder getoonde verdelingen gepresenteerd in Bijlage C.

3.2.2 Betrouwbaarheidsgrafieken loonkloof AVO-onderzoek

In deze paragraaf tonen we voor alle puntschattingen uit Tabel 3.1 de betrouwbaarheidsgrafieken, zoals beschreven in de technische toelichting in voorgaande paragraaf. Allereerst tonen we in Figuur 3.2 de verdeling van het geschatte loonverschil tussen mannen en vrouwen voor het bedrijfsleven en de overheidssector.

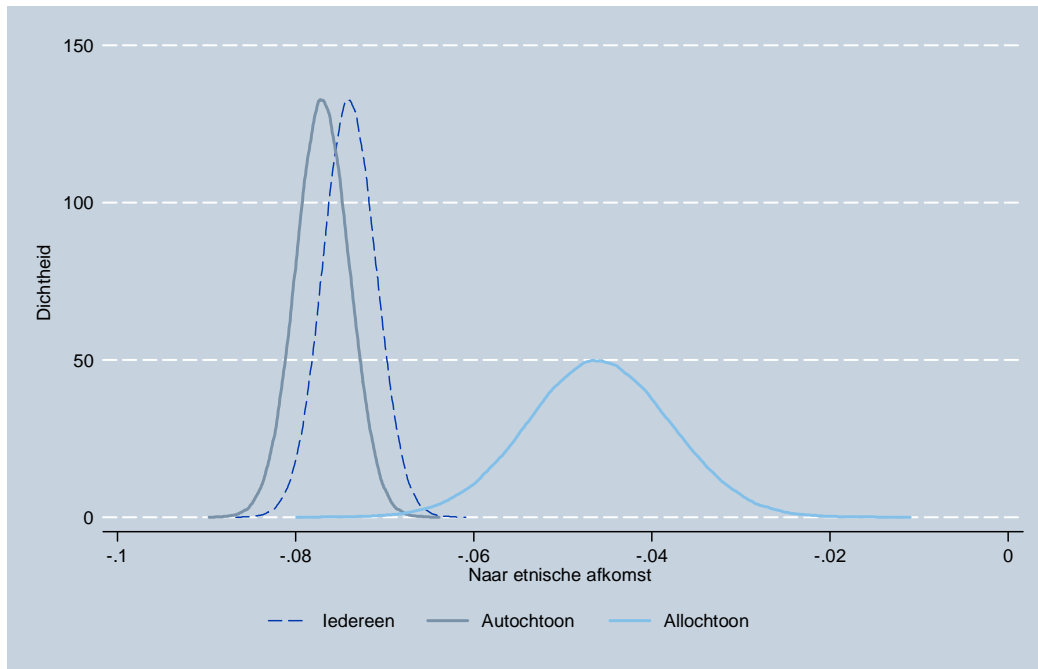
Figuur 3.2 Verdeling van het loonverschil tussen mannen en vrouwen



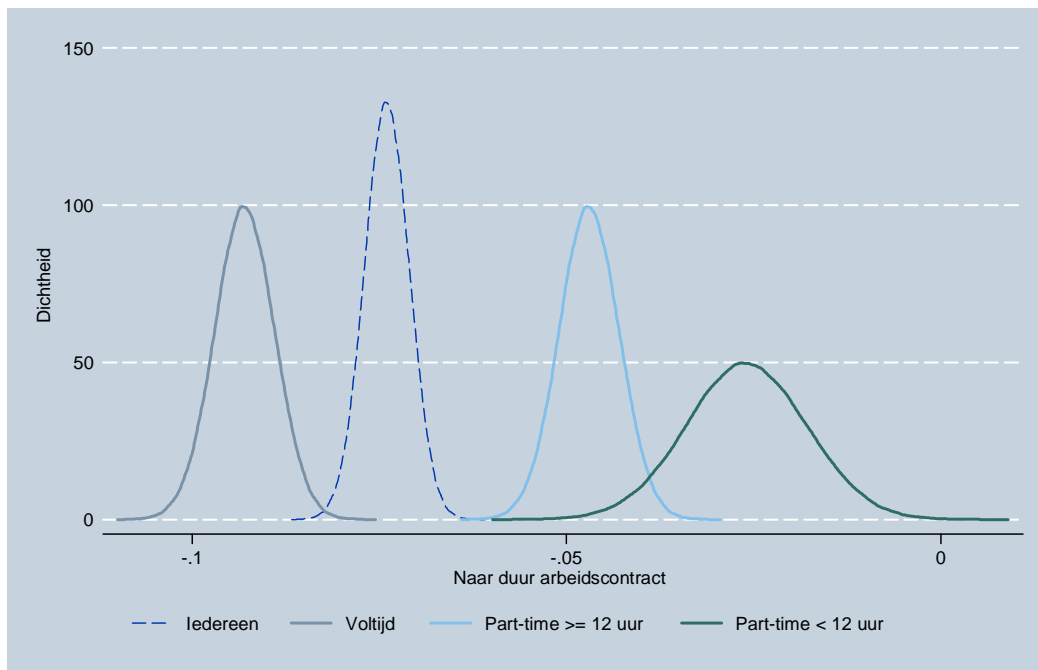
Het loonverschil tussen mannen en vrouwen is niet alleen groter in het bedrijfsleven maar bovendien is de gemeten variatie van de puntschatter groter voor het bedrijfsleven. Het loonverschil tussen mannen en vrouwen wordt daarom minder nauwkeurig gemeten voor het bedrijfsleven. De grotere variatie komt waarschijnlijk doordat de lonen in het bedrijfsleven sowieso veel minder uniform zijn dan in de overheidssector, waar vaker met vaste functiewaarderingsystemen gewerkt wordt.

Figuur 3.3 Verdeling van de puntschatter met betrekking tot het loonverschil tussen mannen en vrouwen in het bedrijfsleven

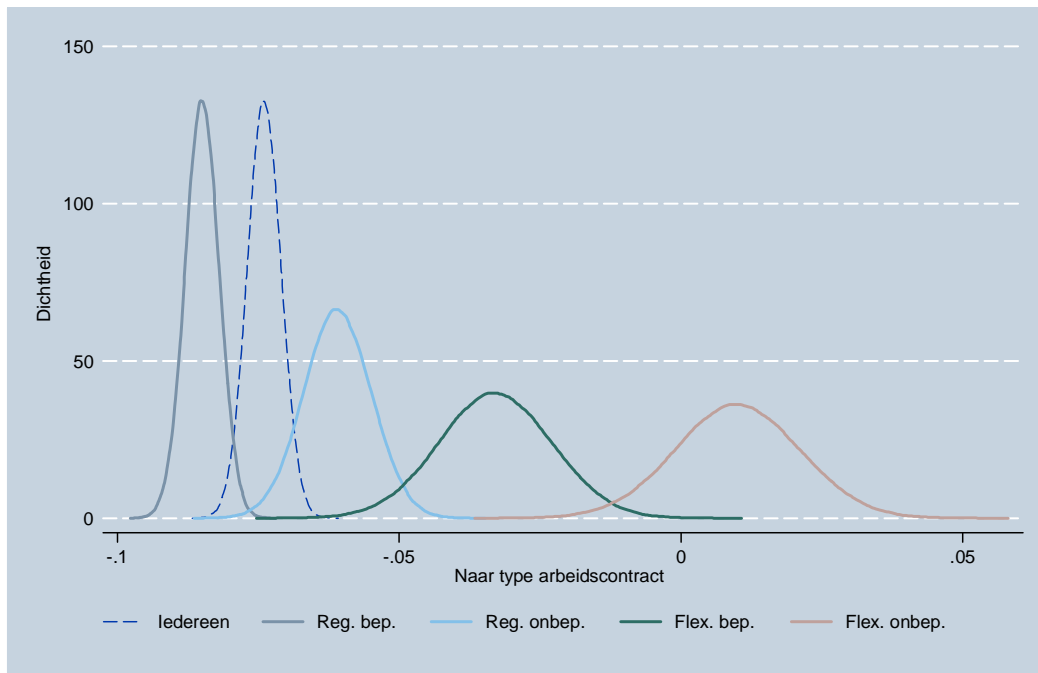
1. Verdelingen naar etnische afkomst:



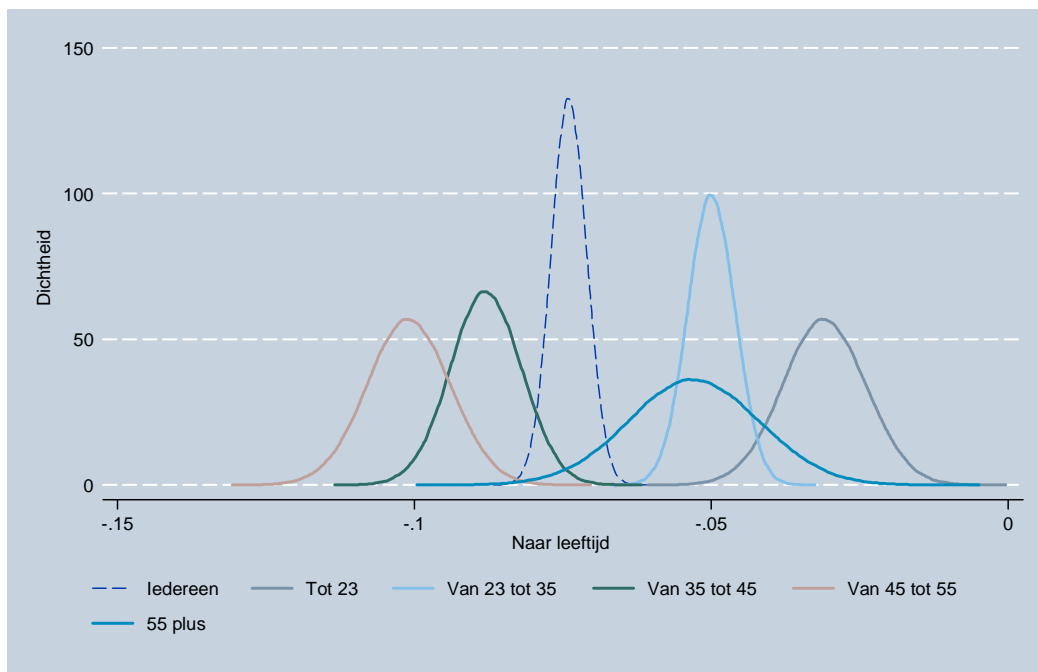
2. Verdelingen naar arbeidsduur:



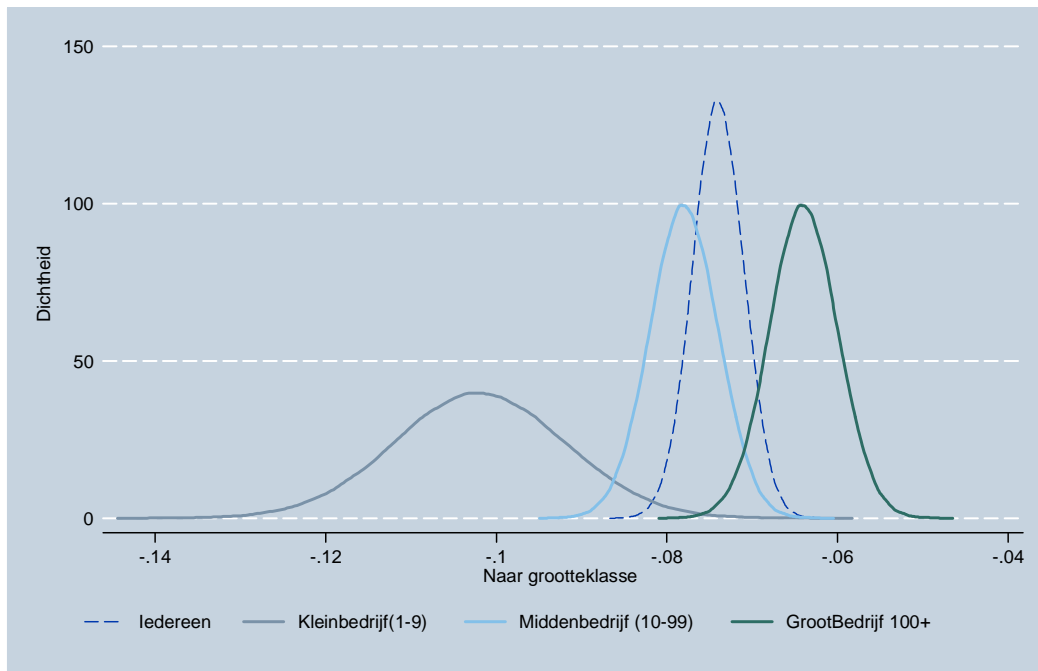
3. Verdelingen naar type arbeidscontract:



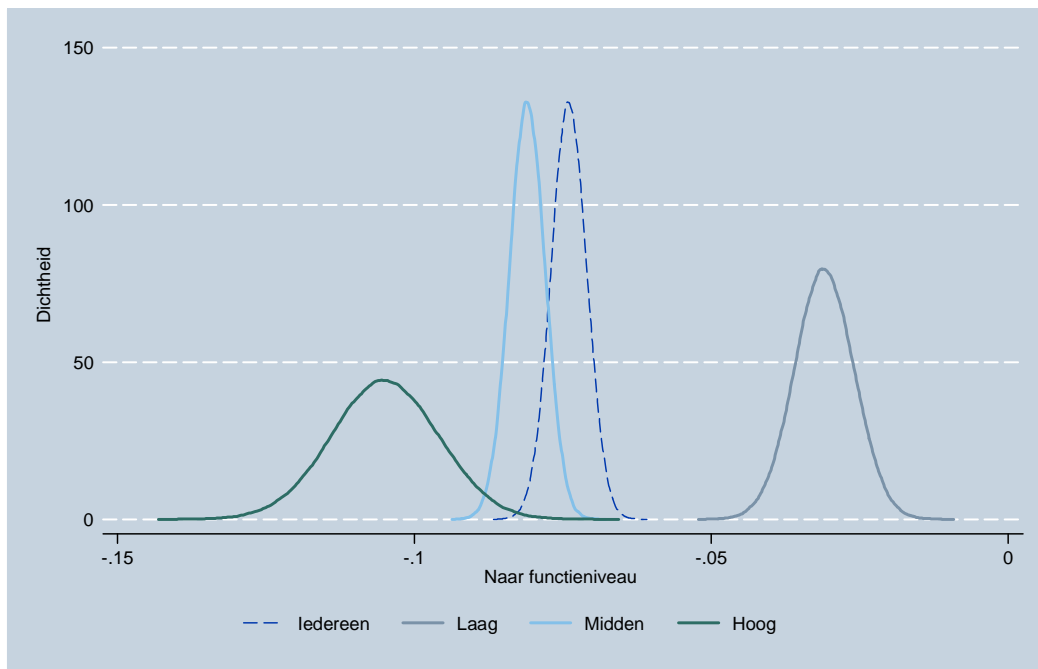
4. Verdelingen naar leeftijdsklassen:



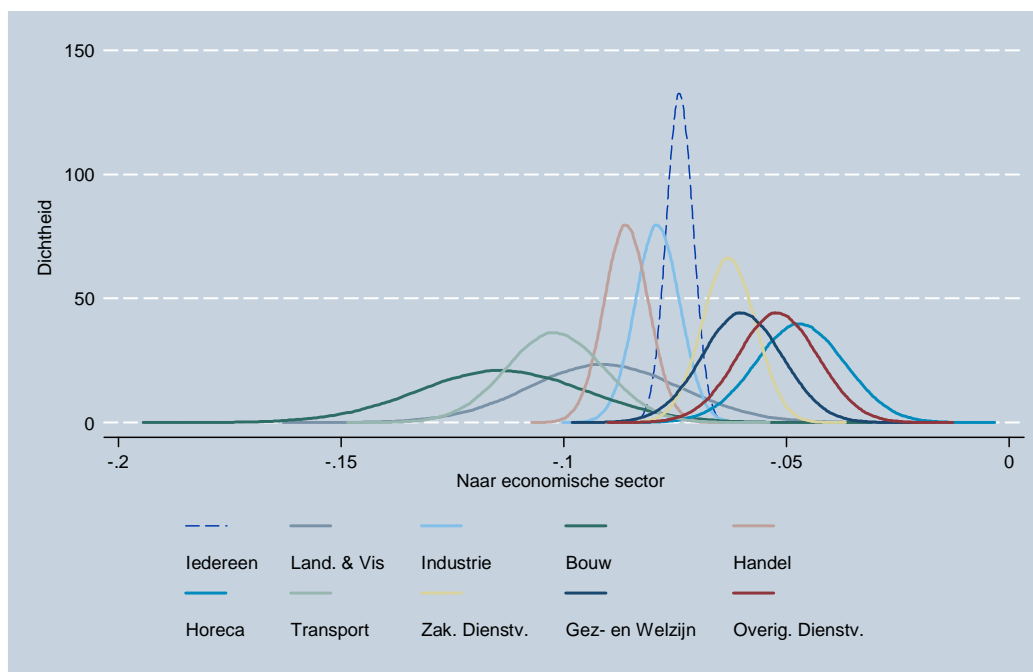
5. Verdelingen naar grootteklasse:



6. Verdelingen naar functieprofiel:



7. Verdelingen naar economische sector:



In Figuur 3.3 worden de verdelingen getoond voor het bedrijfsleven. Voor het gemak verwijzen we naar de verdeling van alle personen in het bedrijfsleven als de verdelingen van de bedrijfspopulatie. We vergelijken de subgroepen met de gehele bedrijfspopulatie.

De verdeling van het geschatte loonverschil voor autochtonen benadert dat van de bedrijfspopulatie, terwijl de loonkloof onder allochtonen relatief kleiner is en onnauwkeuriger gemeten wordt. De grotere spreiding van de verdeling bij allochtonen komt voor een deel doordat het aantal personen in deze groep relatief klein is in vergelijking tot de bedrijfspopulatie. De variatie in de verdeling wordt kleiner en het loonverschil wordt nauwkeuriger gemeten, naarmate het aantal personen in de subpopulatie groter wordt. Dit betekent overigens niet dat het gemiddelde onzuiver gemeten wordt.

Zoals we zagen in paragraaf 3.1 is het loonverschil tussen mannen en vrouwen voor voltijders groter dan bij parttimers. De variatie in de puntschatters voor voltijders en parttimers die meer dan 12 uur werken komen goed overeen, de variatie onder parttimers die minder dan 12 uur werken is aanzienlijk groter. Eigenlijk is er nauwelijks overlap tussen de verschillende categorieën. Dat betekent dat de loonkloof m-v duidelijk en significant verschilt tussen voltijders, grote- en kleine parttimers. Op basis van het type arbeidscontract concluderen we dat de verdeling van de bedrijfspopulatie het meest overeenkomt met de verdeling voor personen die een regulier arbeidscontract hebben voor onbepaalde tijd.

Wanneer we de verdelingen uitsplitsen naar de verschillende leeftijdsklassen zien we dat door de hoeveelheid gedefinieerde subgroepen de variatie in de puntschatter toeneemt ten opzichte van de bedrijfspopulatie en dit komt doordat het aantal waarnemingen per subgroep kleiner is. Voor personen van 55 jaar en ouder is de variatie relatief het grootst. Ondanks dat er relatief weinig

jongeren zijn tot 23 jaar is de variatie in de puntschatter relatief gezien klein. Mogelijkerwijs komt dit door de lagere en meer uniforme lonen die aan hen betaald worden. De diverse verdelingen naar de verschillende leeftijdsklasse komen vrij goed overeen met uitzondering van de leeftijdscategorie 23 tot 35 jaar waarbij de variatie van de puntschatter relatief laag is. Overigens wordt het verschil in betrouwbaarheid van puntschatters mooi geïllustreerd door de verdelingen behorende bij leeftijdscategorieën 23 tot 35 jaar en 55 plus. Voor deze verdelingen ligt namelijk het gemiddelden van beide categorieën vrij dicht bij elkaar terwijl de variatie van de puntschatter van de 55 plus aanzienlijk groter is.

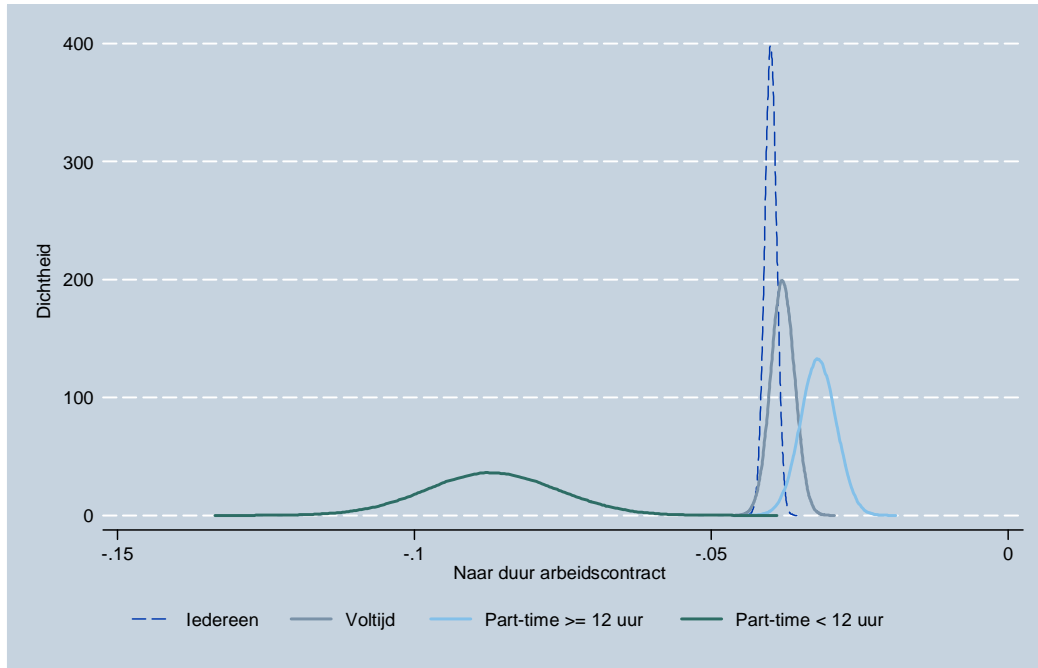
De variatie van het geschatte loonverschil is relatief groot voor het kleinbedrijf terwijl het geschatte loonverschil voor het midden- en grootbedrijf goed overeenkomen. De verdeling van de bedrijfspopulatie komt het meest overeen met die van het middenbedrijf ondanks dat de meeste personen werkzaam zijn in het grootbedrijf (Zie ook Bijlage C). Wanneer we het loonverschil tussen mannen en vrouwen apart schatten voor de verschillende functieprofielen geldt dat de variatie hoger is als het functieniveau hoog of laag is.

Wat betreft de economische sectoren kunnen we grofweg 3 categorieën onderscheiden. De eerste categorie bevat de economische sectoren handel, industrie en zakelijke dienstverlening. De tweede categorie bevat de sectoren zorg, horeca en overige dienstverlening; in deze categorie zijn de loonverschillen relatief klein maar is de variatie groter dan in de eerste categorie. De derde categorie bevat de sectoren bouw, landbouw/visserij en transport. De loonverschillen zijn in deze sector hoger in vergelijking tot de overige sectoren en bovendien is de variatie in de puntschatter aanzienlijk hoger. Kijken we bijvoorbeeld naar landbouw en visserij, waar we op basis van Tabel 3.1 een loonverschil schatten van -0.091 , dan is het in principe mogelijk dat de populatieparameter in werkelijkheid lager ligt dan het geschatte loonverschil in bv. de handel. De reden is dat er relatief weinig vrouwen werken in deze sectoren en zodoende zijn de verkregen loonverschillen onnauwkeurig vast te stellen.

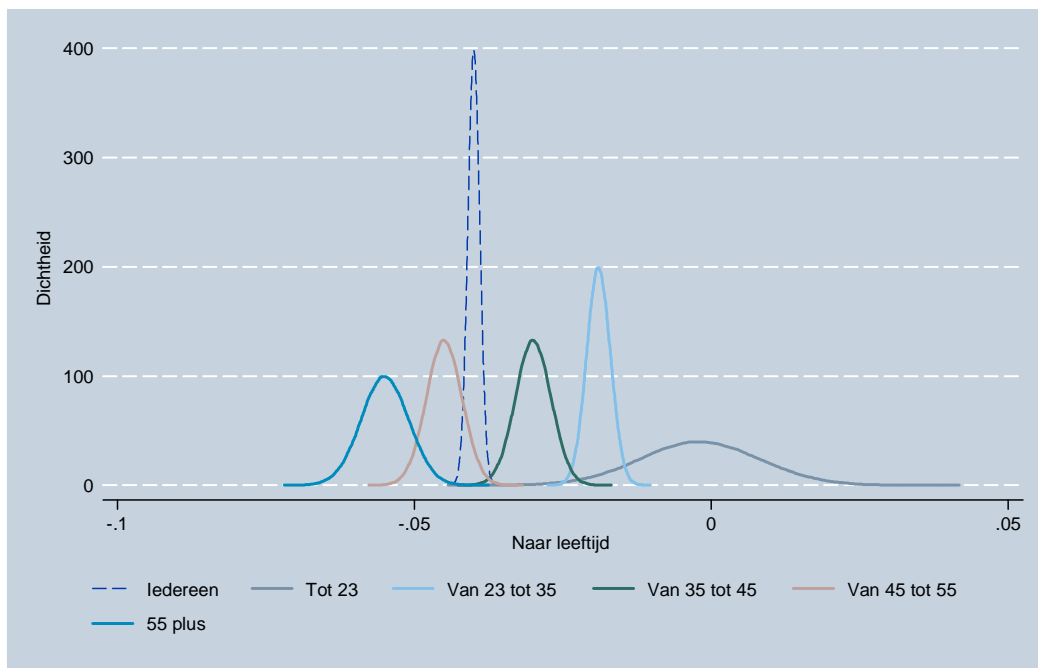
Op dezelfde manier als in Figuur 3.3 tonen we in Figuur 3.4 de verdelingen naar de kenmerken waarvoor informatie beschikbaar is, maar nu voor de overheidssector. Voor het gemak verwijzen we naar de verdeling van alle personen in de overheidssector als de verdelingen van de overheidspopulatie.

Figuur 3.4 Verdeling van de puntschatter met betrekking tot het loonverschil tussen mannen en vrouwen in de overheidssector

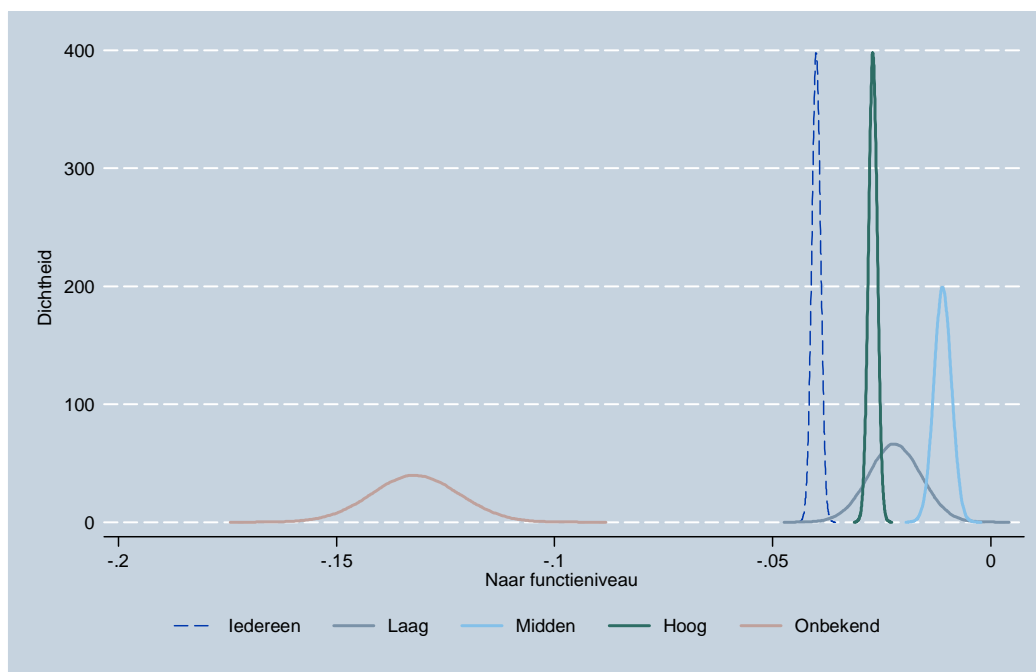
1. Verdelingen naar arbeidsduur:



2. Verdelingen naar Leeftijd:



3. Verdelingen naar functieprofiel:



De verdelingen met betrekking tot de arbeidsduur zijn vrij gelijk aan die van het bedrijfsleven. Doordat er weinig parttimers zijn die minder dan 12 uur werken is deze schatter relatief onbetrouwbaar. De variatie voor parttimers die meer dan 12 uur werken is relatief groter dan die van voltijders wat overeenkomt met de bevindingen voor de bedrijfssector.

De variatie in de geschatte loonverschillen tussen mannen en vrouwen naar de verschillende leeftijdsklassen is het kleinst voor personen tussen de 23 en 35 jaar. Bij de overheid geldt dat de variatie van het geschatte loonverschil het grootst is voor personen jonger dan 23. Dit is wat we op voorhand zouden vermoeden omdat het aantal jongeren in de steekproef relatief klein is en dit zich vertaalt naar een minder betrouwbare schatter.

De verdeling behorende bij het 'functieprofiel' onbekend ligt links van de verdeling voor de hele overheidssector terwijl alle andere verdelingen rechts liggen. Hieruit volgt dat het gemiddelde van het geschatte loonverschil naar functieprofiel voor de overheidssector wordt beïnvloed door de categorie onbekend. De variatie in de verdeling van het functieprofiel hoog lijkt het meest op dat van de hele overheidssector en dit komt omdat relatief de meeste personen in de overheidsector in een hoge functie werken.

3.3 Verbetering modelspecificatie

Om een zuiverdere schatter te verkrijgen van het loonverschil tussen mannen en vrouwen wordt in deze paragraaf de modelspecificatie op twee manieren aangepast. Allereerst worden gebruikte modelvariabelen anders gedefinieerd en worden er aanvullende relevante variabelen in het model opgenomen (paragraaf 3.3.1). Ten tweede worden er verschillende interactietermen en hogere orde variabelen aan de loonvergelijking toegevoegd. Hierbij geven we eerst een leeswijzer die

aangeeft hoe deze complexere termen geïnterpreteerd dienen te worden (paragraaf 3.3.2), om daarna de resultaten te bespreken (paragraaf 3.3.3).

3.3.1 Modeluitbreiding en herdefiniëring variabelen

Loondefinitie

Bij de constructie van het loon in de basisloonschatting is geen gebruik gemaakt van informatie over winstafhankelijke- en niet-winstafhankelijke (jaarlijkse) uitkeringen, terwijl deze uitkeringen wel de hoogte en de variatie van lonen beïnvloeden. Als eerste uitbreiding nemen we daarom deze beide looncomponenten mee in onze loondefinitie. Met niet-winstafhankelijke uitkeringen wordt bedoeld arbeidsmarkttoeslagen en functioneringstoeslagen die doorgaans alleen aan bepaalde categorieën personeelsleden worden uitgekeerd en een 13^e/14^e maand, aldus het arbeidsinspectierapport. Voor het gemak verwijzen we naar de winstafhankelijke uitkering als *WU* en naar de niet-winstafhankelijke uitkering als *NWU*. In Tabel 3.2 worden beschrijvende statistieken weergegeven voor deze beide grootheden.

Het is mogelijk om de omvang van de jaarlijkse uitkeringen te bepalen als percentage van het maandloon (volgens de definitie van de basisloonschatting uit paragraaf 3.1). Op deze manier krijgen we een idee hoe substantieel *NWU* en *WU* zijn. Deze percentages worden als volgt bepaald:

$$NWU_{\% \text{ van maandloon}} = \frac{\frac{100}{12} \cdot NWU}{\text{Bruto maandloon} + \text{som extra loon}}$$

En

$$WU_{\% \text{ van maandloon}} = \frac{\frac{100}{12} \cdot WU}{\text{Bruto maandloon} + \text{som extra loon}}$$

en staan in onderstaande tabel weergegeven.

Tabel 3.2: Beschrijvende Statistieken van *NWU* en *WU* (gewogen)

	% werknemers	% van basisloon	% bij p95	% vrouw	% van basisloon vrouwen
winstafhankelijke uitkering (<i>WU</i>)	9,8	5,7	13,0	33,7	5,4
niet-winstafhankelijke uitkering (<i>NWU</i>)	51,0	4,5	9,8	51,1	4,4

Uit de eerste kolom van de tabel blijkt dat 9.8 procent van de werknemers een winstafhankelijke uitkering ontvangt en 51 procent een niet-winstafhankelijke uitkering ontvangt.¹⁰ Gemiddeld bedragen deze uitkeringen respectievelijk 5.7% en 4.5% van het basisloon (kolom 2). Voor vijf procent van de werknemers is de winstafhankelijke uitkering zelfs groter dan 13% van hun basisloon (kolom 3). Kolom 4 laat zien dat van iedereen die een *WU* geniet slechts eenderde

¹⁰ Bij de overheid is dit bijna 100%, in de marktsector minder dan de helft. Navraag bij de Arbeidsinspectie leert dat in dit bedrag géén vakantiegeld is meegeteld.

vrouw is, terwijl ongeveer evenveel mannen als vrouwen een *NWU* ontvangen. De omvang van de jaarlijkse uitkeringen van vrouwen is bijna gelijk aan die van mannen (kolom 5).

In verhouding tot de loonkloof van 7.4 procent voor het bedrijfsleven en 4 procent voor de overheidssector geldt dat de jaarlijkse *WU* en *NWU* met respectievelijk 5.7 procent en 4.5 procent van het loon aanzienlijk in omvang zijn. Ondanks dat de verschillen in omvang tussen mannen en vrouwen gering zijn, kunnen deze variëren naar achtergrondkenmerken en zodoende een aanvullend deel van het loonverschil verklaren. Het uurloon wordt daarom als volgt gedefinieerd:

$$w_{i,s} = \frac{\text{Bruto maandloon} + \text{som extra loon} + \frac{NWU + WU}{12}}{\text{Aantal arbeidsuren per maand}}$$

In de verdere analyses zal bovenstaande definitie van het uurloon gebruikt worden. Aanvullend worden er als verklarende variabelen twee dummy variabelen opgenomen die aangeven of een persoon een (niet-)winstafhankelijke uitkering ontvangt.

Etniciteit

In de basisanalyses voor het bedrijfsleven we als verklarende variabele de etnische afkomst van een persoon op. Van 1761 personen is echter geen informatie bekend en tot nu toe zijn deze personen beschouwd als autochtoon. Als er in de groep van 1761 personen relatief veel allochtone personen aanwezig zijn dan wordt het effect van etnische afkomst op het loon onzuiver gemeten. Door een extra controlevariabele mee te nemen die aangeeft dat de etnische afkomst onbekend is wordt voor deze onzuiverheid gecorrigeerd.

Type arbeidscontract

Voor het bedrijfsleven onderscheiden we de volgende type arbeidscontracten:

1. Werknemers met een uitzendbaan of een reguliere baan voor bepaalde tijd;
2. Werknemers met een uitzendbaan of een reguliere baan voor onbepaalde tijd;
3. Werknemers met een flexcontract voor onbepaalde tijd;
4. Werknemers met een flexcontract voor bepaalde tijd.

Werknemers met een uitzendbaan kunnen echter niet vergeleken worden met werknemers met een reguliere baan. Daarom maken we in het vervolg onderscheid tussen uitzendkrachten en personen met een reguliere baan zodat we de volgende arbeidscontracten onderscheiden:

1. Werknemers met een reguliere baan voor bepaalde tijd;
2. Werknemers met een reguliere baan voor onbepaalde tijd;
3. Werknemers met een flexcontract voor onbepaalde tijd;
4. Werknemers met een flexcontract voor bepaalde tijd;
5. Werknemers met een uitzendbaan voor onbepaalde tijd;
6. Werknemers met een uitzendbaan voor bepaalde tijd.

Regio

De regio dummy's zijn geconstrueerd aan de hand de provinciecodes die beschikbaar zijn in de data. We maken onderscheid tussen 4 regio's, namelijk: Noord, Oost Zuid en West. Er zijn voor het bedrijfsleven echter 229 observaties waarvoor er geen informatie beschikbaar is. Doordat er geen aparte variabele die aangeeft wanneer er geen regio informatie is geldt dat zowel regio West als de onbekende regio's fungeren als referentiegroep. Het gevolg kan zijn dat we het effect van regio op het loon niet zuiver meten.

In Tabel 3.3 tonen we de schattingsresultaten voor het bedrijfsleven rekening houdende met de hierboven genoemde aanpassingen. Evenals in paragraaf 3.1 nemen we als afhankelijke variabele de logaritme van het bruto uurloon. Omwille van de overzichtelijkheid tonen we alleen de verklarende variabelen waarvan de specificatie is aangepast. We duiden met '*inclusief controle variabelen*' aan dat de ongewijzigde variabelen die eerder gebruikt zijn ook nu weer zijn opgenomen. De complete schattingsresultaten staan vermeld in Bijlage D.

Tabel 3.3 Loonschatting voor het bedrijfsleven – eerste uitbreiding

	Coëfficiënt	Standaard fout	95% B.I	
Vrouw	-0,076	0,003	-0,082	-0,071
Allochtoon	-0,049	0,004	-0,057	-0,040
Etnische afkomst onbekend	-0,016	0,006	-0,028	-0,004
Dummy winstafhankelijke uitkering	0,133	0,003	0,126	0,140
Dummy niet-winstafhankelijke uitkering	0,055	0,002	0,051	0,060
Regulier bepaalde tijd	-0,050	0,004	-0,058	-0,041
Flexibel onbepaalde tijd	-0,053	0,005	-0,063	-0,042
Flexibel bepaalde tijd	-0,073	0,007	-0,087	-0,059
Uitzend onbepaalde tijd	-0,460	0,022	-0,504	-0,416
Uitzend bepaalde tijd	-0,357	0,031	-0,418	-0,297
Regio Noord	-0,049	0,004	-0,058	-0,041
Regio Oost	0,003	0,003	-0,002	0,008
Regio Zuid	-0,014	0,003	-0,020	-0,009
Regio onbekend	-0,017	0,016	-0,048	0,015
Inclusief controle variabelen				
Observaties	52208			
Aangepaste R ²	0,7955			

Alle toegevoegde verklarende variabelen zijn significant met uitzondering van de dummy 'Regio onbekend'. Het opnemen van deze regio dummy is dus wel 'mooier' maar heeft geen significante invloed op de schattingsresultaten. Het gemiddelde loon voor personen wiens etnische afkomst onbekend is, is gemiddeld lager dan dat van autochtonen maar hoger dan dat voor allochtonen. Dit is conform de verwachting omdat deze groep bestaat uit zowel autochtone als allochtone personen.

Uitzendkrachten verschillen duidelijk van personen die een reguliere baan hebben: ze verdienen gemiddeld veel minder. Dit bevestigt het vermoeden dat uitzendkrachten in de basisanalyse onterecht gerekend werden tot de personen met een reguliere baan. Dat de dummyvariabelen

met betrekking tot *NWU* en *WU* significant zijn is triviaal aangezien het loon beide uitkeringen bevat.

Het loonverschil tussen mannen en vrouwen is in de nieuwe specificatie met 7.6 procent iets groter dan het eerdere gemeten loonverschil van 7.4 procent in Tabel 3.1. Op basis van de significantie van de toegevoegde verklarende variabelen concluderen we dat deze variabelen bijdragen aan een betere verklaring van de loonvariatie in het bedrijfsleven. Toch verandert de het geschatte loonverschil m-v nauwelijks.

In Tabel 3.4 tonen we de schattingsresultaten voor de overheidssector rekening houdende met de eerder genoemde veranderingen. Evenals voor de bedrijfssector tonen we alleen de verklarende variabele vrouw en de variabelen die door de eerder genoemde aanpassingen zijn veranderd. De complete schattingsresultaten staan vermeld in Bijlage D.

Tabel 3.4 Loonschatting voor de overheidssector – eerste uitbreiding

	Coëfficiënt	Standaard fout	95% B.I	
Vrouw	-0,038	0,001	-0,041	-0,036
Dummy niet-winstafhankelijke uitkering	0,103	0,048	0,009	0,197
Inclusief controle variabelen				
Observaties	65381			
Aangepaste R ²	0,8078			

Voor de overheidssector is er geen informatie beschikbaar over etniciteit, regio en type arbeidscontract. Omdat er bovendien geen winstafhankelijke uitkeringen plaatsvinden, wordt voor de overheidssector alleen *NWU* opgenomen in het loon en wordt er als verklarende variabele een dummy opgenomen in de analyse die aangeeft of een persoon *NWU* ontvangt. Dat deze *NWU* dummy op een significante manier de variatie in uurloon beïnvloedt is vanzelfsprekend omdat het ontvangen van een uitkering een direct positief effect heeft op het totale uurloon.

In vergelijking tot het geschatte loonverschil tussen mannen en vrouwen in Tabel 3.1 van 4.0 procent daalt het loonverschil nu met 0.2 procent naar 3.8 procent. Wederom concluderen we dat de modelaanpassingen wel bijdragen aan een betere verklaring van het loon, maar dat het eerdere geschatte loonverschil tussen mannen en vrouwen nauwelijks wordt beïnvloed door de toevoeging van de variabelen.

3.3.2 Niet-lineaire specificaties en interactietermen: technische leeswijzer

Niet-lineaire specificaties

Tot nu toe zijn de verklarende variabelen die opgenomen zijn in het model discrete (categoriale) variabelen. Zo worden er op voorhand verschillende leeftijdsklassen gedefinieerd en worden niet de bedragen van de winstafhankelijke uitkering opgenomen als verklarende variabele, maar slechts of personen een winstafhankelijke uitkering ontvangen. In deze paragraaf worden bepaalde categoriale variabelen als continue variabelen opgenomen in de loonvergelijking omdat

we vermoeden dat dit bijdraagt in het verklaren van de variatie in het uurloon en ook mogelijksterwijs het loonverschil tussen mannen en vrouwen deels kan verklaren.

Ter illustratie nemen we bijvoorbeeld de variabele 'leeftijd'. Naarmate personen ouder worden zullen zij meer verdienen, maar deze toename is degressief. Met andere woorden: de loonsprongen die personen op jonge leeftijd maken, zijn groter in vergelijking tot de loonsprongen die zij maken wanneer zij ouder zijn. Wanneer personen tussen de 35 en 44 jaar beschouwd worden als één categorie, gaan we voorbij aan de loonstijging die zal plaatsvinden *tussen* het 35^{ste} en het 44^{ste} levensjaar. In plaats van arbitraire categorieën te kiezen kan er ook voor gekozen worden om de logaritme van leeftijd te gebruiken als verklarende variabele in de loonvergelijking. Hierdoor implementeren we in feite dat personen afnemend meer loon verdienen naarmate zij ouder worden, waardoor de verklarende kracht van het model mogelijk toeneemt. Een andere veelgebruikte mogelijkheid is om leeftijd als continue lineaire variabele op te nemen samen met leeftijd in het kwadraat. Hierdoor laten we in feite de mogelijkheid toe dat het effect van leeftijd op het loon kwadratisch is. We merken overigens op dat in principe het opnemen van een dummyvariabele voor elk levensjaar de voorkeur zou genieten boven een logaritmische of een kwadratische specificatie, mits er voor elk levensjaar voldoende waarnemingen zijn. Zodoende wordt niet op voorhand een functionele vorm wordt opgelegd en zijn de schattingsresultaten daardoor minder afhankelijk van de keuzes van de onderzoeker. In de praktijk is deze oplossing echter zelden mogelijk.

Bij het schatten van de loonvergelijking specificeren we de relatie tussen het loon en de volgende verklarende variabelen: leeftijd, arbeidsduur, het extra loon dat wordt ontvangen en de jaarlijkse uitkeringen. In dit onderzoek wordt gekozen voor een logaritmische specificatie van de loongerelateerde variabelen en een kwadratische specificatie van leeftijd en dienstjaren.

We kunnen bijvoorbeeld de logaritme nemen van het extra loon¹¹ en het effect van het extra loon op het uurloon wordt dan gemeten in procentpunten. Als het geschatte effect van het extra verdiende loon op uurloon 0.034 is, dan betekent dit dat een stijging van het extra loon met 1 procent leidt tot een stijging van het uurloon van ruim 3.4 procent. Voor leeftijd is een logaritmische specificatie minder intuïtief omdat we dan het effect zouden meten op het uurloon als een persoon 1 procent ouder wordt. De logaritmische specificatie is daarom intuïtief voor het extra loon dat wordt ontvangen en de al dan niet ontvangen (niet-)winstafhankelijke uitkeringen terwijl een kwadratische term beter past bij leeftijd en arbeidsduur.

Interactietermen

Het loonverschil tussen mannen en vrouwen is tot nu toe gemeten met een dummyvariabele voor geslacht. Daarmee wordt geïmpliceerd dat het loonverschil tussen mannen en vrouwen gelijk is voor alle mogelijke kenmerken, en bovendien dat *andere verklarende kenmerken hetzelfde effect hebben op loon voor mannen als voor vrouwen*. Deze laatste veronderstelling lijkt niet realistisch, bijvoorbeeld doordat vrouwen en mannen zichzelf anders kunnen selecteren in banen, andere voorkeuren kunnen hebben en andere keuzes maken met betrekking tot het aantal uren dat zij werken. In de analyse moet daarom rekening gehouden worden met verschillende effecten van bepaalde kenmerken voor mannen en voor vrouwen.

¹¹ We merken op dat we in werkelijkheid de logaritme van het extraloon *plus één* zullen opnemen. Dit gebeurt omdat we niet de logaritme van 0 kunnen nemen Deze toevoeging heeft verder geen consequenties voor de zuiverheid van de verkregen schatters.

In dit onderzoek nemen we aan dat het loonverschil tussen mannen en vrouwen kan variëren naar leeftijd, arbeidsduur, etnische afkomst en de sector waarin een persoon werkzaam is. Om deze variatie toe te laten schatten we een loonvergelijking waarbij we interactietermen toevoegen. De werking van deze interactietermen zullen we uitleggen aan de hand van een sterk gesimplificeerd voorbeeld, waarbij we veronderstellen dat uurloon perfect verklaard kan worden door het opnemen van de variabelen geslacht en etniciteit. De loonvergelijking wordt dan gerepresenteerd door:

$$\ln w = \beta_0 + \beta_1 \text{Vrouw} + \beta_2 \text{Etniciteit} + \beta_3 \text{Vrouw} \cdot \text{Etniciteit}$$

De laatste term in de vergelijking is de interactieterm en deze geeft aan hoe de onderlinge relatie tussen etnische afkomst en geslacht het uurloon beïnvloedt. In Tabel 3.5. geven we weer hoe het uurloon varieert op basis van de mogelijke kenmerkcombinaties. In Tabel 3.5 geven de eerste 2 kolommen de combinatie van kenmerken weer die een persoon kan bezitten. De vierde kolom geeft weer hoe de verschillende kenmerken de variatie in uurloon verklaren en kolom 3 laat zien hoe deze effecten berekend zijn.

Tabel 3.5 Effect van kenmerken op uurloon

Kenmerk 1	Kenmerk 2	Berekening	Uurloon
Man	Autochtoon	$\beta_0 + \beta_1 \cdot 0 + \beta_2 \cdot 0 + \beta_3 \cdot 0$	β_0
Man	Allochtoon	$\beta_0 + \beta_1 \cdot 0 + \beta_2 \cdot 1 + \beta_3 \cdot 0$	$\beta_0 + \beta_2$
Vrouw	Autochtoon	$\beta_0 + \beta_1 \cdot 1 + \beta_2 \cdot 0 + \beta_3 \cdot 0$	$\beta_0 + \beta_1$
Vrouw	Allochtoon	$\beta_0 + \beta_1 \cdot 1 + \beta_2 \cdot 1 + \beta_3 \cdot 1$	$\beta_0 + \beta_1 + \beta_2 + \beta_3$

De eenvoudigste manier Tabel 3.5 te interpreteren is door te onderkennen dat het loon van een autochtone man weergegeven wordt door β_0 . Het loonverschil tussen een autochtone man en een allochtone man is β_2 en deze wordt bepaald door het verschil te nemen tussen $\beta_0 + \beta_2$ en β_0 . Op dezelfde manier kan het verschil bepaald worden tussen een autochtone man en een autochtone vrouw en deze is gelijk aan β_1 .

In de nieuwe modelspecificatie is niet langer alleen β_1 van belang, maar ook β_3 ; de ene term kan niet meer los van de andere geïnterpreteerd worden. Door deze nieuwe term meten we het loonverschil tussen mannen en vrouwen naar etnische afkomst. Het verschil tussen een autochtone vrouw en een allochtone man is $(\beta_0 + \beta_1) - (\beta_0 + \beta_2) = \beta_1 - \beta_2$. Het verschil tussen een allochtone vrouw en een allochtone man is $(\beta_0 + \beta_1 + \beta_2 + \beta_3) - (\beta_0 + \beta_2) = \beta_1 + \beta_3$.

3.3.3 Niet-lineaire specificaties en interactietermen: resultaten

In Tabel 3.6 tonen we de schattingsresultaten voor het bedrijfsleven, waarbij we wederom alleen de uitkomsten tonen voor de variabelen die aan verandering onderhevig zijn geweest ten opzichte van de vorige paragraaf. De kwadratische termen voor dienstjaren en leeftijd zijn weliswaar significant maar het effect is zeer klein. Gemiddeld genomen verdienen allochtonen een lager uurloon dan autochtonen maar op basis van de interactieterm concluderen we dat dit

verschil in uurloon kleiner wordt als het een vrouwelijke allochtoon betreft. De effecten van extra looncomponenten, *NWU* en *WU* zijn zoals verwacht aanzienlijk maar in vergelijking tot Tabel 3.3 en Tabel 3.4 zijn de effecten gematigder. De reden hiervoor is dat een logaritmische specificatie in vergelijking tot het opnemen van een dummy beter corrigeert voor mogelijke extreme waarden die voorkomen.

Tabel 3.6 Loonschatting voor het Bedrijfsleven met toevoeging van interactietermen en niet lineaire verklarende variabelen⁴

	Coëfficiënt	Standaardfout	95% B.I	
Vrouw	-0,127	0,028	-0,182	-0,073
Leeftijd	0,067	0,001	0,066	0,069
Leeftijd ²	-0,001	0,000	-0,001	-0,001
Dienstjaren	0,000	0,000	-0,001	0,001
Dienstjaren ²	0,000	0,000	0,000	0,000
Allochtoon	-0,051	0,001	-0,062	-0,040
Sector (bouw = referentie)				
<i>landbouw en visserij</i>	-0,062	0,023	-0,080	-0,044
<i>industrie</i>	-0,082	0,018	-0,093	-0,071
<i>handel</i>	-0,103	0,018	-0,115	-0,091
<i>horeca</i>	-0,093	0,020	-0,110	-0,075
<i>transport / communicatie</i>	-0,084	0,019	-0,098	-0,070
<i>zakelijke dienstverlening</i>	-0,089	0,018	-0,101	-0,076
<i>zorg en welzijn</i>	-0,070	0,020	-0,092	-0,048
<i>overig</i>	-0,092	0,019	-0,108	-0,076
Log(extra loon+1)	0,028	0,001	0,027	0,029
Log(<i>WU</i>)	0,034	0,001	0,033	0,036
Log(<i>NWU</i>)	0,020	0,001	0,019	0,021
Vrouw geïnteracteerd met:				
Leeftijd	-0,008	0,001	0,002	0,007
Leeftijd ²	0,000	0,000	0,000	0,000
Dienstjaren	0,004	0,001	0,002	0,006
Dienstjaren ²	-0,000	0,000	0,000	0,000
Allochtoon	0,029	0,009	0,011	0,047
Sector (bouw = referentie)				
<i>landbouw en visserij</i>	n.s.	0,023	-0,039	0,050
<i>industrie</i>	0,042	0,018	0,007	0,076
<i>handel</i>	n.s.			
<i>horeca</i>	0,042	0,020	0,003	0,081
<i>transport / communicatie</i>	n.s.			
<i>zakelijke dienstverlening</i>	n.s.			
<i>zorg en welzijn</i>	0,099	0,020	0,060	0,138
<i>overig</i>	0,073	0,019	0,036	0,110
Inclusief controle variabelen				
Observaties	52208			
Aangepaste R ²	0,7895			

Het loonverschil tussen mannen en vrouwen is door de toevoeging van de interactietermen niet langer gelijk aan de parameter behorende bij de variabele vrouw. Tevens hangt het verschil af van de gemiddelde leeftijd en ervaring (en etnische afkomst) in een subsector. De loonkloof m-v kan

alleen geanalyseerd worden in combinatie met de andere interactievariabelen of voor het gemiddelde hiervan. De sectoreffecten zijn relatief ten opzichte van de bouwsector en in interactie met geslacht, ervaring en etnische afkomst waardoor geen recht-toe-recht-aan interpretatie gegeven kan worden. Gemiddeld genomen is het uurloon voor mannen in de bouwsector hoger dan dat in andere sectoren. Op basis van de bijbehorende interactieterm blijkt dat verschillen in uurloon voor vrouwen in enkele sectoren lager uitvalt dan in de bouw.

Om een duidelijk te maken wat deze resultaten in de praktijk betekenen, laat Tabel 3.7 de loonverschillen zien voor de gemiddelde werknemer in elke subsector van het bedrijfsleven.¹² Een negatieve waarde duidt op een loonachterstand voor vrouwen. Het gemiddelde loonverschil tussen mannen en vrouwen daalt van 7.4 procent in de basisloonschatting en 7.6 procent in de aangepaste specificatie uit de vorige paragraaf naar 6.6 procent in bovenstaande uitgebreide modelspecificatie.¹³ We concluderen dat de gemaakte modelaanpassingen zorgen voor een lagere waarde van het resulterende loonverschil tussen mannen en vrouwen in het bedrijfsleven.

Splitsen we de loonverschillen uit naar subsector, dan zien we daar aanzienlijke verschillen. Er komt duidelijk naar voren dat in sectoren waar veel vrouwen werken, zoals de zorg en overige dienstverlening, seksegerelateerde loonverschillen veel kleiner of zelfs geheel afwezig zijn.¹⁴

Tabel 3.7 Loonkloof naar sector, bedrijfsleven

Landbouw en visserij	-8.4%
Industrie	-5.0%
Bouw	-9.3%
Handel	-10.4%
Horeca	-4.3%
Transport	-10.6%
Zakelijke dienstverlening	-6.8%
Gezondheids & welzijnszorg	0.1%
Overige dienstverlening	-2.5%
Totaal	-6.6%

In Tabel 3.8 tonen we de schattingsresultaten voor de overheidssector. Evenals in het bedrijfsleven zijn de kwadratische termen wel significant maar verwaarloosbaar in grootte. Personen verdienen zodoende een hoger uurloon naarmate zij ouder worden en meer ervaring krijgen.

¹² Hierbij wordt voor elke subsector uitgegaan van de gemiddelde leeftijd, ervaring en etniciteit in die sector. Dit sluit het best aan bij de separate analyses per subsector zoals die zijn uitgevoerd door de Arbeidsinspectie.

¹³ Dit verschil is bepaald voor de gemiddelde respondent in de steekproef, werkzaam in het bedrijfsleven.

¹⁴ Deze verschillen zijn bepaald voor de gemiddelde respondent in de steekproef, werkzaam in de subsector.

Tabel 3.8 Loonschatting voor de overheidssector met toevoeging van interactietermen en niet lineaire verklarende variabelen

	Coëfficiënt	Standaardfout	95% B.I	
Vrouw	-0,054	0,018	-0,090	-0,019
Leeftijd	0,026	0,001	0,025	0,028
Leeftijd ²	0,000	0,000	0,000	0,000
Dienstjaren	0,003	0,000	0,003	0,004
Dienstjaren ²	0,000	0,000	0,000	0,000
Sector (Rijk = referentie)				
Onderwijs	-0,060	0,002	-0,065	-0,055
Rechterlijke Macht	0,022	0,006	0,010	0,033
Defensie	-0,037	0,002	-0,042	-0,033
Log(extra loon+1)	0,018	0,000	0,017	0,018
Log(<i>NWU</i>)	0,080	0,001	0,078	0,082
Vrouw geïnteracteed met:				
Leeftijd	0,003	0,001	0,002	0,005
Leeftijd ²	-0,000	0,000	0,000	0,000
Dienstjaren	0,001	0,000	0,000	0,002
Dienstjaren ²	n.s.			
Sector (Rijk = referentie)				
<i>Onderwijs</i>	-0,034	0,003	-0,039	-0,029
<i>Rechterlijke Macht</i>	-0,020	0,007	-0,034	-0,006
<i>Defensie</i>	0,033	0,004	0,025	0,041
Inclusief controle variabelen				
Observaties	65381			
Aangepaste R ²	0,8311			

Zoals verwacht zijn de extra looncomponenten en *NWU* significant en gematigder dan in Tabel 3.4. Om een betere indruk te krijgen van wat deze resultaten in de praktijk betekenen, laat Tabel 3. de loonverschillen zien voor de gemiddelde werknemer in elke subsector van de overheid. Een negatieve waarde duidt op een loonachterstand voor vrouwen.

Tabel 3.9 Loonkloof naar sector, overheid

Rijk	-1.4%
Onderwijs	-5.8%
Rechterlijke Macht	-4.2%
Defensie Burg. Personeel	2.3%
Totaal	-2.7%

De gemiddelde loonkloof daalt van 4.0 procent in de basisloonschatting en 3.8 procent in Tabel 3.4 naar gemiddeld 2.7 procent in de huidige, uitgebreide modelspecificatie.¹⁵ We concluderen dus dat ook voor de overheid de gemaakte modelaanpassingen zorgen voor een lagere waarde van het resulterende loonverschil tussen mannen en vrouwen. De loonverschillen tussen mannen

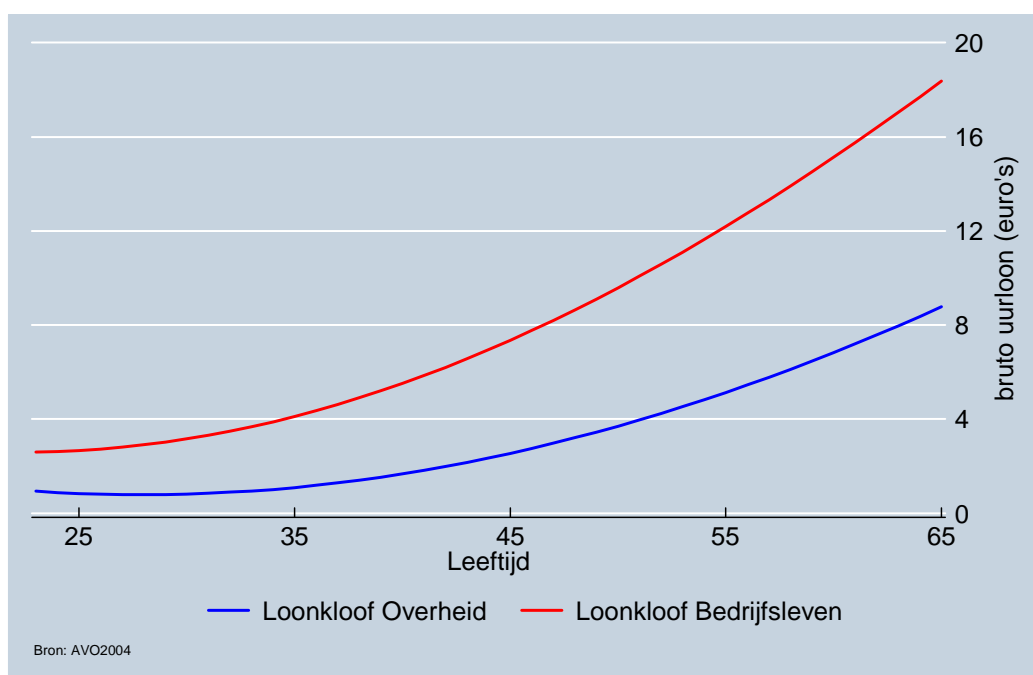
¹⁵ Dit verschil is bepaald voor de gemiddelde respondent in de steekproef, werkzaam bij de overheid.

en vrouwen zijn het grootst in het onderwijs, terwijl er bij het burgerpersoneel van defensie sprake is van een klein beloningsvoordeel voor vrouwen. Verdere uitsplitsing binnen de omvangrijke subsector onderwijs was uiteraard wenselijk geweest: het betreft namelijk een heterogene groep die in totaal 6 verschillende CAO's omvat, variërend van primair- tot universitair onderwijs. Helaas kan dit relevante onderscheid in de AVO-data niet gemaakt worden.

Puur leeftijdseffect en beperkingen aan de traditionele schattingsmethode

Tot slot blijkt dat wanneer de gemiddelde leeftijd relatief hoog is, de loonachterstand van vrouwen ook relatief hoog is. Dit resultaat vinden we zowel bij de overheid als bij het bedrijfsleven. Nu is dit deels te verklaren uit het feit dat oudere vrouwen gemiddeld ook een lager aantal dienstjaren hebben dan mannen, maar zowel Tabel 3.6 als Tabel 3.8 laten zien dat er meer speelt. Ook onafhankelijk van het aantal dienstjaren speelt het pure leeftijdseffect mee: de interactietermen van $\text{vrouw} \cdot \text{leeftijd}^2$ zijn immers beide negatief en significant. Hieronder is dit grafisch weergegeven voor twee referentiepersonen, de gemiddelde werknemer van de overheid en de gemiddelde werknemer van het bedrijfsleven.

Figuur 3.5 Loonverschillen naar leeftijd, referentiepersonen overheid & bedrijfsleven



Een duidelijk nadeel van de traditionele multivariate schattingsmethodes (in dit geval OLS op een logaritmische specificatie) komt uit de grafiek naar voren: de beperking van de resultaten door de gekozen functionele vorm. In de basisspecificatie zoals ook in het Arbeidsinspectie-onderzoek gehanteerd was dat een loglineaire relatie, in de meest uitgebreide specificatie in deze paragraaf is het een loglineaire specificatie met interactie- kwadratische termen, maar het fundamentele probleem blijft hoe dan ook bestaan.

Figuur 3.5 illustreert dit: vanwege de kwadratische relatie tussen leeftijd en loon, rapporteert de schattingsmethode dat het loonverschil m-v stijgt met de leeftijd. Op je 26^e is het loonverschil bij de overheid bijna nihil, op je 50^e is het bijna 4 procent. Trek je deze relatie door, dan zou het loonverschil op je 64^e zelfs bij de overheid 8% moeten zijn. In de praktijk is dat waarschijnlijk niet het geval, maar omdat er maar weinig vrouwen van 64 zijn die nog werken, is het lastig om in deze extreme gevallen nog een goede schatting te maken. De relatie tussen leeftijd en loon wordt eigenlijk vooral bepaald door de grote groep vrouwen tussen 23 jaar en 50 jaar oud. Daarbuiten hangt de relatie sterk af van de functionele vorm van de loonvergelijking. Theoretisch is het mogelijk om in een dergelijke analyse bijvoorbeeld jonge vrouwen met oude mannen te vergelijken, ook als ze eigenlijk niet goed vergelijkbaar zijn. Om dit nadeel te overwinnen nemen we in het volgende hoofdstuk onze toevlucht tot een modernere schattingsmethode: matching.

3.4 Conclusie uitbreiding modelspecificatie

De onzekerheidsmarge van het door Arbeidsinspectie gevonden loonverschil tussen mannen en vrouwen is groter voor het bedrijfsleven dan voor de overheid. Het loonverschil tussen mannen en vrouwen wordt in de marktsector dus minder nauwkeurig gemeten. Deze grotere variatie komt waarschijnlijk doordat de lonen in de marktsector sowieso veel minder uniform zijn dan in de overheidssector, waar vaker met vaste functiewaarderingsystemen gewerkt wordt. Bovendien bevat het bedrijfsleven in deze definitie ook de zorgsector, een sector die relatief weinig karakteristieken gemeen heeft met de andere subsectoren.

Overigens bevat de hier gebruikte AVO-dataset voor de overheidssector veel minder geschikte variabelen dan voor de marktsector, en zijn zelfs niet alle subsectoren van de overheid vertegenwoordigd: gemeenten, provincies, waterschappen, militairen en politie ontbreken. Bij de interpretatie van de resultaten is het belangrijk om dit voor ogen te houden. Alle hierna vermelde resultaten kunnen niet zonder meer vergeleken worden *tussen* beide sectoren, en waar we spreken van de ‘overheid’ bedoelen we in feite alleen de in het onderzoek vertegenwoordigde sectoren.

Uit de basisreplicatie voor het bedrijfsleven resulteert het reeds bekende gecorrigeerde loonverschil van 7,4%. Vrouwen in het bedrijfsleven ontvangen relatief minder vaak dan mannen een winstuitkering, maar als ze die krijgen is hij wel net zo groot. Naast deze aanvullende informatie dragen ook de meeste andere toegevoegde variabelen bij in de nauwkeurigere ontleding van loonverschillen in het bedrijfsleven. Desondanks verandert het onverklaarde loonverschil door deze uitbreiding echter nauwelijks in grootte.

Het gemiddelde loon voor personen wiens etnische afkomst onbekend is, is gemiddeld lager dan dat van autochtonen maar hoger dan dat voor allochtonen. Dit is conform de verwachting omdat deze groep bestaat uit zowel autochtone als allochtone personen. Uitzendkrachten verschillen duidelijk van personen met een reguliere baan: ze verdienen gemiddeld veel minder. Dit bevestigt het vermoeden dat uitzendkrachten in de basisanalyse onterecht gerekend werden tot de personen met een reguliere baan.

Door de toevoeging van interactie- en hogere ordetermen wint de analyse van loonverschillen nog meer aan dimensie: in subsectoren bouw, handel en transport (typische mannensectoren) blijkt het beloningsverschil tussen mannen en vrouwen groter te zijn dan in de zorgsector (een typische vrouwensector). Allochtonen ontvangen een lager uurloon dan autochtonen maar op basis van de interactieterm concluderen we dat dit verschil onder vrouwen kleiner is. Gemiddeld resulteert in de meest uitgebreide analyse een loonverschil van 6,6%, minder dus dan in de basisreplicatie, maar nog steeds substantieel.

Tabel 3.10 Overzicht loonverschillen, AVO-data, OLS.

	<i>Bedrijfsleven</i>
Basisloonschatting / replicatie	7,4
+ uitbreiding nieuwe variabelen	7,6
+ interactie- & hogere orde effecten	6,6
	<i>Overheid</i>
Basisloonschatting / replicatie	4,0
+ uitbreiding nieuwe variabelen	3,8
+ interactie- & hogere orde effecten	2,7

Bij de overheidssector geeft de basisreplicatie het loonverschil van 4,0%. Niemand ontvangt een winstgerelateerde uitkering en nagenoeg iedereen een niet-winstgerelateerde uitkering (zoals bijvoorbeeld een eindejaarsuitkering). Na de uitbreiding met extra variabelen en toevoeging van interactie- en hogere ordetermen neemt deze grootte af naar gemiddeld 2,7%. We concluderen daarom dat ook voor de overheid de gemaakte modelaanpassingen zorgen voor een lagere waarde van het resulterende loonverschil tussen mannen en vrouwen. Ook hier zijn er echter belangrijke verschillen tussen de subsectoren, met het grootste loonverschil in de subsector Onderwijs. Verdere uitsplitsing binnen de omvangrijke subsector onderwijs was uiteraard wenselijk geweest: het betreft namelijk een heterogene groep die in totaal 6 verschillende CAO's omvat, variërend van primair- tot universitair onderwijs. Helaas kan dit relevante onderscheid in de AVO-data niet gemaakt worden.

Een duidelijk nadeel van de traditionele multivariate schattingsmethodes (zoals OLS) is de beperking van resultaten door de gekozen functionele vorm. Theoretisch is het mogelijk om in een dergelijke analyse bijvoorbeeld jonge vrouwen met oude mannen te vergelijken, ook als ze eigenlijk niet goed vergelijkbaar zijn. Om dit nadeel te overwinnen nemen we in het volgende hoofdstuk onze toevlucht tot een modernere schattingsmethode: matching.

4 Alternatieve schattingsmethode: ‘matching’

Onderzoek naar loonverschillen tussen mannen en vrouwen gebeurt in de regel met behulp van een zogenaamde ‘Mincer-vergelijking’, een multivariate regressie van loonbepalende factoren op de logaritme van het uurloon. In hoofdstuk 3 en in diverse andere studies wordt met deze methode wordt de variatie in het loon toegeschreven aan de variatie in allerlei kenmerken. Personen worden hier nooit direct met elkaar vergeleken. In theorie zou met behulp van de alternatieve matching-methode nauwkeuriger loonvergelijkingen uitgevoerd moeten kunnen worden. In dit hoofdstuk worden de resultaten beschreven van een dergelijke analyse, toegepast op de AVO-data. In paragraaf 4.1 bespreken we de kenmerken van een dergelijke analyse in algemene zin: hoe gaat matching in zijn werk, welke verschillende methoden zijn hiervoor beschikbaar en waar moet je op letten? Deze paragraaf dient om de lezer te informeren over deze relatief onbekende onderzoeksmethode. Lezers die slechts geïnteresseerd zijn in de uiteindelijke resultaten van de analyse naar loonverschillen op basis van de AVO-data, kunnen direct doorgaan naar de paragrafen 4.2 en 4.3.

4.1 Kenmerken van matching

Om te bepalen of er een oorzakelijk verband bestaat tussen het geslacht en het loonniveau zouden we het loon van vrouwen moeten vergelijken met het loon dat zij ‘*zouden verdienen waren zij mannen geweest*’. Het loon dat een vrouw zou verdienen als zij een man was geweest nemen we echter niet waar; hier wordt in de literatuur naar verwezen als de *counterfactual*.¹⁶ De achterliggende gedachte van matching is dat het oorzakelijke verband tussen geslacht en het loonniveau gemeten kan worden in twee stappen. In de eerste stap wordt er voor de groep vrouwen in de steekproef een controlegroep gesimuleerd. Dit gebeurt door mannen aan vrouwen te koppelen op basis van kenmerken, anders dan geslacht, die het loon beïnvloeden. De controlegroep bestaat dus uit mannen die, met uitzondering van het geslacht, identiek zijn aan de vrouwen in de steekproef. Naar de groep vrouwen in de analyse zullen we verwijzen als de testgroep, omdat we voor deze groep willen testen of er een effect is. In een tweede stap kan het loonniveau van deze groep vrouwen worden vergeleken met dat van de groep mannen in de controlegroep.

In de literatuur wordt een controlegroep beschouwd als ‘goed’, als er aan de volgende voorwaarden voldaan wordt:

- *Volledigheid van de koppelkenmerken*: Bij het koppelen van mannen aan vrouwen moeten, met uitzondering van het geslacht, alle kenmerken die het loonniveau kunnen beïnvloeden worden opgenomen.¹⁷

¹⁶ Dit probleem werd voor het eerst op deze manier gedefinieerd door Roy (1951), Rubin (1974), Holland (1986) and (Splawa)-Neyman (1990) en zij werkten ook de eerste oplossingsmethoden uit.

¹⁷ Deze voorwaarde staat in de literatuur bekend als de ‘unconfoundedness’ voorwaarde (Imbens, 2005) of de ‘ignorability’ voorwaarde (Rubin, 1978; Wooldridge, 2001).

- *Exogeniteit van geslacht*: Als personen gekoppeld zijn op basis van relevante kenmerken moet het verschil in samenstelling tussen de testgroep en de controlegroep onafhankelijk zijn van geslacht.¹⁸
- *Gelijke basis van kenmerken*: De waargenomen kenmerken in de testgroep moeten vergelijkbaar zijn met die in de controlegroep. Zodoende wordt er bewerkstelligd dat de controlegroep op alle dimensies een goede vergelijkingsgroep is voor de vrouwen.¹⁹

Bovenstaande voorwaarden zullen we nader toelichten. Het is niet mogelijk om na te gaan of er aan de voorwaarde *volledigheid van de koppelkenmerken* voldaan wordt. Dit komt omdat het meestal onduidelijk is wat precies de relevante factoren zijn. Bovendien kan het gebeuren dat een factor als relevant wordt beschouwd maar dat er geen informatie beschikbaar is. Het niet opnemen van relevante kenmerken zal ten koste gaan van de vergelijkbaarheid van de controlegroep met de testgroep; het effect van geslacht op het loon wordt hierdoor minder zuiver gemeten.²⁰

Aan de voorwaarde dat geslacht exogeen is zal in vrijwel alle gevallen voldaan zijn: de hoogte van het loon bepaald immers niet het geslacht. Deze voorwaarde meer van belang in een situatie waarin men bijvoorbeeld het effect van een reïntegratietraject op de werkloosheidsduur zou willen meten. Wanneer personen gemotiveerder zijn, zullen zij eerder deelnemen aan een reïntegratietraject en zodoende een hoger loon hebben. Er geldt dan dat personen zichzelf selecteren in een traject waardoor het traject niet meer als exogene variabele gekenmerkt kan worden.

De voorwaarde *gelijke basis van kenmerken* zullen we toelichten aan de hand van een voorbeeld. In dit voorbeeld worden mannen en vrouwen puur aan elkaar gekoppeld op het kenmerk opleidingsniveau, gemeten in drie categorieën: laag, midden en hoog. Als alledrie de opleidingscategorieën voorkomen voor de personen in de testgroep dan moeten al deze opleidingscategorieën ook voorkomen voor personen in de controlegroep. Het testen voor *gelijke basis van kenmerken* wordt lastiger naarmate het aantal koppelingskenmerken (en dus het aantal dimensies) toeneemt of als er continue variabelen gebruikt worden als koppelingsvariabelen. We zullen in dit rapport verder niet ingaan op de technische details maar merken op dat we alleen vrouwen in de testgroep zullen gebruiken als er bij de koppeling voldaan wordt aan de voorwaarde *gelijke basis van kenmerken*.

4.1.1 Methoden van matchen

In deze paragraaf bespreken we de verschillende manieren waarop personen aan elkaar gekoppeld kunnen worden. Allereerst zijn er twee afstandsmaten beschikbaar aan de hand waarmee we, op basis van relevante kenmerken, kunnen bepalen hoever personen van elkaar af liggen. Vervolgens worden er verschillende koppelingsregels geformuleerd waarmee mannen gekoppeld kunnen

¹⁸ Ofwel, de conditionele kans om een man of een vrouw te zijn mag niet gemiddeld genomen niet verschillen tussen de test- en de controlegroep. Deze voorwaarde staat in de literatuur bekend als de ‘conditional mean independence assumption’.

¹⁹ In de literatuur staat deze voorwaarde bekend als de ‘common support’ aanname. Het idee is dat wanneer zowel uit de proefgroep als de controlegroep een persoon wordt gekozen op basis van de koppelingskenmerken, beide personen getrokken moeten worden uit dezelfde kansverdeling.

²⁰ Eenzelfde probleem doet zich ook voor bij het schatten van een Mincer-vergelijking als niet alle relevante kenmerken in de schatting worden opgenomen. In de literatuur wordt hiernaar verwezen als de ‘Omitted Variable Bias’.

worden aan vrouwen met behulp van een van de twee afstandsmaten. We zullen in deze paragraaf verwijzen naar ‘de relevante kenmerken’; hiermee bedoelen we alle kenmerken (anders dan geslacht) die het loonverschil tussen mannen en vrouwen mede verklaren.

In de AVO dataset is het niet mogelijk om koppelingen zodanig te laten plaatsvinden dat er voor elke vrouw een man wordt gevonden die exact dezelfde relevante kenmerken heeft. Daarom wordt er op basis van relevante kenmerken twee afstandsmaten gedefinieerd worden, namelijk:

1. de propensity score afstand; en
2. de Mahalanobis afstand.

Rosenbaum en Rubin (Biometrika, 1983) hebben aangetoond dat het koppelen van mannen aan vrouwen kan plaatsvinden op basis van de kans dat een persoon een vrouw is. Deze kans is voor alle personen te bepalen door middel van een eenvoudige probit-schatting op basis van de relevante kenmerken. Naar deze kans wordt verwezen als de ‘propensity score’. Een voordeel van de propensity score als afstandsmaat is dat een vrouw uiteindelijk gekoppeld wordt aan een man op basis van een ééndimensionale kansvariabele, onafhankelijk van het aantal relevante kenmerken. Dit maakt het bijvoorbeeld eenvoudig om voor het koppelen te testen of voldaan wordt aan de eerder genoemde voorwaarde *gelijke basis van kenmerken*. Een nadeel van deze propensity score afstandsmaat is echter dat de voorspelde kans afhangt van de gekozen functionele vorm bij het schatten van een probit model; daarmee wordt de analyse niet-parametrisch. Deze nadelen zullen het grootst zijn als het aantal relevante kenmerken beperkt is.

Een alternatieve afstand om de koppeling te laten plaatsvinden is de ‘Mahalanobis afstand’, waarbij geldt dat de afstand tussen twee personen direct wordt bepaald op basis van de relevante kenmerken.²¹ Vergeleken met de propensity score als afstandsmaat, heeft de Mahalanobis afstand als voordeel dat deze niet afhankelijk is een gekozen functionele vorm. Een nadeel van deze afstandsmaat is dat de voorwaarde *gelijke basis van kenmerken* moeilijk te testen is. In de wetenschappelijke literatuur zijn geen duidelijke regels geformuleerd wanneer welke afstandsmaat beter is. Een aanbeveling is wel dat de propensity score meegenomen wordt als relevant kenmerk indien koppeling plaatsvindt op basis van de Mahalanobis afstand, omdat er dan getest kan worden op *gelijke basis van kenmerken* zonder dat de analyse zijn niet-parametrische kracht verliest.

Onderscheid makende tussen de hierboven genoemde afstandsmaten, zullen we in dit onderzoek de volgende koppelingsmanieren toepassen:

Op basis van de ‘propensity score’ afstandsmaat:

1. Nearest neighbor matching zonder teruglegging;
2. Nearest neighbor matching met teruglegging
3. Nearest neighbor matching aan de 5 nearest neighbors;
4. Nearest neighbor matching aan de 10 nearest neighbors.
5. Radius matching met een straal van 0.01;

²¹ Formeel geldt dat de Mahalanobis-afstand tussen vrouw i en man j op basis van kenmerken X uitgedrukt kan worden als: $D_{ij} = (X_i - X_j)' \Sigma^{-1} (X_i - X_j)$, waarbij Σ^{-1} de covariantiematrix voorstelt.

6. Radius matching, met een straal van 0.05
7. Radius matching, met een straal van 0.001
8. Kernel matching met een Epanechnikov verdelingsfunctie

Op basis van de Mahalanobis afstandsmaat:

9. Mahalanobis matching;
10. Mahalanobis matching met de propensity score opgenomen als koppelingsvariabele.

De manieren van matching zoals hierboven weergegeven kunnen samengevat worden in 4 groepen, namelijk: nearest neighbor matching, radius matching, kernel matching en mahalanobis matching. Hieronder zullen we de verschillende manieren van matching kort uiteenzetten.

'Nearest neighbor'

Deze koppelingsregel rangschikt personen in de testgroep en de controlegroep willekeurig. Vervolgens wordt voor de eerste vrouw uit de testgroep een man uit de controlegroep gevonden die het minst afwijkt in termen van propensity score. Omdat deze regel alleen kijkt naar de relatieve verschillen tussen mannen en vrouwen is het mogelijk dat een vrouw gekoppeld wordt aan een mannelijke 'buurman' die in termen van propensity score op een behoorlijke afstand ligt. Om de robuustheid van de verkregen loonverschillen te testen zullen we verschillende varianten van 'nearest neighbor' matching toepassen:

- Koppeling met de nearest neighbor met en zonder teruglegging;
- Koppeling met de eerste 5 nearest neighbors;
- Koppeling met de eerste 10 nearest neighbors.

Als de controlegroep bestaat uit oneindig veel mannen met een oneindig aantal verschillende combinaties aan kenmerken, is er geen verschil tussen koppeling met teruglegging en koppeling zonder teruglegging. In de AVO-data geldt dat de ratio man/vrouw voor het bedrijfsleven 1.7 is en voor de overheidssector 1.1. er zijn dus meer mannen in de dataset aanwezig, maar niet heel veel meer. De verwachting is dat koppeling zonder teruglegging minder nauwkeurig wordt naarmate er meer mannen gekoppeld zijn, omdat de kans op een nauwkeurige volgende koppeling steeds kleiner wordt. We zullen dit verduidelijken aan de hand van een voorbeeld. Stel dat er 2 identieke vrouwen zijn (mevrouw A en mevrouw B) voor wie we op zoek zijn naar een zo goed mogelijk gelijkende man. Voor mevrouw A vinden we een man die heel erg op haar lijkt. Als we koppelen zonder teruglegging dan is het niet mogelijk om ook mevrouw B te koppelen aan deze man, ook zou dit de beste match zijn op basis van de afstandsmaat. De koppeling voor mevrouw B kan dus nooit beter zijn dan die voor mevrouw A; er ontstaat 'ruis' bij het matchen en die beïnvloedt de zuiverheid van de schatter. Naarmate de controlegroep relatief kleiner is zal deze 'ruis' groter zijn.

Een alternatieve koppelingsmethode is die waarbij we een vrouw aan meer dan één nearest neighbor koppelen. De achterliggende gedachte is dat het loon van de 5 nearest neighbors een loon geeft dat representatiever is dan dat van slechts één nearest neighbor. Aan de andere kant is het zo dat de koppelingen gebaseerd zijn op de relatieve afstand, dus als een vrouw gekoppeld wordt aan 5 mannen kan het zo zijn dat de 4^e en 5^e nearest neighbor vrij ver van de vrouw afliggen in termen van de afstandsmaat. Het opnemen van meer dan één nearest neighbor kan

thus zorgen voor meer onzuiverheid in de analyse. Door alle verschillende varianten van nearest neighbor te analyseren krijgen we een idee over de robuustheid van de verkregen schatter, en over de impact van de verschillende onderliggende aannames die per variant kunnen verschillen.

Radius Matching

Bij radius matching wordt iedere vrouw gekoppeld aan alle mannen die minder dan een bepaalde kanshoeveelheid van hen afwijken. Omdat het hier gaat om een afwijking in kanshoeveelheid vindt deze koppeling plaats op basis van de propensity score. Het loon van elke vrouw wordt vervolgens vergeleken met dat van het gemiddelde loon van de mannen die aan haar gekoppeld zijn. De gedachte achter radius matching is dat het gemiddelde loon van meerdere matches representatiever is dan dat van één enkele match. Merk op dat deze gedachte gelijk is aan die bij nearest neighbor matching waar een vrouw gekoppeld wordt aan meer dan één man. Een belangrijk verschil is dat er bij radius matching een grens wordt gesteld in de hoeveelheid dat een man zou mogen afwijken, wil hij in aanmerking komen als match. Een nadeel is dat deze grens arbitrair is: op voorhand is niet duidelijk hoe groot de maximale afwijking zou moeten zijn. In dit onderzoek analyseren we de volgende afwijkingsgrenzen:

- Radius matching met een straal van 0.01;
- Radius matching met een straal van 0.05;
- Radius matching met een straal van 0.001

Hoe kleiner de gekozen straal (de tolerantie) hoe kleiner naar verwachting het aantal mogelijke koppelingen per vrouw. Met andere woorden, naarmate de straal kleiner gekozen wordt zullen er steeds minder mannen op een vrouw lijken. Voor elke vrouw bestaat er een straal waarbij geen enkele man als potentiële koppelingspartner wordt aangemerkt.

Kernel matching

Bij kernel matching worden alle vrouwen in de testgroep gekoppeld aan een gewogen gemiddelde van alle mannen in de controlegroep. Bij een vrouw in de testgroep wordt een man uit de controlegroep gekozen met een bepaalde kans, die omgekeerd evenredig is met de gekozen afstandsmaat. Hoe verder weg, hoe kleiner de kans. Omdat alle mannen in principe een kans hebben om ‘mee te tellen’ in de evaluatie, resulteert kernel matching in een koppeling waarbij de afwijking in de propensity score een lagere variantie heeft. Er moet bij kernel matching een verdelingsfunctie gekozen worden (wat de analyse niet-parametrisch maakt), in de praktijk blijkt dat de efficiëntie van de schatters niet beïnvloed wordt door het type verdelingsfunctie dat gebruikt wordt (Dehejia en Wahba, 1999)²². In dit onderzoek zullen we gebruik maken van de Epanechnikov verdelingsfunctie, omdat er in de wetenschappelijke literatuur naar deze functie wordt verwezen als de meest efficiënte (zie Cammeron en Trivedi, pag 303).

Mahanobis matching

Deze manier van koppelen is gelijk aan nearest neighbor matching maar als afstandsmaat wordt de mahalanobis-afstand gebruikt. In dit onderzoek passen we 2 varianten toe:

- Mahalanobis matching; en

²² Zie ook <http://www.quantlet.com/mdstat/scripts/spm/html/spmhtmlnode16.html>

- Mahalanobis matching met de propensity score opgenomen als matching-variabele.

Het voordeel van beide varianten is dat het loonverschil tussen mannen en vrouwen niet parametrisch bepaald wordt en dus niet afhangt van een gekozen functionele vorm. Door de propensity score toe te voegen in het tweede alternatief kan er bovendien getest worden of aan de voorwaarde *gelijke basis van kenmerken* voldaan wordt.

Ten slotte merken we op dat in de huidige matching literatuur geen duidelijke regels zijn geformuleerd over welke koppelingsregel het beste is: iedere koppelingsregel heeft zijn voordelen maar ook zijn nadelen. Door bovenstaande koppelingsregels toe te passen krijgen we inzicht over de verschillen in uitkomst tussen de verschillende koppelingsregels en testen we tevens de robuustheid van de schatters verkregen met matching.

4.2 Matching resultaten basisloonschatting

In deze paragraaf bespreken we de schattingsresultaten van de loonverschillen tussen mannen en vrouwen die verkregen zijn op basis van de 10 koppelingsvarianten die in de vorige paragraaf zijn geïntroduceerd. Wederom maken we onderscheid tussen het bedrijfsleven en de overheidssector. Als uitgangspunt voor de koppeling nemen we de verklarende variabelen die ook gebruikt worden in paragraaf 3.1. Daardoor zijn de uitkomsten in deze paragraaf vergelijkbaar met de loonverschillen in de basisloonschatting zoals beschreven in Hoofdstuk 3 (zie Tabel 3.1). De schattingsresultaten voor het bedrijfsleven staan beschreven in Tabel 4.1, waarbij we voor de volledigheid ook de eerdere uitkomst o.b.v. OLS uit Tabel 3.1 vermelden in de tabel. Alle vermelde resultaten zijn statistisch significant.

Tabel 4.1 Loonverschillen m-v in het bedrijfsleven: matching met verklarende variabelen analoog aan basisloonschatting

Matching Methode	Loonverschil	Standaardfout	95% B.I	
1-op-1 matching zonder teruglegging	-0,175	0,005	-0,185	-0,166
1-op-1 matching met teruglegging	-0,038	0,011	-0,061	-0,017
1-op-5 matching	-0,035	0,009	-0,053	-0,018
1-op-10 matching	-0,032	0,008	-0,049	-0,016
Mahalanobis matching	-0,058	0,010	-0,078	-0,038
Mahalanobis matching met propensity score	-0,057	0,010	-0,077	-0,036
Radius matching (r = 0.01)	-0,034	0,008	-0,050	-0,018
Radius matching (r = 0.05)	-0,033	0,008	-0,049	-0,017
Radius matching (r = 0.001)	-0,036	0,008	-0,053	-0,019
Kernel matching (Epanechnikov)	-0,033	0,008	-0,049	-0,017
Loonverschil o.b.v. OLS (Tabel 3.1)	-0,074	0,003	-0,082	-0,071

Het loonverschil bij 1-op-1 matching zonder teruglegging wijkt af van de overige geschatte loonverschillen. Dit verschil ontstaat doordat in deze variant kwaliteit van een koppeling minder is wanneer de controlegroep niet veel groter is dan de testgroep²³ (zie uitleg in de vorige paragraaf). We richtten ons daarom op de overige koppelingsvarianten.

²³ Om deze verklaring te testen hebben we een beperkt aantal vrouwen getrokken uit de dataset en deze gekoppeld aan mannen in de steekproef op basis van 1-op-1 matching zonder teruglegging. Dit is eerst

Uit de tabel blijkt dat we de schattingen van het loonverschil tussen mannen en vrouwen kunnen verdelen in twee groepen. Voor de eerste groep wordt het loonverschil geschat op basis van Mahalanobis matching en is de uitkomst gemiddeld 5.8 procent. Voor de tweede groep wordt het loonverschil geschat op basis van de propensity score en voor deze groep wordt het loonverschil geschat op 3.5 procent. Voor beide groepen zijn de geschatte loonverschillen kleiner dan het loonverschil volgens de traditionele methode. Toch blijft er wel degelijk nog een onverklaard loonverschil bestaan. Bij propensity score matching hebben we vrij uiteenlopende koppelingmethoden toegepast, maar desondanks blijken de voorspelde loonverschillen hiervoor zeer robuust te zijn.

In Tabel 4.2 presenteren we voorspelde loonverschillen voor de overheidssector. De statistisch significante resultaten zijn weergegeven met een sterretje. Evenals voor het bedrijfsleven is het geschatte loonverschil voor de overheidssector veel groter wanneer er een 1-op-1 koppeling plaatsvindt zonder teruglegging. De koppeling is via deze methode is echter van onvoldoende kwaliteit, zodat we in het vervolg de schattingsresultaten van deze koppelingsvariant niet meer bespreken.

Tabel 4.2 Loonverschillen m-v in de overheidssector: matching met verklarende variabelen analoog aan basisloonschatting

Matching Methode	Loonverschil	Standaardfout	95% B.I	
1-op-1 matching zonder teruglegging	-0,133*	0,003	-0,138	-0,128
1-op-1 matching met teruglegging	-0,061	0,036	-0,134	0,012
1-op-5 matching	-0,048*	0,016	-0,079	-0,017
1-op-10 matching	-0,048*	0,011	-0,070	-0,027
Mahalanobis matching	-0,021	0,036	-0,093	0,050
Mahalanobis matching met propensity score	-0,021	0,036	-0,093	0,050
Radius matching (r = 0.01)	-0,052*	0,004	-0,060	-0,043
Radius matching (r = 0.05)	-0,054*	0,004	-0,062	-0,046
Radius matching (r = 0.001)	-0,046*	0,004	-0,055	-0,037
Kernel matching (Epanechnikov)	-0,054*	0,004	-0,062	-0,046
Loonverschil o.b.v. OLS (Tabel 3.1)	-0,040*	0,001	-0,042	-0,037

Het gemiddelde loonverschil van alle propensity score koppelingsvarianten, met uitzondering van de variant zonder teruglegging, is 5.2 procent. Dit loonverschil is groter dan het loonverschil van 4 procent in de traditionele methode uit het vorige hoofdstuk.

De Mahalanobis matchingsvarianten zijn niet significant. Bovendien vallen de geschatte loonverschillen verkregen met andere methoden binnen het betrouwbaarheidsinterval van de Mahalanobis puntschatter. Met andere woorden: er kan niet met 95 procent zekerheid gesteld worden dat het loonverschil na Mahalanobis matching anders is dan het loonverschil op basis van andere varianten. Van de verschillende toegepaste koppelingsvarianten hebben Radius matching en Kernel matching de kleinste standaardfouten.

gedaan voor 10 procent van de vrouwen in de steekproef, daarna voor 20, 30, etc. totdat het aantal gelijk is aan alle vrouwen in de steekproef. Uit deze test komt naar voren dat het geschatte loonverschil goed overeenkomt met het loonverschil in de andere koppelingsvarianten zolang het aantal getrokken vrouwen kleiner is dan 1/3 maal de omvang van de controlegroep. Daarboven neemt het verschil snel en monotoon toe.

4.3 Matching resultaten modeluitbreiding

In de paragrafen 3.2 en 3.3 worden diverse modelverbeteringen voorgesteld ten opzichte van de basisloonschatting, waardoor de variatie in het uurloon beter wordt verklaard. Deze uitbreidingen zullen we nu ook implementeren wanneer we de 10 varianten van matching toepassen. Daardoor kunnen we de uitkomsten vergelijken met de uitkomsten in paragraaf 3.2 en 3.3, maar ook met die uit de vorige paragraaf. Zo krijgen we ook informatie over de veranderingen in voorspelde loonverschillen na verbetering van de modelspecificatie. We merken op dat er bij het matchen geen interactietermen opgenomen kunnen worden. In Tabel 4.3 tonen we de loonverschillen op basis van de verschillende koppelingsvarianten voor het bedrijfsleven. In deze tabel zijn alle vermelde resultaten statistisch significant.

Wederom leidt een verbetering van de modelspecificatie tot lagere loonverschillen. Met een propensity score matching methode wordt een gemiddeld loonverschil gevonden voor het bedrijfsleven van -2.8 procent. Net als in de basisspecificatie blijkt ook hier dat in het bedrijfsleven het loonverschil m-v kleiner is bij de matching dan bij de traditionele OLS-schatting uit Hoofdstuk 3.

Tabel 4.3 Loonverschillen m-v in het bedrijfsleven: matching inclusief modelverbetering

Matching Methode	Loonverschil	Standaardfout	95% B.I	
1-op-1 matching met teruglegging	-0,031	0,011	-0,053	-0,009
1-op-5 matching	-0,030	0,009	-0,048	-0,011
1-op-10 matching	-0,027	0,009	-0,045	-0,009
Radius matching (r = 0.01)	-0,027	0,008	-0,044	-0,010
Radius matching (r = 0.05)	-0,026	0,008	-0,043	-0,010
Radius matching (r = 0.001)	-0,028	0,009	-0,046	-0,011
Kernel matching (Epanechnikov)	-0,026	0,008	-0,043	-0,009
Loonverschil o.b.v. OLS (Tabel 3.6)	-0,066	0,021*	-0,665	-0,656

De uitkomsten voor de overheidssector staan vermeld in

Tabel 4.4 en deze zijn allen statistisch significant. Hier is het loonverschil bij propensity score matching juist groter dan het loonverschil in de OLS-schatting.

Tabel 4.4 Loonverschillen m-v in de overheidssector: matching inclusief modelverbetering

Matching Methode	Loonverschil	Standaardfout	95% B.I	
1-op-1 matching met teruglegging	-0,041	0,005	-0,052	-0,031
1-op-5 matching	-0,043	0,004	-0,051	-0,034
1-op-10 matching	-0,044	0,004	-0,053	-0,036
Radius matching (r = 0.01)	-0,045	0,004	-0,053	-0,037
Radius matching (r = 0.05)	-0,047	0,004	-0,055	-0,039
Radius matching (r = 0.001)	-0,045	0,004	-0,053	-0,037
Kernel matching (Epanechnikov)	-0,047	0,004	-0,055	-0,039
Loonverschil o.b.v. OLS (Tabel 3.6)	-0,027	0,012**	-0,028	-0,027

Op basis van de standaardfout en de robuustheid concluderen we dat matching een betere methode kan zijn om het effect te bepalen van geslacht op het verdiende uurloon, mits de juiste

vorm gehanteerd wordt. De geschatte loonverschillen op basis van propensity score matching zijn bij benadering vergelijkbaar, maar omdat kernel matching op theoretische gronden het meest geschikt is en bovendien de laagste standaardfout geeft geniet deze koppelingmethode de voorkeur.

Het theoretische voordeel van matching bestaat eruit dat alleen mannen en vrouwen met elkaar vergeleken worden die daadwerkelijk gelijk zijn. Vrouwen en mannen met dermate extreme kenmerken dat ze niet goed vergelijkbaar zijn (outliers) worden buiten beschouwing gelaten, waardoor de resultaten minder vertekend worden. Tegelijkertijd kunnen man-vrouw vergelijkingen aan de randen van de verdeling, bv. 23-minners en 55-plussers nauwkeuriger uitgevoerd worden omdat ze niet meer gebonden zijn aan een functionele vorm (zie de opmerkingen aan het eind van het vorige hoofdstuk).

Naar aanleiding van eerdere commentaren hebben nog eens uitgebreid bekeken welke mannen er nu eigenlijk ‘overblijven’, welke dus niet gekoppeld worden aan een vergelijkbare vrouw.²⁴ Dat blijkt vooral bepaald te worden door de deeltijdfactor, en met name bij de overheid. In die sector werken veel vrouwen in deeltijd, maar mannen relatief minder. Deze mannen worden dus in de matching-methode minder belangrijk, terwijl zij in de traditionele methode wellicht voor wat vertekening kunnen zorgen. Het feit dat dit meer speelt bij de overheid dan bij het bedrijfsleven, is een extra indicatie dat we niet de resultaten van de overheid en het bedrijfsleven onderling kunnen vergelijken, maar alleen binnen de sectoren.

4.4 Conclusie alternatieve schattingsmethode

Het schatten van loonverschillen op de AVO-data door middel van *propensity score matching* biedt enkele belangrijke technische voordelen. Het voordeel van matching bestaat eruit dat alleen mannen en vrouwen met elkaar vergeleken worden die daadwerkelijk gelijk zijn. Vrouwen en mannen met dermate extreme kenmerken dat ze niet goed vergelijkbaar zijn (outliers) worden buiten beschouwing gelaten, waardoor de resultaten minder vertekend worden. Tegelijkertijd kunnen man-vrouw vergelijkingen aan de randen van de verdeling, bv. 23-minners en 55-plussers nauwkeuriger uitgevoerd worden omdat ze niet meer gebonden zijn aan een functionele vorm. Als uitgangspunt voor de koppeling van mannen en vrouwen nemen we dezelfde variabelen als in de basisloonschatting, zodat de uitkomsten vergelijkbaar zijn met de loonverschillen in Hoofdstuk 3.

Bij de propensity score matching hebben we allerlei uiteenlopende koppelingmethoden toegepast, maar de voorspelde loonverschillen blijken hiervoor zeer robuust te zijn. Voor het bedrijfsleven wordt het loonverschil door matching geschat op 3,5 procent, dat is kleiner dan het loonverschil volgens de traditionele methode maar wel degelijk significant. Het verschil in uitkomsten tussen de traditionele schattingsmethode en matching suggereert dat in de traditionele methode vaker vrouwen en mannen worden vergeleken die eigenlijk slecht vergelijkbaar zijn; dit wordt blijkbaar niet voldoende gecorrigeerd door de functionele vorm. De modeluitbreidingen die in hoofdstuk 3 zijn toegepast, implementeerden we ook in de matching-analyses. Zo krijgen we ook informatie over de veranderingen in voorspelde loonverschillen na verbetering van de

²⁴ Indien van toepassing. Met kernel matching wordt uiteraard iedereen meegenomen, zij het met verschillende gewichten.

modelspecificatie. Net als bij de traditionele schattingsmethode leidt ook hier een verbetering van de modelspecificatie tot iets lagere loonverschillen (2,8% in plaats van 3,5%). Kortom, beide schattingsmethoden geven aan dat er uit de AVO-dataset nog meer te halen valt dan tot nu toe bekend was.

Tabel 4.5 Overzicht loonverschillen bedrijfsleven, AVO-data, OLS en matching.

	OLS	matching
<i>Bedrijfsleven</i>		
Basisloonschatting	7,4	3,5
Meest uitgebreide model	6,6	2,8

Voor de overheid wordt het gemiddelde loonverschil tussen mannen en vrouwen geschat op ongeveer 5 procent, dus iets groter dan met de traditionele methode. Helaas bevat de AVO-dataset voor de overheid veel minder variabelen, zijn ook lang niet alle subsectoren vertegenwoordigd en is het detailniveau veel lager. Vandaar dat de analyses van beide sectoren niet zomaar met elkaar vergeleken kunnen worden.

Tabel 4.6 Overzicht loonverschillen overheid, AVO-data, OLS en matching.

	OLS	matching
<i>Overheid</i>		
Basisloonschatting	4,0	5,0
Meest uitgebreide model	2,7	4,5

5 Alternatieve verklaringen: literatuurstudie

5.1 Definitie en geschiedenis van de loonkloof

5.1.1 Definities en begrippen

Het is internationaal gebruikelijk om de loonkloof te definiëren als het verschil tussen het (gemiddelde) mannen- en vrouwenloon, uitgedrukt in percentages van het mannenloon: $((m-v) \text{ in } \% m) \times 100\%$. Doorgaans wordt gerekend in bruto-uurlonen. De uitkomst van deze berekening wordt het ongecorrigeerde beloningsverschil genoemd. In Nederland was dit volgens het CBS eind 2005 18,3%. Dat betekent dat het gemiddelde uurloon van een vrouwelijke werknemer in ons land (€ 16,17) 18,3% lager was dan het gemiddeld uurloon van een mannelijke werknemer (€ 19,80). Onderstaande tabel laat zien dat in ieder geval in de industrie voor werknemers van 21 jaar en ouder (de enige categorie die over zo lange tijd vergeleken kon worden) de loonkloof m-v vanaf 1954 uiteindelijk fors is gedaald.

Tabel 1 Ontwikkeling loonkloof m-v in Nederland, werknemers 21 jaar en ouder in de industrie, 1947-2005

Jaar	(m-v) in % m
1947	38
1954	44
1960	40
1970	29
1980	21
1995	25
2000	21
2005	17

Bronnen: 1947-1970: Tegelaar, 1972; 1980-2005: CBS, diverse statistieken

Internationaal slaat Nederland niettemin geen goed figuur. Al langere tijd bivakkeert Nederland op de 17^e-19^e plaats onder de tot voor kort 25 EU-lidstaten. Rond 2005 lag de loonkloof volgens de Europese Commissie het laagst, rond 10%, in België, Polen en Hongarije, gevolgd door onder meer Frankrijk en Spanje met een kloof tussen 10 en 15%. Hoger dan Nederland kwamen Engeland, Finland en Duitsland uit, met een loonkloof van 20% of nog iets hoger (EC, 2007b; zie ook EC, 2003, en EF, 2006b). Ook de VS en Canada kennen een loonkloof van nationaal meer dan 20%. (OECD, 2005).

Een gecorrigeerd beloningsverschil ontstaat door te corrigeren voor verschillen in achtergrondkenmerken tussen werknemers, bijvoorbeeld voor verschillen in leeftijd, arbeidsduur, beroeps- of functieniveau en sector tussen mannen en vrouwen. Als we aldus corrigeren voor

achtergrondkenmerken, dan daalt het beloningsverschil uiteraard aanzienlijk. De omvang van het gecorrigeerde verschil hangt mede af van de gebruikte statistische analyse die wordt gebruikt en van de achtergrondkenmerken die worden meegenomen. Volgens onderzoek van de Arbeidsinspectie bedraagt het beloningsverschil in het particuliere bedrijfsleven in ons land na correctie al een aantal jaren tot 7%, terwijl het in 2004 bij de overheid na correctie 4% is (Spijkerman, 2000; Hoeben en Venema, 2004; Erdem *et al*, 2006, 59-60). “Dat er na correctie voor verschillen in achtergrondkenmerken nog een beloningsverschil bestaat, kan het gevolg zijn van het niet meenemen van alle relevante achtergrondkenmerken die van invloed kunnen zijn op de beloning (bijvoorbeeld de gezinssituatie, de motivatie of het arbeidsverleden bij andere werkgevers). Voorts kan er sprake zijn van een wat te grove meting en indeling bij enkele achtergrondkenmerken. Tot slot kan er sprake zijn van beloningsdiscriminatie”, aldus recent de Arbeidsinspectie (Erdem *et al*, 2006, 6). Hierop komen wij in later terug.

Er moet een duidelijk onderscheid worden gemaakt tussen twee begrippen: beloningsongelijkheid en beloningsverschillen. Het begrip beloningsongelijkheid is een juridische term: er mag tussen individuele mannen en vrouwen geen sprake zijn van ongelijk loon voor gelijkwaardig werk, zo is in ons land vastgelegd in de Wet Gelijke Behandeling van mannen en vrouwen (WGB). Het begrip beloningsverschil is een term uit de economische wetenschap. Door het gemiddelde uurloon van de mannelijke en de vrouwelijke beroepsbevolking uit te rekenen, kan het beloningsverschil tussen populaties van beide worden aangegeven. Deze verschillende begrippen worden herhaaldelijk door elkaar gebruikt (Tijdens, 2001).

5.1.2 Wetgeving en overheidsinitiatieven

Het streven naar gelijk loon tussen mannen en vrouwen kent een lange geschiedenis. Al in 1951 werd door de International Labour Organisation (ILO) Conventie no. 100 inzake gelijke beloning voor gelijkwaardige arbeid aangenomen. Pas twintig jaar later, in 1971, bekrachtigde de Nederlandse regering deze Conventie. De tussenliggende periode, vooral de jaren zestig, werd gebruikt om CAO-afspraken ongedaan te maken waarin vrouwenlonen werden uitgedrukt als een percentage van mannenlonen (Van Eijl *et al*, 1998, 183-184). In 1971 is in de CAO Wasserijen de laatste ongelijkheid van deze aard ongedaan gemaakt (Tegelaar, 1972). In 1975 werd vervolgens de Wet Gelijk Loon voor vrouwen en mannen (WGL) van kracht, die in 1980 opging in de Wet Gelijke Behandeling van mannen en vrouwen (WGB). In 1997 is aan het BW artikel 7: 646 toegevoegd, waarin het is verboden om onderscheid te maken op grond van geslacht bij (onder meer) de arbeidsvoorwaarden. Voor zover het gaat om loon, is in de WGB uitgewerkt hoe een vergelijking moet worden gemaakt. Wanneer een vrouw stelt dat zij ongelijk wordt beloond ten opzichte van een man, moet haar loon worden vergeleken met het loon dat een mannelijke collega ontvangt voor arbeid van (nagenoeg) gelijke waarde in – in principe – dezelfde onderneming. Het loon van de vrouw wordt geacht gelijk te zijn aan het loon dat een mannelijke collega voor ‘gelijkwaardig werk’ pleegt te ontvangen. Klachten terzake kunnen aan de Commissie Gelijke Behandeling (CGB) en aan de rechter worden voorgelegd. Blijkens haar jaarverslagen betreffen klachten die bij de CGB binnenkomen, slechts voor een beperkt deel ongelijk loon.

Een centraal begrip in de gelijk-loon discussie is ‘gelijkwaardig werk’. De wetgever verstaat hieronder nagenoeg gelijkwaardige taken bij dezelfde werkgever, of, in zeldzame gevallen, gelijkwaardige taken bij verschillende maar vergelijkbare werkgevers. Van oudsher is de objectieve

beoordeling van wat arbeid van gelijke waarde ofwel wat taken van dezelfde zwaarte zijn een centraal probleem geweest. Al in 1947 vond de Stichting van de Arbeid functiewaardering de meest aangewezen methode om de gelijkwaardigheid van functies aan te tonen. Inderdaad is de invoering van functiewaarderingssystemen belangrijk gebleken voor het verkleinen van de loonkloof. In de jaren '80 laat het Ministerie van Sociale Zaken en Werkgelegenheid onderzoeken uitvoeren naar mogelijke seksediscriminatie in functiewaarderingssystemen. In de jaren '90 is hierin belangrijke vooruitgang geboekt. Door intensieve discussies met de houders van functiewaarderingssystemen over de seksneutraliteit van systemen zijn de 'zorg'-aspecten in de arbeid, die met name, zoals in de gezondheidszorg, vrouwen raken, opnieuw geformuleerd. Alle veel in ons land gebruikte functiewaarderingssystemen zijn in de jaren daarna getest op seksneutraliteit (Van Klaveren *et al*, 2007).

De Europese eenwording is voor het streven naar gelijke beloning m-v niet onbelangrijk geweest. Het beginsel van gelijke beloning werd in 1957 verankerd in art. 119 van het EG-Verdrag (nu art. 141). Het laatste decennium heeft de Europese Unie (EU) zich ontwikkeld tot een motor achter pogingen om de loonkloof te verkleinen. De Europese Commissie (EC) publiceert regelmatig over de loonkloof en doet aanbevelingen tot verkleining. De sociale partners in de EU zijn in maart 2005 in een raamovereenkomst (*Framework of Actions on Gender Equality*) overeengekomen om de loonkloof te verkleinen. In 2006 brengt de EC *Een routekaart voor de gelijkheid van mannen en vrouwen* uit, lopend van 2006 tot 2010, waarin een samenhangende analyse van het effect van maatregelen om de gelijkheid van vrouwen en mannen dichterbij te brengen, wordt bepleit. Benadrukt wordt de samenhang tussen verschillende vormen van beleid ten aanzien van gelijkheid, zoals die ten aanzien van beloning, *work-life balance* en training (EC, 2006a, 2006b). Ook in 2006 wordt het *European Institute for Gender Equality* opgericht, dat technische ondersteuning kan bieden bij pogingen om de loonkloof m-v te verkleinen. In 2007 dringt de EC er hernieuwd bij de lidstaten op aan de loonkloof te bestrijden, en daartoe nationale doelstellingen en tijdslijnen vast te leggen (EC, 2007a).

5.2 Verklarende factoren

Er zijn vele honderden studies gedaan naar de 'gender pay gap', het beloningsverschil tussen mannen en vrouwen. De meesten daarvan betreffen de USA, maar er zijn ook studies bekend uit landen als Vietnam, Madagascar of ruraal Rusland. Dit hoofdstuk vat een aantal toonaangevende studies samen, waarbij de verklarende variabelen gegroepeerd zijn naar economische, sociologisch en psychologische verklaringen.

5.2.1 Economische factoren: menselijk kapitaal, sector en arbeidsduur

De economische wetenschap kent twee belangrijke, lange tradities van onderzoek naar lonen. De eerste onderzoekt het effect van het menselijk kapitaal (*human capital*) van werknemers. De tweede onderzoekt de loonvorming tussen sectoren. We bespreken daarnaast nog een derde, meer recente onderzoekslijn in de economische wetenschap, die ingaat op de arbeidsduur. Verklaringen voor beloningsverschillen tussen mannen en vrouwen bestaan begripsmatig uit de effecten van een viertal elementen:

- de samenstelling (compositie) van de desbetreffende groepen bij een bepaald kenmerk;
- de opbrengsten (returns) van een bepaald kenmerk;

- de preferentieverschillen tussen mannen en vrouwen bij een bepaald kenmerk;
- de selectievoorkeuren van werkgevers bij een bepaald kenmerk.

Het neoklassieke economische model veronderstelt een direct effect tussen preferentieverschillen en loonverschillen: vrouwen krijgen gemiddeld een lager loon omdat ze minder investeren in menselijk kapitaal. Vier indicatoren van menselijk kapitaal zijn: opleiding; werkervaring; dienstjaren bij de huidige werkgever (anciënniteit) en *on-the-job-training*. Voor alle vier geldt dat des te hoger c.q. langer de opleiding, ervaring, dienstjaren of training, des te hoger het loon, maar dat na verloop van een aantal jaren het effect afvlakt of zelfs negatief wordt. Om aan te geven of en hoe deze factoren doorwerken op de loonverschillen m-v moeten allereerst de compositie-effecten worden onderzocht. Weliswaar hebben vrouwen gemiddeld een lagere opleiding, minder bedrijfstraining, en minder ervaring dan mannen, maar deze verschillen worden in veel landen kleiner. Minder eenduidig zijn de resultaten wat betreft de m-v- verschillen in opbrengsten van deze factoren. Daarnaast zijn de effecten van zelf-selectie en werkgeversvoorkeuren vrijwel niet onderzocht.

Het onderzoek naar loonverschillen m-v tussen sectoren komt uit een onderzoekstraditie, waarin niet de kenmerken van werknemers maar de werking van de markt centraal staat. Gannon *et al* (2007), bijvoorbeeld, analyseren de loonverschillen tussen mannen en vrouwen in zes Europese landen op basis van loonverschillen binnen en tussen sectoren met behulp van de *European Structure of Earnings Survey* uit 1995. In deze zes landen bestaan er voor beide geslachten aanzienlijke loonverschillen tussen sectoren, maar de loonverdeling daarbinnen is zeer vergelijkbaar voor mannen en vrouwen, zij het dat in landen met gedecentraliseerde arbeidsverhoudingen de spreiding aanzienlijk groter is.

Andere studies bewandelen de omgekeerde weg en betrekken in onderzoek naar de invloed van beroepssegregatie op de loonverschillen m-v ook inter-sectorale verschillen. Zo concluderen Fields en Wolff (1995), op basis van een analyse van sectoren op 3-digit-niveau, dat 12-22% van de totale loonverschillen m-v in de VS kunnen worden verklaard uit de (meer)opbrengsten van het werken in een bepaalde sector en 15-19% uit de verschillende m-v-compositie van sectoren. Groshen (1991) toont voor vijf Amerikaanse sectoren aan dat de sekse-compositie van het beroep, van de bedrijfsvestiging en daarbinnen van de *job cell* allen significante effecten hebben op de loonverschillen tussen mannen en vrouwen. Verschillen in beroep verklaren de helft tot tweederde van de loonverschillen m-v binnen de sector, na controle voor de sekse-compositie. Deze uitkomsten worden ondersteund door die van Petersen en Morgan (1995) in hun onderzoek in 16 Amerikaanse sectoren en in tien beroepen. Hun voornaamste bevinding is dat beroepssegregatie (*occupational segregation*) de belangrijkste determinant van de loonverschillen m-v is. Macpherson en Hirsch (1995) komen tot de conclusie dat de loonkloof tussen mannen en vrouwen niet het product is van discriminatie binnen functies op vestigingsniveau. Dit resultaat ondersteunt de oudere conclusie van Oaxaca (1973) dat ongelijke beloning voor gelijkwaardig werk maar een klein gedeelte van de loonkloof kan verklaren.

De effecten van arbeidsduur op de loonverschillen m-v zijn diverse keren onderzocht. In Engeland wordt meestal een negatief effect gevonden van deeltijdarbeid (Ermish en Wright, 1992, 1993; Tam, 1997). Ermish en Wright laten zien dat ten aanzien van de verwachte lonen in vol- of deeltijdbanen het feit belangrijk is dat vrouwen voltijd werkten, terwijl het inkomen van de echtgenoot vooral bepalend is voor de vraag of vrouwen überhaupt gaan werken. Niet alleen

blijkt dat de lonen in deeltijdbanen lager zijn dan in voltijdbanen, ook waren de opbrengsten van opleiding en werkervaring groter in voltijdbanen. In Nederland wordt een soortgelijk effect alleen pas gevonden voor kleine deeltijdbanen (15 uur per week of minder). Indien gecorrigeerd voor diverse factoren, treden geen loonverschillen op tussen grotere deeltijdbanen en voltijdbanen van vrouwen (Hu, 2005).

5.2.2 Psychologische factoren: een 'fair' loon, taakbeschrijving en glazen plafond

In de psychologie bestaat een lange traditie van onderzoek naar gender-verschillen, maar weinig traditie als het gaat om loonverschillen tussen mannen en vrouwen. Wel is er onderzoek gedaan naar de waardering van het loon: vindt een werknemer dat hij/zij een eerlijk loon ontvangt voor de geleverde prestaties. Daarnaast is er in de arbeidspsychologie een traditie van onderzoek naar functieclassificatie en taakbeschrijvingen. En tenslotte is onderzoek verricht naar het glazen plafond. Deze drie clusters bespreken we hier.

Binnen de psychologie houdt een onderzoekslijn zich bezig met de perceptie van werknemers of hun loon een 'fair' (rechtvaardig) loon is voor een functie en hun inspanningen. Het begrip 'fairness' is hier niet identiek aan 'pay satisfaction', maar de twee begrippen liggen wel in elkaars verlengde. Uit onderzoek naar de percepties van 'self-pay', dus het loon waarop een individu meent recht te hebben, ook wel 'wage-entitlement' genoemd, blijkt dat mannen vaker dan vrouwen vinden dat zij een hoger loon waard zijn (Paul, 2006). Zoals veel psychologisch onderzoek is ook het onderzoek naar 'fair pay' veelal gebaseerd op experimenten, meestal bij studenten. Op grond van twee onderzoeken onder mannelijke studenten laten Hogue, Yoder en Sigleton (2007) zien dat de relatie tussen geslacht en 'self-pay-beoordelingen' loopt via 'self-reliance/self-assurance' naar 'self-competence ratings' en 'self-promotion entitlement', en tenslotte naar 'self-pay-beoordelingen'. In een onderzoek naar 'self-pay-beoordelingen' onder vrachtwagenchauffeurs in een bedrijf in de VS laten Verhoogen *et al* (2002) zien dat het relatieve loon (gemeten als het verschil tussen het feitelijke loon en het verwachte loon buiten het bedrijf) en de werkloosheidspercentage in de regio de twee belangrijkste determinanten van de *wage-fairness* perceptie waren. Mogelijk zijn *wage-fairness* percepties een factor in het arbeidsmarktgedrag van individuen.

In een studie op basis van de *British Social Attitudes Survey* (BSA) toont Paul (2006) aan dat bijna 37% van de werknemers in Engeland hun loon als unfair laag beschouwen en bijna 8% hun loon als unfair hoog. Vrouwen vinden gemiddeld hun loon vaker unfair dan mannen, maar als gecontroleerd wordt voor sector verdwijnt het verschil tussen mannen en vrouwen. Na controle voor meer variabelen slaat het m-v verschil om in zijn tegendeel: in dezelfde werksituaties percipiëren vrouwen vaker dan mannen dat zij fair beloond worden. Dit is in overeenstemming met eerder onderzoek, door Paul geciteerd, dat laat zien dat vrouwen in werksituaties waarin de ongecorrigeerde loonverschillen tussen mannen en vrouwen groot zijn dit toch niet als discriminatie of unfair pay percipiëren.

Het onderzoek naar functieclassificatie en functiebeschrijvingen vindt zijn oorsprong in de VS. Daar is van oudsher veel aandacht voor functiebeschrijvingen, vooral omdat in dat land een minder duidelijke relatie bestaat tussen het type opleiding en een beroep. De toelidingskanalen zijn niet erg beroepsspecifiek, zoals ze dat in Duitsland en in mindere mate in Nederland wel zijn. Al vrij snel is er in de VS aandacht voor de vraag of vrouwenberoepen anders worden beschreven

dan mannenberoepen, en voor de vraag of vrouwenberoepen dimensies kennen die minder beloond worden dan de dimensies van mannenberoepen, de zogenaamde *equal worth* benadering.

Het onderzoek naar het glazen plafond is vooral gebaseerd op gegevens van personeelsbestanden van bedrijven, omdat met deze bestanden de loopbanen van werknemers per maand of per jaar gevolgd kunnen worden. Naast het bestaan van een glazen plafond blijkt dat vrouwen daarenboven aan zwaardere promotiecriteria moeten voldoen, aldus analyses van personeelsbestanden van twee bedrijven door Jones en Makepeace (1996). Ook al zijn hun bevindingen statistisch significant, het effect van de verschillende behandeling van vrouwen is relatief gering in vergelijking met het effect van bijvoorbeeld anciënniteit. De auteurs berekenen dat als vrouwen en mannen gelijk behandeld zouden worden in de organisatie, het aandeel vrouwelijke managers zou stijgen van 1 naar 3%, maar als anciënniteit gelijk zou worden behandeld, zou het stijgen naar 20%. De effecten van het glazen plafond zijn moeilijk onderzoekbaar. Arulampalam *et al* (2007) volgen daarom een heel andere werkwijze. In navolging van Albrecht *et al* (2003) onderzoeken zij of de loonverschillen m-v groter of kleiner zijn in de verschillende strata van de loonverdeling. Zij interpreteren grotere loonverschillen m-v aan de top van de loon-distributie als een 'glazen plafond', terwijl ze grotere loonverschillen m-v aan de onderkant van de loonverdeling als een 'plakkerige vloer' betitelen. In hun onderzoek naar 11 landen blijkt dat slechts in twee sprake is van een plakkerige vloer, terwijl voor de andere landen een glazen plafond wordt geconstateerd.

5.2.3 Sociologische factoren: de impact van gezinsopbouw en segregatie naar sekse

In de sociologie is veel onderzoek gedaan naar de effecten op loonverschillen m-v van gezinsopbouw, van loopbaanonderbreking en herintreding, en van de verdeling van zorgarbeid binnen het huishouden. Een tweede onderzoekslijn houdt zich bezig met de effecten van sekse-segregatie naar beroep en workplace. Een derde lijn onderzoekt de effecten van promotie, leidinggeven, functieniveau en hiërarchische structuur in een organisatie. Bijna alle studies laten een effect zien van gezinsopbouw: het krijgen van kinderen heeft een nadelig effect op het loon van vrouwen en een voordelig effect op het loon van mannen. De belangrijkste verklaring is dat het arbeidsmarktgedrag van vrouwen en mannen gaat verschillen als er eenmaal kinderen geboren zijn in hun huishouden. In een studie naar Zweden en Japan toont Kumlin (2007) aan dat de sekse-verschillen in de financiële opbrengsten van kinderen in beide landen bijdragen aan het loonverschil m-v, maar dat dit effect in Japan groter is dan in Zweden. Over de verdeling van zorgarbeid binnen huishoudens en het effect daarvan op de loonkloof is niet zoveel onderzoek verricht, voornamelijk omdat de causale relatie hierbij niet duidelijk is.

Het onderzoek naar de effecten van sekse-segregatie naar beroep, functieniveau en workplace wijst op grote loonverschillen. Vrouwen in door vrouwen gedomineerde beroepen verdienen minder, ook als gecontroleerd wordt voor menselijk kapitaal en andere aspecten (England, 1992; le Grand, 1997; Bakker *et al*, 1999; de Ruijter, 2002). Het meten van de seksesegregatie van beroepen is niet onproblematisch, vooral als het gaat om de gedetailleerdheid waarmee de beroepen worden gemeten: worden beroepen in enkele grote groepen ingedeeld of in een paar duizend functies, of iets daar tussen in? Voor de bepaling van de sekse-compositie van beroepen is dit een belangrijke vraag, die echter empirisch lang niet altijd goed op te lossen is. Op basis van kleine databestanden kunnen alleen geaggregeerde beroepsgroepen worden gebruikt. Naast de

seks-compositie van het beroep heeft ook de seks-compositie van de werkplek een negatief effect op beloning, (Groshen, 1991; Simon en Russell, 2004). Kumlin (2007) laat zien dat in Zweden de relatieve bijdrage van de seks-compositie van de werkplek groter is dan in Japan. In vergelijking met mannen ondervinden vrouwen in Zweden een groter negatief effect op hun loon als zij werkzaam zijn in vestigingen of afdelingen met minder mannen.

De belangrijkste onderzoeksresultaten tot nu toe laten zien dat de loonverschillen m-v voor een groter deel toegeschreven moeten worden aan de tussen-werkgevers en tussen-beroepen effecten dan aan de zogenaamde binnen beroep-werkgever effecten. Opnieuw blijkt het onwaarschijnlijk dat werkgevers systematisch loondiscriminatie toepassen. Loonverschillen m-v moeten eerder toegeschreven worden aan de allocatieprocessen tussen beroepen en tussen werkgevers. Onderzoeken naar de effecten van beroepensegregatie op lonen tonen steeds aan dat lonen substantieel lager zijn in door vrouwen gedomineerde beroepen. Macpherson en Hirsch (1995) concluderen dat de bijdrage van de gemeten en de ongemeten beroepskwalificaties en baankenmerken grotendeels de loonverschillen tussen beroepen verklaart, en dus niet zozeer de seks-compositie van een beroep. Tam (1997) vindt steun voor de gespecialiseerde menselijk kapitaal-hypothese: de effecten van de seks-compositie van beroepen hangen grotendeels af van de looneffecten van beroepsspecifieke training. Anders gezegd, beroepsvaardigheden en beroepstraining verklaren grotendeels het effect van de seks-compositie van beroepen op de loonverschillen m-v.

5.3 Onderzoek naar loonverschillen in Nederland

5.3.1 Inleiding

In dit hoofdstuk geven we een redelijk uitputtend overzicht van de onderzoeken naar de loonverschillen m-v in Nederland. Het betreft 20 studies. De bijlage bevat een tabel over de gebruikte databronnen, de onderzoekspopulaties, eventuele schattingen op deelpopulaties, en de gebruikte schattingsmethode. In dit hoofdstuk bespreken we deze onderzoeken, net als in het vorige hoofdstuk, ingedeeld naar economische, psychologische en sociologische factoren. Alvorens dat te doen gaan we hier kort in op de gehanteerde statistische methoden.

Loonverschillen m-v kunnen op diverse manieren worden geanalyseerd. Bivariate analyses geven veel inzicht in mogelijke verschillen tussen de seksen, bijvoorbeeld door het uurloon van vrouwelijke en mannelijke voltijders en deeltijders naast elkaar te zetten. In multivariate analyses kunnen deze verschillen gecontroleerd worden voor overige kenmerken: blijft het verschil in uurloon naar arbeidsduur en seks bestaan als bijvoorbeeld gecontroleerd wordt voor opleiding? En blijft dit verschil bestaan als ook het effect van sector of dienstjaren in de analyse wordt betrokken? Multivariate analyses zijn de meest gebruikte manier om loonverschillen m-v te analyseren. De meest eenvoudige methode is om een loon-regressiemodel te schatten voor een gehele populatie. Het effect van geslacht wordt dan geanalyseerd door een sex-dummy op te nemen in het model. Het effect van geslacht op het uurloon wordt ook wel het onverklaarde deel van de loonvergelijking genoemd, dus na controle voor overige factoren. Het nadeel van zo'n analyse is dat het compositie-effect (bijv. mannen hebben vaker dan vrouwen een leidinggevende positie) niet onderscheiden kan worden van het zogenoemde opbrengsten (*returns*) -effect (bijv. voor mannen brengt het financieel meer op om in een leidinggevende positie te werken dan voor

vrouwen). Daarom is het gebruikelijker om eerst een loon-regressiemodel te schatten voor mannen en voor vrouwen afzonderlijk. Dan kunnen wel de groottes van de coëfficiënten, dus het opbrengsten-effect, vergeleken worden tussen mannen en vrouwen. Als daarna nog een decompositie wordt uitgevoerd, dan kan zowel per kenmerk als voor het totale loonverschil de effecten van de compositie en van de opbrengsten tussen mannen en vrouwen worden onderscheiden, zoals ontwikkeld door Blinder (1973) en Oaxaca (1973).

Er speelt nog wel een andere discussie, namelijk of de allocatie van mannen en vrouwen naar bepaalde posities op de arbeidsmarkt een gevolg is van discriminatie dan wel van zelfselectie. Daarmee komt de causaliteit van de positie op de arbeidsmarkt naar het uurloon ter discussie. Hu (2005), bijvoorbeeld, corrigeert eerst voor zelfselectie in deeltijd- en voltijdbanen en schat daarna pas het uurloon. Albrecht *et al* (2004) zijn van mening dat het echter onjuist is om te corrigeren voor functieniveau, omdat loonverschillen m-v hier mogelijk een gevolg zijn van seksediscriminatie. Deze discussie is nog niet beslecht.

5.3.2 Economische factoren

Invloed van menselijk kapitaal: opleiding, werkervaring, anciënniteit en on-the-job training

Het menselijk kapitaal bestaat uit opleiding, werkervaring, anciënniteit en on-the-job training. In de geïndustrialiseerde landen geldt dat werknemers met een groter menselijk kapitaal meer verdienen. Dat is ook in Nederland het geval. Daar staat tegenover dat de mannelijke en vrouwelijke beroepsbevolking in Nederland verschilt van elkaar in termen van menselijk kapitaal: mannen hebben gemiddeld meer opleiding, werkervaring, anciënniteit en on-the-job training dan vrouwen. Dit leidt tot een compositie-effect in het beloningsverschil m-v. Daarenboven is de vraag of de financiële opbrengsten uit menselijk kapitaal voor mannen en vrouwen verschillend zijn. Een methodologisch probleem is dat in veel databestanden niet alle vier kenmerken van menselijk kapitaal aanwezig zijn en dus ook niet opgenomen kunnen worden in de analyse. In sommige analyses wordt bij gebrek aan beter leeftijd genomen als proxy voor werkervaring of anciënniteit. Dat is voor vrouwen problematischer dan voor mannen, omdat vrouwen veel vaker een langdurige loopbaanonderbreking hebben gehad, waardoor leeftijd een minder goede indicator is voor hun werkervaring. We bespreken hier de onderzoeken die op dit onderwerp ingaan.

Op basis van de bestanden 2002 en 2004 van het Arbeids Voorwaarden Onderzoek (AVO) concludeert de Arbeidsinspectie dat in de het bedrijfsleven bij elke opleidingscategorie de financiële opbrengsten van opleiding groter zijn voor mannen dan voor vrouwen (Hoeben en Venema, 2004; Erdem *et al*, 2006). Op basis van het bestand van de Enquête Werkgelegenheid en Lonen (EWL) voor de overheidssector vinden Rienstra en Hagoort (2006) eveneens dat de opbrengsten van opleiding groter zijn voor mannen dan voor vrouwen, met uitzondering van de Hbo-opleiding. In een onderzoek naar hoogopgeleide mannen en vrouwen vinden Kalmijn en Van der Lippe (1997) dat de opbrengsten van door mannen gedomineerde opleidingen hoger zijn dan van door vrouwen gedomineerde opleidingen. Op basis van het bestand van het Sociaal-

Economisch Panelonderzoek (SEP) zijn volgens Bakker *et al* (1999) daarentegen de financiële opbrengsten van opleiding voor vrouwen hoger dan voor mannen.²⁵ Dat geldt ook voor opleidingsrichting. Bij alle onderzoeken is gecontroleerd voor functieniveau en voor opleidingsrichting (SEP, EWL) en voor functiesoort (AVO). Bij alle onderzoeken is het loon-effect van functieniveau voor vrouwen kleiner dan voor mannen.

Wetzels en Tijdens (2002) hebben alleen voor vrouwen een loonregressie geschat. De resultaten laten zien dat, zoals verwacht, elk extra opleidingsjaar en elk extra leeftijdjaar een substantieel positief effect hebben op het uurloon (respectievelijk 0.055 en 0.022). Leeftijd kwadraat, indicierend dat na een bepaalde leeftijd het effect van leeftijd op uurloon negatief wordt, heeft een klein negatief effect op het loon (-0.007). Ook heeft elk dienstjaar bij de huidige werkgever extra een klein positief effect op het uurloon. Op basis van de SEP1996 data concludeert De Ruijter (2002) dat in haar model het aantal dienstjaren voor vrouwen een groter effect op het netto uurloon heeft dan voor mannen. Ook Bakker *et al* (1999) vinden dat dienstjaren voor vrouwen een groter effect hebben dan voor mannen. In de AVO2002 en 2004 onderzoeken worden voor de marktsector dezelfde uitkomsten gevonden, maar in het EWL onderzoek naar de overheidssector heeft een extra jaar alleen een groter effect voor vrouwen bij 6 tot en met 20 dienstjaren.

Uitgesplitst naar leeftijdsgroep blijkt uit onderzoek onder werknemers in de overheidssector dat in de leeftijdsgroep 16 tot 23 jaar het uurloon van vrouwen hoger is dan het uurloon van mannen. In alle overige leeftijdscategorieën is het uurloon van mannen hoger dan dat van vrouwen. Naarmate de leeftijd hoger is, is het beloningsverschil tussen mannen en vrouwen groter (Rienstra en Hagoort, 2006).

Invloed van sector, bedrijfsgrootte, CAO en beloningsbeleid

In het vorige hoofdstuk is ingegaan op de beloningsverschillen tussen sectoren en op het feit dat vrouwen vaker werken in sectoren met een laag loon. Ook in Nederland is dat het geval. De ongecorrigeerde loonkloof m-v is in sommige sectoren hoog: ruim 30% in de financiële instellingen en 25-35% in de verschillende delen van de gezondheidszorg (Van Klaveren *et al*, 2007). Internationaal blijkt de financiële dienstverlening allereerste een grote loonkloof te vertonen, terwijl de collectieve sector in het algemeen een wat lagere kloof laat zien dan de marktsector (Plantenga en Remery, 2006).

Er zijn verschillende methoden om de variabele sector op te nemen in statistische analyses. Het meeste onderzoek in Nederland neemt sector op als een sterk geaggregeerde variabele op 1-digit niveau, waarmee uitgaande van de referentiegroep het looneffect voor elke sector afzonderlijk kan worden onderzocht. Zo onderscheidt de Arbeidsinspectie in het bedrijfsleven 9 sectoren.²⁶ Andere studies nemen slechts het onderscheid tussen de collectieve en de marktsector op in hun analyse.

²⁵ Een reden hiervoor zou een verschil in meetmethode kunnen zijn. Hoeben en Erdem meten opleidingsniveau als dummy variabele met vijf categorieën en 'onbekend', waarbij de laagste categorie als referentiegroep is opgenomen. Van een aanzienlijk deel (19%) van de observaties is geen opleidingsniveau bekend. Rienstra en Hagoort hebben zeven opleidingscategorieën gebruikt met de middelste als referentiegroep. Bakker *et al* (1999) behandelen opleidingscategorie als een interval-variabele. Hier zijn alleen observaties genomen waarvan het opleidingsniveau bekend is.

²⁶ Het AVO-bestand bevat geen gegevens van de overheidssector, wel van de gezondheids- en welzijnszorg.

Is er een verband tussen de loonkloof en het aandeel van vrouwen in een bepaalde sector? Is in sectoren waar relatief veel vrouwen werken de loonkloof groter dan in sectoren die door mannen worden gedomineerd? Volgens Van Klaveren *et al* (2007) is bivariaat geen duidelijk patroon te ontdekken voor de geselecteerde 7 sectoren in hun onderzoek. Gemiddeld genomen is de loonkloof iets groter in sectoren met een klein aandeel van vrouwen dan in sectoren met een groot aandeel. Opvallend is verder dat de patronen tussen 1995 en 2005 nauwelijks zijn veranderd. Ander Nederlands onderzoek naar de relatie tussen het percentage vrouwen in een sector en het uurloon is niet aangetroffen.

Uit eerdere beloningsonderzoeken van SEO is gebleken dat vrouwen in de collectieve en de marktsector minder goed worden betaald dan mannen, maar vooral in voltijdbanen (Berkhout *et al*, 2004; Berkhout *et al*, 2007a; Berkhout *et al*, 2007b; Heyma *et al*, 2004; Heyma *et al*, 2005). Mannen hebben over het algemeen een relatief beloningsnadeel in de collectieve sector t.o.v. de marktsector, voor vrouwen is juist de collectieve sector relatief gunstig. Al geldt dat laatste alleen nog voor vrouwen in voltijdbanen, want die hebben in de marktsector een nog grotere achterstand op mannen dan bij de overheid. Opmerkelijk is dus dat het beloningsverschil m-v voor deeltijdwerkers gelijk is over beide sectoren. Uitgesplitst naar subsector verdienen werknemers in het openbaar bestuur iets meer, in het onderwijs hetzelfde, maar in de zorg 6% minder dan in de marktsectoren. Vrouwen zijn beter af in de collectieve sector, want in alle subsectoren van de overheid verdienen zij na correctie meer dan wanneer zij in de private sector zouden hebben gewerkt. Uit het onderzoek onder vrouwelijke werknemers van Wetzels en Tijdens in (2002) blijkt dat werken in de zorgsector bij vrouwen een klein positief effect heeft op het loon. Bakker *et al* (1999) concluderen dat werken in de collectieve sector een negatieve invloed op het loon van mannen, niet op dat van vrouwen. Als mannen en vrouwen gelijk verdeeld zouden zijn over de sectoren zou het beloningsverschil m-v kleiner zijn.

De beloning in sectoren verschilt ook van elkaar omdat beloningssystemen per sector variëren. Ook werken vrouwen vaker dan mannen in bedrijven waar een beloningssysteem met salarisschalen wordt toegepast (67% versus 51%). Volgens de AVO-data is in het bedrijfsleven een significant negatief effect zichtbaar van de toepassing van salarisschalen op het uurloon. Afzonderlijk voor vrouwen en mannen geanalyseerd blijkt uit AVO-2002 dat de toepassing van salarisschalen voor mannen een loondrukkende werking heeft, maar voor vrouwen dit juist financieel profijtelijk is (Hoeben en Venema, 2004, 176). Volgens AVO-2004 echter ondervinden zowel mannen als vrouwen een significant negatief effect van de toepassing van salarisschalen (Erdem *et al*, 2006, 178).

De beloning in sectoren verschilt ook van elkaar omdat de Cao-werking per sector varieert. Daarbij blijken compositieverschillen. Vrouwen werken vaker dan mannen onder de werkingsfeer van een Cao, al dan niet via een AVV, algemeen-verbindend-verklaring (78% versus 72%). Voor het uurloon maakt het noch voor vrouwen, noch voor mannen, noch in 2002 noch in 2004 uit of een werknemer, onder een ondernemings- of bedrijfstak-cao valt, wel pakt het uit nadeling uit voor het uurloon van vrouwen als een werknemer onder de AVV valt (Hoeben en Venema, 2004; Erdem *et al*, 2006). Tijdens *et al* (2002) vinden op basis van Loonwijzer-data wel een positief effect van de werkingsfeer van de Cao op het uurloon van vrouwen.

Ook de bedrijfsgrootte is van invloed op het uurloon. Werknemers in grote bedrijven verdienen per uur gemiddeld meer dan werknemers in kleine bedrijven. In lijn hiermee vinden Wetzels en

Tijdens (2002) dat bedrijfsgrootte een positief effect heeft op het loon van vrouwelijke werknemers: naarmate het bedrijf groter is, is hun uurloon hoger, namelijk 0.02 voor elke additionele categorie, lopend van 1, minder dan 10 werknemers, tot 6, meer dan duizend werknemers. Alleen voor herintreedsters is het effect van bedrijfsgrootte op het uurloon kleiner. Uit de analyses op AVO-2002 blijkt dat het effect van werken in een groot bedrijf in het bedrijfsleven voor mannen iets groter is dan voor vrouwen, terwijl dit voor AVO-2004 net andersom is (Hoeben en Venema, 2004; Erdem *et al*, 2006).

Variabele beloningselementen zijn meestal niet in de statistische analyses van de loonkloof opgenomen. De AVO-bestanden bevatten een dichotome variabele 'extra looncomponenten'. Het uurloon van mannen wordt door deze extra looncomponenten meer verhoogd dan dat van vrouwen (Hoeben en Venema, 2004; Erdem *et al*, 2006). Een schriftelijke enquête van PWC (2002) onder HRM-functionarissen geeft aan dat vooral functiecategorieën waarin relatief weinig vrouwen werken, zoals de buitendienst en de hogere functieniveaus, in aanmerking komen voor variabele elementen; soms geldt ook de voorwaarde van een minimaal aantal dienstjaren. Uit het onderzoek blijkt dat vrouwen in de sectoren Voeding en Finance sterker vertegenwoordigd zijn in bedrijven zonder variabele beloning, ongeacht de bedrijfsomvang. Deze verschillen tussen bedrijven met en zonder variabele beloning kunnen komen doordat vrouwen relatief meer dan mannen voor een werkgever kiezen die geen variabele beloning toepast, of doordat bedrijven waar relatief veel vrouwen werken minder snel kiezen voor een variabel beloningssysteem.

Invloed van arbeidsduur

Meer dan in enig ander geïndustrialiseerd land wordt in Nederland in deeltijd gewerkt. Driekwart van de vrouwen en ruim een-tiende van de mannen heeft een deeltijd baan. De effecten van arbeidsduur, na controle voor overige kenmerken, zijn in de meeste hier besproken onderzoeken geanalyseerd. Heeft het werken in deeltijd een negatief effect op het uurloon als gecontroleerd wordt voor andere kenmerken?

Erdem *et al* (2006) concluderen dat grote deeltijders (12 uur per week of meer) bij de vrouwen *ongecorrigeerd* 10% en kleine deeltijders (minder dan 12 uur per week) 47% minder dan voltijders verdienen. Bij de mannen zijn deze verschillen nog groter: *ongecorrigeerd* verdienen grote deeltijders 30% en kleine deeltijders 88% minder dan voltijders. Na correctie voor overige kenmerken verdienen bij de vrouwen de grote deeltijders 1,3% minder en de kleine deeltijders 0% minder dan de voltijders. Bij de mannen is dit 6,5% minder voor de grote deeltijders en 3,6% minder voor de kleine deeltijders. Op basis van hun analyses op de SEP-data stellen Bakker *et al* (1999) vast dat noch het uurloon van vrouwen noch dat van mannen significant wordt beïnvloed door het aantal uren dat zij werken. Op basis van hun analyses op de Vrouwenloonwijzer concluderen Wetzels en Tijdens (2002) dat het aantal arbeidsuren een zeer klein negatief effect heeft op het uurloon van vrouwen. In hun onderzoek naar de lonen in de overheidssector stellen Rienstra en Hagoort (2006) vast dat bij vrouwen het verschil tussen uurlonen in deeltijd banen (12 uur en meer) en in voltijd banen niet significant is, terwijl dit bij mannen 7 procent lager is.

Op basis van data van het *European Communities Household Panel* (ECHP) heeft Hu (2005) de effecten van het aantal gewerkte uren voor kleine deeltijders (12–21 uur per week), grote deeltijders (22–29) en voltijders (30 uur of meer) onderzocht. Na controle voor endogeniteit (de hoogte van het loon hangt niet alleen af van de arbeidsduur van de baan, maar de hoogte van het loon bepaalt ook of een werknemer fulltime of parttime wil werken) blijkt uit de decompositie

dat het beloningsverschil tussen kleine deeltijd en voltijd ongeveer 11 procent bedraagt, het verschil tussen kleine deeltijd en grote deeltijd ongeveer 7.25 procent en tussen grote deeltijd en voltijd ongeveer 3 procent. Deze data zijn niet verder uitgesplitst naar geslacht.

5.3.3 Psychologische factoren

Invloed van fair loon

Op basis van een vragenlijst onderzoekt Van Silfhout (2000) bij tien bedrijven de opvattingen van werknemers over loonongelijkheid en beloning op basis van zwaarte van de functie of van prestatie. Mannen hebben vaker opvattingen ten gunste van een grotere loonongelijkheid en ten gunste van beloning naar zwaarte en verdienste dan vrouwen. Als echter gecontroleerd wordt voor overige factoren blijkt het sekse-effect kleiner te worden en met name de invloed van de eigen sociale positie groter. Werknemers in minder gunstige posities vinden dat er meer gelijkheid in beloning behoort te zijn.

Er is weinig onderzoek gedaan of vrouwen anders onderhandelen over hun salaris dan mannen. Dit lijkt inderdaad het geval te zijn, al blijkt uit Berkhout & van der Werff (2008) dat dit effect onder hoogopgeleide starters duidelijk significant maar ook in omvang beperkt is.

Functieclassificatie

Veldman *et al* (2002) gebruiken een methode van paarsgewijze loonvergelijking tussen twee of meer vergelijkbare werknemers binnen een bedrijf. Bij drie bedrijven, een productiebedrijf, een gemeente, en een uitvaartbedrijf, hebben zij het bruto functieloon per maand van vergelijkbare werknemers vergeleken, waarbij in totaal 41 paarsgewijze loonvergelijkingen zijn uitgevoerd om discriminatie op grond van geslacht, ras of arbeidsduur te achterhalen. Lonen van deeltijders zijn daarbij omgerekend naar een voltijdsdienstverband. Bij de 18 loonvergelijkingen tussen de mannelijke en vrouwelijke werknemers varieert het verschil in functieloon van 1,2% tot 26,1%, waarbij in 13 van de 18 gevallen het vermoeden van discriminatie bestaat, waarbij 11 van de 13 ten nadele van vrouwen. Het beloningsverschil tussen de allochtone en autochtone werknemers binnen een paar varieert van 0,7% tot 18,6%, waarbij in 6 van de 11 cases een vermoeden van discriminatie bestaat, alle ten nadele van de allochtonen. Tussen de deeltijders en voltijders gaat het om verschillen van 0,9% olopend tot 13,4%. Op basis van de 11 loonvergelijkingen naar arbeidsduur worden wel loonverschillen aangetroffen die niet verklaarbaar zijn door de toegepaste beloningsmaatstaven, maar is het niet aantoonbaar dat dit onderscheid tot een benadeling op grond van arbeidsduur leidt. De verschillen in functieloon ontstaan vooral omdat bedrijven bij nieuwe medewerkers het eerste periodiek ad hoc toekennen. Datzelfde geldt voor de toekenning van anciënniteitperiodieken. Bij een bedrijf was inschaling in de lagere aanloopschaal -of nog een schaal lager- eerder regel dan uitzondering. Dat geldt ook wanneer de werknemer beschikt over de voor de functie vereiste opleiding en ervaring.

Glazen plafond

In Nederland zijn enkele onderzoeken gedaan naar het bestaan van het glazen plafond. Net als Albrecht *et al* (2003) onderzoekt Van Vuuren (2004) of de loonverschillen m-v groter zijn in de hogere strata van de loonverdeling. Op grond van het OSA aanbodpanel 1992 en 1998 constateert hij dat er inderdaad sprake is van een glazen-plafondeffect. De loonverschillen zijn het gevolg van zowel selectie-effecten als verschillen in karakteristieken, waarbij het effect van de laatste het grootste is. Fouarge *et al* (2004) concluderen op grond van het OSA aanbodpanel dat

vrouwen in het bovenste en in het onderste kwartiel een significant lager uurloon hebben dan mannen, indien gecontroleerd wordt voor overige variabelen. Gautier en Zijl (1998) schatten op basis van AVO1995 voor mannen en vrouwen afzonderlijk het functieniveau gebaseerd op een ordered logit model. Hoog opgeleide vrouwen hebben gemiddeld een twee tot drie maal zo grote kans om op een laag functieniveau terecht te komen en een ruim drie maal zo kleine kans om een leidinggevende baan te bezetten dan mannen. Ze concluderen dat het glazen plafond lijkt nog weinig barsten te vertonen.

5.3.4 Sociologische factoren

Invloed van gezinsopbouw en loopbaanonderbreking

In de vorige paragraaf is ingegaan op de 'child penalty', dat wil zeggen dat vrouwen met kinderen gemiddeld minder verdienen dan vergelijkbare mannen. In dit hoofdstuk gaan we in op Nederlandse resultaten. Daarbij kunnen we helaas niet verwijzen naar de analyses van de Arbeidsinspectie op de AVO-bestanden en niet naar de analyses van het CBS op de EWL-bestanden, omdat deze geen gegevens bevatten de huishoudenskenmerken van de werknemers. Gelukkig bevatten het SEP, de OSA en de Loonwijzer wel deze gegevens.

Uit de analyses op het SEP blijkt dat het ontberen van een partner voor mannen wel maar voor vrouwen niet financieel nadelig is en dat het hebben van kinderen voor vrouwen wel maar voor mannen niet financieel nadelig is (Bakker *et al.*, 1999). Uit de Loonwijzer-data komen voor thuiswonende kinderen vergelijkbare bevindingen (Tijdens *et al.*, 2002). Een aanzienlijk deel van de vrouwen met kinderen heeft haar loopbaan enige tijd onderbroken, en dat heeft een negatief effect op hun uurloon. Het onderzoek naar de effecten van loopbaanonderbreking hinkt op twee gedachten. Mannen en vrouwen kunnen in dit opzicht vergeleken worden, maar dan doet zich het probleem voor dat mannen vrijwel geen onderbrekingen hebben in verband met de zorg voor kinderen, maar wel veel vaker in verband met werkloosheid. De onderbrekingsduur voor de zorg voor kinderen is ook nog eens gemiddeld veel langer dan de die vanwege werkloosheid. Bakker *et al.* (1999) concluderen dat loopbaanonderbreking wegens terugtrekking van de arbeidsmarkt in financiële zin negatief uitpakt voor vrouwen en niet voor mannen. Daarentegen pakt loopbaanonderbreking wegens werkloosheid financieel negatief uit voor mannen en niet voor vrouwen.

Uit analyses op de Vrouwenloonwijzer 2000/01 blijkt dat, na controle voor overige factoren, het pure feit van herintreding een sterk negatief effect (-0.084) heeft op het uurloon (Wetzels en Tijdens 2002). De loonvergelijking is ook geschat voor alleen de herintreedsters, want dan kan gekeken worden naar het effect van de onderbrekingsduur. Elk extra onderbrekingsjaar heeft een negatief effect op het loon (-0.009). Dat effect is bijvoorbeeld groter dan het positieve effect van elk extra dienstjaar.

Tot nu toe hebben we geen aandacht geschonken aan beloningsverschillen tussen allochtone en autochtone werknemers. Hier lijkt het de aangewezen plek om daar toch iets over te zeggen, omdat de studies van de Arbeidsinspectie (Hoeben en Venema, 2004; Erdem *et al.*, 2006) en van het CBS (Rienstra en Hagoort, 2006) uitgebreid aandacht besteden aan dit onderwerp.²⁷ Volgens

²⁷ Het is niet helemaal duidelijk in hoeverre in deze studies dezelfde definitie van allochtoon wordt gebruikt. Het CBS onderscheidt autochtonen, westerse allochtonen en niet-westerse allochtonen. De Arbeidsinspectie sluit in haar rapporten aan bij de definitie van de Wet Samen. Hier gaan we ervan uit dat de allochtonen volgens de Arbeidsinspectie overeenkomen met de niet-westerse allochtonen van het CBS.

de Arbeidsinspectie verdienen in 2004 in het bedrijfsleven allochtone werknemers in het bedrijfsleven, na correctie voor overige factoren, gemiddeld 4,8% minder dan hun autochtone collega's (Erdem *et al*, 2006, 176). Voor vrouwen is dit verschil kleiner dan voor mannen: 3,8% respectievelijk 5,1%. Kijken we alleen naar de groep allochtonen, dan is het beloningsverschil m-v 4,6%. Dit is aanmerkelijk kleiner dan het beloningsverschil m-v binnen de groep autochtonen, waar het 7,7% is. In het onderzoek op basis van de EWL naar de overheidssector is bij de autochtonen en westerse allochtonen het gecorrigeerde beloningsverschil tussen mannen en vrouwen ruim 10 procent, terwijl er in de subpopulatie niet-westerse allochtonen geen significant verschil in uurloon tussen mannen en vrouwen is. In tegenstelling tot de subpopulaties autochtonen en westerse allochtonen is voor niet-westerse allochtonen de variabele leeftijd bijna niet van invloed op het uurloon. (Rienstra en Hagoort, 2006).

Invloed van beroep

De beloning in vrouwenberoepen is gemiddeld lager dan in mannenberoepen. De segregatie naar beroep en sector blijkt een belangrijke oorzaak van de loonverschillen m-v. Er zijn verschillende methoden om beroep op te nemen in statistische analyses. Het meest gebruikt is een sterk geaggregeerde indeling van beroepen op 1-digit niveau, waardoor het looneffect voor elke beroepsgroep (minus de referentiegroep) afzonderlijk kan worden onderzocht. Een tweede mogelijkheid is om het aandeel vrouwen in een beroep op te nemen, maar omdat de grote meerderheid van beroepen dan wel een heel groot dan wel een heel klein aandeel vrouwelijke beroepsbeoefenaren heeft, is deze variabele zeer scheef verdeeld naar de twee uiteinden. Een oplossing is schaling van het aandeel vrouwen (Bakker *et al*, 1999) of een categorisering met dummyvariabelen: een variabele 'neutraal beroep' en een variabele 'door vrouwen gedomineerd beroep' (De Ruijter, 2002). De Ruijter gebruikt ook een derde variant. Ze construeert de absolute concentratiegraad van vrouwen in een beroep, waarin het aantal vrouwen in een beroep gerelateerd is aan het totaal aantal vrouwen in de beroepsbevolking. Als deze absolute concentratiegraad vervolgens gedeeld wordt door de absolute concentratiegraad van alle werkzame personen ontstaat de relatieve concentratiegraad van vrouwen in beroepen: de mate van concentratie van de vrouwelijke beroepsbevolking in een beroep, gerelateerd aan de omvang van dat beroep.

Bakker *et al* (1999) concluderen dat de beloning in vrouwenberoepen lager is, maar dat er binnen beroepen nauwelijks beloningsverschillen zijn tussen mannen en vrouwen. De sekse-compositie van beroepen draagt substantieel bij aan het beloningsverschil. Als de vrouwelijke en de mannelijke beroepsbevolking gelijk zou zijn naar functieniveau, leidinggevende taken en sekse-compositie van het beroep, en de opbrengsten van deze factoren gelijk zouden zijn geweest, dan zouden de loonverschillen m-v geheel zijn verdwenen.

Kooreman (2007) concludeert uit de Nationaal Scholierenonderzoeken 1984-2001 dat in vier 'job types' (werken in winkel of supermarkt, baby-sitting, kranten rondbrengen, en werken in een restaurant of café) jongens meer verdienen dan meisjes. Dat komt vooral omdat het baan type een grote invloed op de loonverschillen m-v heeft, vooral omdat meisjes veel meer in de baan types met de laagste lonen werken, en dan vooral babysitten. Hij constateert dat er binnen de baan types geen loonverschillen m-v zijn, en concludeert hieruit dat er geen steun is voor de veronderstelling dan meisjes minder assertief zouden onderhandelen over hun beloning dan jongens. Ook doen het aantal arbeidsuren en de participatie van jongens en meisjes niet ter zake. Bovendien constateert hij dat tussen 1994 en 2001 het loon van het laagst betaalde job type

relatief is achtergebleven bij dat van de andere job types. Gedurende de hele periode 1984-2001 bleef de sekse-compositie van de vier baan types vrijwel onveranderd.

Ook De Ruijter (2002) vindt in haar onderzoek op de Structure of Earnings data 1997 van het CBS dat mannen in alle beroepen significant hogere lonen verdienen dan vrouwen. Gemiddeld verdienen vrouwen maar 85% van het loon van mannen in hetzelfde beroep. Gecontroleerd voor menselijk kapitaal verdienen vrouwen in mannenberoepen gemiddeld 89% van het gemiddelde loon, in vrouwenberoepen is dit percentage 94%. Het loonverschil is daarmee het grootst in mannenberoepen en het kleinst in vrouwenberoepen. Interessant is dat beroepen gemeten zijn met het begrip concentratiegraad. Dit is het percentage vrouwen in het beroep, opgedeeld in tien percentielen, waarbij deze gegevens zijn gehaald uit de EBB data van het CBS. Mannen en vrouwen in vrouwenberoepen verdienen slechts 93% van het loon van mannen en vrouwen in mannenberoepen. Ongeveer éénvijfde van dit verschil kan worden verklaard door verschillen in de samenstelling van deze beroepen wat betreft menselijk kapitaal. Mannen en vrouwen in mannenberoepen beschikken over meer menselijk kapitaal dan mannen en vrouwen in vrouwenberoepen. Bovendien is het vereiste opleidingsniveau van vrouwenberoepen lager dan dat in mannenberoepen. De opbrengsten van het vereiste opleidingsniveau zijn in vrouwenberoepen lager dan in mannenberoepen.

De invloed van de sekse-compositie van de werkplek is in Nederland veel minder onderzocht dan die van sekse-compositie van beroepen. Alleen de Loonwijzer bevat een survey vraag over de sekse-compositie van de afdeling. Uit analyses op de Vrouwenloonwijzer 2000/01 blijkt dat als vrouwen merendeels mannelijke collega's hebben, dat een substantieel positief effect op hun loon heeft (Wetzels en Tijdens, 2002). Vrouwen zonder loopbaanonderbreking profiteren in financiële termen nog meer van het werken met merendeels mannelijke collega's dan herintreedsters. Tijdens *et al* (2002) vinden in een decompositie van de loonvergelijking op basis van de Loonwijzer data dat vrouwen in vergelijking met mannen een positief effect op hun loon ondervinden van het werken met merendeels mannelijke collega's.

Invloed van functieniveau, leidinggeven en hiërarchische structuur

Indelingen naar beroep vormen veelal de basis voor een groepering in 8 functieniveaus zoals CBS en Arbeidsinspectie hanteren. Vrouwen vervullen vaker een functie van laag niveau dan mannen; in de drie laagste onderscheiden niveaus zit 49% van de vrouwen en 34% van de mannen. Op de drie hoogste onderscheiden niveaus bevindt zich 10% van de vrouwen en 25% van de mannen (Erdem *et al*, 2006). De analyses van de AVO-data van 2002 en van 2004 laten zien dat een hoger functieniveau tot een hoger uurloon leidt, voor mannen en voor vrouwen. Op de meeste functieniveaus zijn de opbrengsten van functieniveau hoger voor mannen dan voor vrouwen, alleen op functieniveau III en op de hoogste functieniveaus, VII en VIII, verdienen vrouwen een hoger loon dan mannen, alle overige kenmerken constant gehouden. Bakker *et al* (1999) constateren eveneens dat vrouwen als ze op hetzelfde functieniveau werken als mannen iets minder verdienen. In deze analyse is niet naar elk niveau afzonderlijk gekeken, en kunnen dus geen uitspraken per niveau worden gedaan.

Onder werknemers in banen met een elementair of lager beroepsniveau in de overheidssector bestaat geen loonverschil m-v, terwijl in banen met een middelbaar beroepsniveau het uurloon van vrouwen, na correctie voor andere variabelen, 17 procent lager is dan dat van mannen. Voor werknemers in banen met een hoger of wetenschappelijk beroepsniveau is het uurloon van

vrouwen met 13 procent ook aanzienlijk lager dan het uurloon van mannen. Vrouwen in de overheidssector hebben, na controle voor overige variabelen, in hogere beroepen een uurloon dat nauwelijks verschilt van dat in wetenschappelijke beroepen; voor mannen in de overheidssector is het uurloon in wetenschappelijke beroepen echter beduidend hoger dan in hogere beroepen (Rienstra en Hagoort, 2006).

Bakker *et al* (1999) concluderen dat mannen vaker dan vrouwen een leidinggevende positie hebben én als vrouwen al een leidinggevende positie hebben, zij daar financieel veel minder profijt van hebben dan mannen. Voor vrouwen loont het, in letterlijke zin, niet om een leidinggevende positie te vervullen. De Ruijter (2002) komt tot eenzelfde conclusie: zowel op individueel niveau als op beroepsniveau wordt in vrouwenberoepen leidinggeven minder beloond dan in mannenberoepen.

5.4 Conclusie onderzoek naar de loonkloof in Nederland

De Nederlandse onderzoeken overziend moeten we constateren dat de economische effecten goed in kaart gebracht zijn, met uitzondering van het effect van sector. Beloningsverschillen tussen sectoren worden in Nederland geïnstitutionaliseerd door Cao's, maar de relatie tussen de institutionele beloningsstructuur en de loonverschillen m-v zijn tot nu toe eigenlijk niet onderzocht. Een beter inzicht in de sekse-compositie van sectoren en de invloed daarvan op het loonverschil zal het onverklaarde deel daarvan zeker doen verminderen. Ook effecten van sekse-specifieke productiviteitsverschillen en verschillen ziekteverzuim zijn weinig onderzocht.

Kijken we naar de psychologische effecten, dan zijn de effecten van mogelijke sekse-verschillen in individuele loononderhandelingen tussen werkgever en werknemer nog vrijwel niet onderzocht. Een probleem hierbij is de meting van de onderhandelingsruimte, vooral betreffende de secundaire arbeidsvoorwaarden. Het design van een dergelijk onderzoek vereist inzicht in personeelsbestanden van bedrijven en in personeelsdossiers. Op voorhand kan geen inschatting gemaakt worden van het mogelijke effect hiervan op de verkleining van het onverklaarde deel van het loonverschil m-v.

Kijken we naar de sociologische effecten, dan zijn de effecten van de beroepensegregatie goed onderzocht, maar die van de sekse-compositie en de hiërarchische structuur van organisaties veel minder. Ook zouden de gedragsbeslissingen die samenhangen met de opeenvolgende gezinsfasen en het effect daarvan op het loonverschil m-v wat beter onderzocht kunnen worden. Gedrags- en allocatieverschillen laten na langere tijd grote effecten zien. Een beter begrip van het loonverschil m-v in de oudere leeftijdsgroepen zou daarom kunnen bijdragen aan een verkleining van het onverklaarde deel.

Literatuur hoofdstuk 5

- Albrecht, J., A. Bjorklund, S. Vroman (2003) Is There a Glass Ceiling in Sweden? *Journal of Labor Economics*, Vol. 21, No. 1 (January 2003)
- Albrecht, J., A. van Vuuren, S. Vroman (2004) *Decomposing the Gender Wage Gap in the Netherlands with Sample Selection Adjustments*. Bonn: IZA, Discussion Paper 1400
- Arulampalam, W., A. L. Booth, M. L. Bryan (2007) Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution, *Industrial and Labor Relations Review*, Vol. 60, No. 2 (January 2007).
- Bakker, B.F.M., K.G. Tijdens, J.W. Winkels (1999) Gender, occupational segregation and wages in the Netherlands, *Netherlands Official Statistics*, 14, Winter, 36-41
- Berkhout, E., A. Heyma, W. Salverda (2006). *Beloningsverschillen tussen de marktsector en collectieve sector in 2004*. Amsterdam: SEO Economisch Onderzoek, SEO rapport nr. 889
- Berkhout, E., M. Biermans, W. Salverda, K. Tijdens (2007a) *Internationale beloningsverschillen van wetenschappelijke personeel*. Amsterdam: SEO Economisch Onderzoek, SEO-rapport nr. 981
- Berkhout, E.E., P.H.G. Berkhout, M.L. Biermans (2007b) *Studie & Werk 2006*. Amsterdam: SEO Economisch Onderzoek, 2007. SEO rapport nr. 902
- Berkhout, E.E., S.G. van der Werff (2008) *Studie & Werk 2008*. Amsterdam: SEO Economisch Onderzoek, 2008. SEO rapport nr. 2008-28.
- Blinder, A.S. (1973) Wage discrimination: reduced form and structural estimates. *Journal of Human Resources*, 8(4): 436-455
- Brouns, M., L. Halsema, J. de Bruijn (1996) *Waardering van werk. Opstellen over functiewaardering en sekse*. Amsterdam: VU Uitgeverij
- Eijl, C. van, K. Tijdens, J. Schippers (1998) 'Wie heeft het recht mij mijn broodwinning te ontnemen'. Arbeidsloon, arbeidsomstandigheden en arbeidstijden, in Pott-Buter, H., K. Tijdens (red.) *Vrouwen - leven en werk in de twintigste eeuw*. Amsterdam: Amsterdam University Press, 179-195
- England, P. (1992). *Comparable Worth*. New York: Aldine De Gruyter.
- Erdem, O., M.A. Bos, J. Hoeben (2006) *De arbeidsmarktpositie van werknemers in 2004. Verschillen in beloning en mobiliteit*. Den Haag: Arbeidsinspectie
- European Commission (2003) *Gender pay gaps in European labour markets. Measurement, analyses and policy implications*. Brussels: SEC (2003)937, 4.9.2003
- European Commission (2006a) *Mededeling. Een routekaart voor de gelijkheid van mannen en vrouwen*. Brussel: COM (2006)92def, 01.3.2006

- European Commission (2006b) *Commission Staff Working Document. Annex to the Communication A roadmap for equality between women and men 2006-2010. Impact Assessment*. Brussels: COM(2006)92final, 01.3.2006
- European Commission (2007a) *Bridging the Gender Pay Gap*. Brussels: MeMo/07/297, 18.7.2007
- European Commission (2007b) Communication. Tackling the gender pay gap between women and men. Brussels: COM (2007)XXX
- Fields, J., E. N. Wolff (1995). Interindustry wage differentials and the gender wage gap. *Industrial and Labor Relations Review*. 49: 105-120
- Fouarge, D., R. Grim, M. Kerkhofs, A. Roman, T. Wilthagen (2004) *Tendrapport Aanbod van arbeid 2003*. Tilburg, OSA
- Gannon, B., R. Plasman, F. Rycx, I. Tojerow (2007) Inter-industry Wage Differentials and the Gender Wage Gap: Evidence from European Countries, *Economic and Social Review*, 38(1): 135-55
- Gautier, P.A., M. Zijl (1998) Het glazen plafond zichtbaar gemaakt. *Economisch Statistische Berichten*, 83e jaargang, nr. 4168, 724
- Groshen, E. L. (1991) The Structure of the Female/Male Wage Differential. *Journal of Human Resources*. 26: 457-472
- Heyma, A., E. Berkhout, W. Salverda, M. Biermans (2004) *Beloningsverschillen tussen de marktsector en collectieve sector in 2001*. Amsterdam: SEO Economisch Onderzoek
- Heyma, A., K. Tijdens, E. Berkhout, L. Janssens (2005) *De aantrekkelijkheid van de collectieve sector als werkgever: een vergelijking van beloningsprofielen tussen collectieve sector en marktsector*. Amsterdam: SEO Economisch Onderzoek, SEO rapport nr. 797
- Hoeben, J., P.M. Venema (2004) *De arbeidsmarktpositie van werknemers in 2002. Een onderzoek naar de verschillen in beloning en mobiliteit tussen groepen werknemers*. Den Haag: Arbeidsinspectie
- Hogue, M., J.D. Yoder, S.B. Singleton (2007) The Gender Wage Gap: An Explanation of Men's Elevated Wage Entitlement, *Sex Roles: A Journal of Research*, 56(9-10): 573-579
- Hu, Y. (2005) *Essays on Labour Economics: Empirical Studies on Wage Differentials across Categories of Working Hours, Employment Contracts, Gender and Cohorts*. Amsterdam: Tinbergen Institute / University of Amsterdam
- Jones, D.R., G.H. Makepeace (1996) Equal Worth, Equal Opportunities: Pay and Promotion in an Internal Labour Market, *Economic Journal*, March, 106 (435), 401-09
- Kalmijn, M., T. van der Lippe (1997). Type of Schooling and Sex Differences in Earnings in the Netherlands. *European Sociological Review*. 13: 1-15

- Klaveren, M. van, W. Sprenger, K. Tijdens (2007). *Dicht de loonkloof!! Verslag van het CLOSE (Correctie LOonkloof in SEctoren) onderzoek voor de FNV, ABVAKABO FNV en FNV Bondgenoten*. Eindhoven, STZ advies & onderzoek/Amsterdam, Universiteit van Amsterdam/AIAS
- Kooreman, P. (2007) *The Early Inception of Labor Market Gender Differences*. Tilburg University, revised paper
- Kumlin, J. (2007) The Sex Wage Gap in Japan and Sweden: The Role of Human Capital, Workplace Sex Composition, and Family Responsibility, *European Sociological Review*, 23, 2, 203-22
- Meer, Peter van der, 2005, Is the Gender Pay Gap declining in the Netherlands?
- Oaxaca, R. (1973) Male-Female wage differentials in urban labour markets, *International Economic Review*, 14, 693-709
- OECD (2005) *Employment Outlook*. Paris: OECD
- Paul, M. (2006) A Cross-Section Analysis of the Fairness-of-Pay Perception of UK Employees. *The Journal of Socio-Economics*, 35(2): 243-267
- Petersen, T., L. A. Morgan (1995) Separate and unequal, occupational-establishment sex segregation and the gender wage gap. *American Journal of Sociology*. 101: 329-365
- Plantenga, J., C. Remery (2006) The Gender Pay Gap – Origins and policy responses. A comparative review of 30 European countries. Brussels: European Commission
- Price Waterhouse Consulting (PWC) (2002) *Gelijke monniken (m-v), gelijke kappen? Resultaten van een kwantitatief onderzoek naar gelijke beloning van mannen en vrouwen bij variabele beloning*. Almere: PWC Consulting B.V.
- Rienstra, M., K. Hagoort (2006) *Beloningsverschillen verklaard? Verschillen in uurloon bij de overheid, 2004*. Voorburg/Heerlen: CBS, Centrum voor Beleidsstatistiek
- Ruijter, J. de (2002) *Occupational Wage differences: A Gender Approach*. Rotterdam: Erasmus University Rotterdam, dissertation
- Silfhout, R. van(2000) *Inequality in pay within organizations: normative and instrumental perspectives*. Tilburg University, dissertation
- Simon, H., H. Russell (2004) *Firms and the Gender Pay Gap: A Cross-National Comparison*. London: School of Economics, PIEP Working Paper.
- Spijkerman, R. (2000) *De positie van mannen en vrouwen in het bedrijfsleven en bij de overheid 1998*. Den Haag: Arbeidsinspectie Ministerie van Sociale Zaken en Werkgelegenheid
- Tam, T. (1997). Sex segregation and occupational gender inequality in the United States: Devaluation or specialized training? *American Journal of Sociology*. 102:1652-92

- Tegelaar, P.J.C. (1972) Gelijke beloning voor mannen en vrouwen, *Mens en Onderneming*, 26(3): 150-166
- Tijdens, K.G. (2001) Beloningsongelijkheid en beloningsverschil, *Zeggenschap*, 12(1): 35-38
- Tijdens, K.G., A. Dragstra, D. Dragstra, M. van Klaveren, P. Osse, C.M.M.P. Wetzels, A. Zorlu (2002) *Loonwijzers 2001/2002. Werk, lonen en beroepen van mannen en vrouwen in Nederland*. Amsterdam University of Amsterdam, AIAS Research Report RR02/10
- Veldman, A.G., S. Schalkwijk, m.m.v. A. van Doorne-Huiskes (2002) *Gelijke beloning naar sekse, ras en arbeidsduur op ondernemingsniveau. Een methode voor onderzoek naar de naleving van de Nederlandse gelijkebeloningswetgeving*. Onderzoek verricht in opdracht van het Ministerie van Sociale Zaken en Werkgelegenheid door de Universiteit Utrecht, sectie Arbeidsrecht en Sociaal Beleid
- Vuuren, A. van (2004) Hogere participatie van vrouwen leidt tot grotere beloningsverschillen, *Economisch Statistische Berichten*, 89: 537-539
- Wetzels, C.M.M.P., K.G. Tijdens (2002) Women's career break due to motherhood and the effects on wages. *Cahiers Economiques de Bruxelles* 45 (1), 169-189

6 Alternatieve verklaringen: een empirische verkenning

Naast een gerichte inventarisatie van de wetenschappelijke literatuur (kwalitatief) is het ook zinvol om voor de Nederlandse situatie te onderzoeken in hoeverre het aanvankelijke gecorrigeerde loonverschil nog verder ontleed kan worden met behulp van extra variabelen (kwantitatief). Is het ‘onverklaarde loonverschil’ bijvoorbeeld te verklaren door het feit dat vrouwen meer belang hechten aan bepaalde subjectieve factoren (zoals de combinatie werk-gezin, leuke collega’s, reistijd) of functiekenmerken (zoals afwisselend werk, ploegendienst)? Nemen ze in functies met afwisselend werk en weinig reistijd genoegen met minder salaris, omdat de compenserende waarde van het afwisselende werk of de geringe reistijd voor hen belangrijker is?

Dergelijke vragen zijn helaas niet te beantwoorden met het AVO-bestand, omdat daarin geen subjectieve factoren aanwezig zijn. Maar voor de analyse van verdere significante verklaringen van seksegerelateerde loonverschillen zijn de antwoorden op bovenstaande vragen wel degelijk van belang. Gelukkig hebben we dergelijke variabelen wel in de Loonwijzer dataset en kunnen we dus die informatie aan het model toevoegen, om zodoende het onverklaarde deel van het loonverschil verder te verklaren en (waarschijnlijk) kleiner te maken. Basis van deze analyse is een herschatting van het model uit het Arbeidsinspectie-rapport op de Loonwijzer (paragraaf 6.2). Omdat de Loonwijzer met een ander doel en volgens een andere methodiek is opgezet zal hieruit in absolute termen waarschijnlijk een ander gecorrigeerd loonverschil tevoorschijn komen.²⁸ Maar dit absolute verschil nemen we als gegeven, om ons vervolgens te richten op de verdere uitsplitsing van het ‘onverklaarde’ deel. De focus is hier dus niet meer op het absolute percentage loonverschil man-vrouw maar op het gedeelte onverklaard loonverschil dat alsnog verklaard kan worden met behulp van nieuwe variabelen.

6.1 Data Loonwijzer

Sinds 2000 is de Loonwijzer een website waar informatie wordt gegeven over lonen en arbeidsvoorwaarden. Bezoekers worden uitgenodigd om een vragenlijst in te vullen over de eigen beloning, op basis waarvan informatie in de Loonwijzer wordt geüpdate. Ongeveer 1 op de 100 bezoekers vult de vragenlijst ook daadwerkelijk in.

De vragenlijst bevat vragen over het beroep, het bedrijf en de opleiding, de werkplek, de arbeidsgeschiedenis, lonen en de persoonlijke situatie. Ondertussen is de Loonwijzer uitgegroeid tot een permanent onderzoek op internet, met een voortdurend groeiend bestand met gedetailleerde informatie over lonen, bonussen en subjectieve gegevens over carrièreperspectieven en loopbaanontwikkelingen in alle sectoren van de Nederlandse economie. Het gegevensbestand telt inmiddels meer dan 160 duizend geldige observaties van uurlonen,

²⁸ Belangrijk methodisch verschil is dat de Loonwijzer geen steekproefkader heeft, maar gebaseerd is op zelfselectie van respondenten, via Internet. Daarnaast zijn voor de overheid dezelfde variabelen beschikbaar als voor het bedrijfsleven, dat is in de AVO-dataset niet zo. Ook is de overheid hier in principe volledig vertegenwoordigd, dus ook de subsectoren die in de AVO-data ontbreken. Tot slot zijn enkele variabelen anders geoperationaliseerd, met andere definities en antwoordcategorieën. Het is dan niet zinvol om de absolute loonverschillen te analyseren, daarvoor is waarschijnlijk de AVO-dataset geschikter. We kijken echter naar de relatieve bijdrage van de subjectieve factoren in de totale verklaring van het loonverschil.

gemeten over de periode 2000-2006. Vanwege de beperkte beschikbaarheid van enkele relevante variabelen is echter de eerste release niet bruikbaar, en moeten we ons derhalve richten op het vierde kwartaal 2004 tot en met het vierde kwartaal 2006. Daarbij resteren nog altijd ruim 93 duizend waarnemingen.

Tabel 6.1 Aantal waarnemingen uurloon in Loonwijzer

	Frequentie
2000	12.821
2001	14.047
2002	11.217
2003	12.822
2004	26.823
2005	47.842
2006	37.118
Totaal 2000 - 2006	162.690
Totaal vierde kwartaal 2004 – vierde kwartaal 2006	93.334

In de Loonwijzer dataset wordt, naast de waardering van het loon, de waardering van een flink aantal arbeidsvoorwaarden en functiekenmerken gevraagd. Behalve deze subjectieve looncompenserende factoren (o.a. belang reistijd & -afstand, deeltijdmogelijkheid, vast contract, opleidingsmogelijkheden, hoog salaris, leuke collega's etc.) is er in de Loonwijzer ook objectieve informatie beschikbaar over de huishoudensituatie die van belang kan zijn. Immers, voor alleenstaande vrouwen (en mannen) zal het inkomensmotief belangrijker zijn dan voor vrouwen (en mannen) met een werkende partner. Het belang van een hoog salaris is wellicht anders als er in het huishouden al iemand anders is met een inkomen. Ook de aanwezigheid van jonge kinderen is waarschijnlijk van belang: de 'tweede' partner accepteert wellicht een baan met een lager salaris als hij/zij daardoor thuis kan zijn wanneer de jonge kinderen dat ook zijn. Immers, het alternatief is het betalen van kinderopvang of huishoudelijke hulp waardoor van het hoge salaris minder overblijft. Of de tweede partner zoekt juist naar een baan met een hoog salaris, omdat werken ook betekent dat er hogere kosten voor kinderopvang gemaakt moeten worden. Het hebben van betaalde huishoudelijke hulp kan dus een reden zijn om alleen een baan met een hoog salaris te accepteren. Helaas blijkt de Loonwijzer wel gegevens over allerlei redenen van loopbaanonderbreking te bevatten, maar helaas alleen voor de eerste jaargang en dus niet voor de jaargangen die wij in onze analyse gebruiken (waar juist weer andere relevante variabelen aanwezig zijn).

6.2 Replicatie basisloonschatting

In deze paragraaf voeren we een soortgelijke loonschatting uit als in paragraaf 3.3 op basis van de Loonwijzer data. In deze loonschatting is het bruto uurloon de te verklaren variabele. De volgende kenmerken zullen in de loonvergelijking worden gebruikt als verklarende variabelen voor zowel het bedrijfsleven als de overheidssector.

- geslacht
- leeftijd
- dienstjaren bij huidige werkgever

- arbeidsduur
- permanent contract
- uitzendkracht
- opleidingsniveau (ISCED)
- beroepsniveau
- functiesoort
- economische sector
- bedrijfsgrootte
- regio
- etnische afkomst
- cao aanwezig
- jaardummies

In Bijlage A wordt uiteengezet hoe de bovenstaande variabelen zijn gedefinieerd zijn en welke categorieën er worden onderscheiden. We benadrukken dat de Loonwijzer met een ander doel en volgens een andere methodiek is opgezet en dat we als gevolg hiervan een ander gecorrigeerd loonverschil vinden. Echter, dit absolute verschil nemen we als gegeven. Het belangrijkste doel van deze paragraaf is om een referentiekader te stellen d.m.v. een basisloonschatting. In dit referentiekader bepalen we het loonverschil tussen mannen en vrouwen. Vervolgens toetsen we voor verschillende aanvullende factoren of deze een significantie bijdrage leveren aan de variatie van het uurloon en of het gemeten loonverschil m-v hierdoor verandert. Tabel 6.2 en Tabel 6. bevatten de resultaten van de basisloonschatting voor respectievelijk het bedrijfsleven en de overheidssector. Het loonverschil m-v in het bedrijfsleven is 14.5 procent, in de overheidssector 11 procent.

Tabel 6.2 Basisloonschatting voor het bedrijfsleven, OLS, o.b.v. Loonwijzer

	Coëfficiënt ^a	Standaardfout	95% B.I	
Man	<i>referentie</i>			
Vrouw	-0,145	0,003	-0,150	-0,140
Leeftijd	0,050	0,001	0,048	0,051
Leeftijd kwadraat	-0,000	0,000	0,000	0,000
Dienstjaren	0,003	0,000	0,002	0,004
Dienstjaren kwadraat	-0,000	0,000	0,000	0,000
Voltijd	<i>referentie</i>			
Deeltijd, 12 uur of meer	0,060	0,003		
Deeltijd, minder dan 12 uur	0,125	0,012		
Permanent contract	0,095	0,003	0,090	0,101
Uitzendkracht	-0,058	0,005	-0,067	-0,049

vervolg Tabel 6.2 Basisloonschatting voor het bedrijfsleven, OLS, o.b.v. Loonwijzer

	Coëfficiënt*	Standaardfout	95% B.I	
ISCED 0-1 (Primair)	-0,129	0,008	-0,145	-0,114
ISCED 2 (Lager secundair)	-0,080	0,003	-0,086	-0,074
ISCED 3 (Hoger secundair)	0,008	0,005	-0,001	0,017
ISCED 4 (Post-secundair)	<i>referentie</i>			
ISCED 5-6 (Tertiair)	0,183	0,003	0,178	0,188
Opleiding onbekend	n.s.	0,017	-0,014	0,053
Elementaire beroepen	-0,203	0,012	-0,227	-0,178
Lagere beroepen	-0,164	0,011	-0,185	-0,142
Middelbare beroepen	-0,073	0,011	-0,094	-0,052
Hogere beroepen	0,020	0,011	-0,002	0,042
Wetenschappelijke beroepen	0,071	0,012	0,047	0,094
Onbekend	<i>referentie</i>			
Technische functie	n.s.	0,011	-0,017	0,025
Administratieve functie	0,038	0,011	0,016	0,059
Automatiseringsfunctie	n.s.	0,012	-0,009	0,036
Commerciële functie	0,071	0,011	0,050	0,093
Dienstverlenende functie	0,026	0,011	0,004	0,048
Docent functie	0,039	0,023	-0,006	0,085
Bestuurlijke functie	0,089	0,015	0,060	0,118
Leidinggevende functie	-0,051	0,031	-0,111	0,010
Landbouw en visserij	-0,035	0,008	-0,051	-0,020
Industrie	0,025	0,003	0,019	0,031
Bouwnijverheid	0,018	0,005	0,009	0,027
Handel	-0,042	0,004	-0,049	-0,034
Horeca	-0,061	0,006	-0,072	-0,050
Transport	-0,026	0,004	-0,035	-0,018
Zakelijke dienstverlening	<i>referentie</i>			
Zorg	-0,020	0,004	-0,028	-0,012
Overige diensten	-0,021	0,005		
Bedrijfsgrootte: 1-9	-0,153	0,004	-0,161	-0,146
Bedrijfsgrootte: 10-99	-0,089	0,003	-0,094	-0,083
Bedrijfsgrootte: 100-499	-0,041	0,003	-0,047	-0,035
Bedrijfsgrootte: 500 of meer	<i>referentie</i>			
Bedrijfsgrootte: onbekend	-0,088	0,031	-0,148	-0,028
Regio Noord	-0,066	0,004	-0,073	-0,059
Regio Oost	-0,043	0,003	-0,048	-0,037
Regio West	<i>referentie</i>			
Regio Zuid	-0,038	0,003	-0,043	-0,033
Regio Onbekend	-0,057	0,014	-0,085	-0,028
Autochtoon	<i>referentie</i>			
Allochtoon	-0,009	0,003	-0,015	-0,003

vervolg Tabel 6.2 Basisloonschatting voor het bedrijfsleven, OLS, o.b.v. Loonwijzer

	Coëfficiënt*	Standaardfout	95% B.I	
Etniciteit onbekend	-0,002	0,011	-0,022	0,019
CAO aanwezig	-0,017	0,002	-0,022	-0,012
2004	-0,044	0,004	-0,051	-0,037
2005	-0,061	0,002	-0,066	-0,057
2006	<i>referentie</i>			
constante	1,621	0,018	1,585	1,656
Aantal observaties	66354			
Aangepaste R-kwadraat	0,457			

*De afkorting n.s. (niet significant) geeft aan dat het kenmerk geen significante invloed heeft op het uurloon.

Tabel 6.3 Basisloonschatting voor de overheidssector, OLS, o.b.v. Loonwijzer

	Coëfficiënt*	Standaardfout	95% B.I	
Man	<i>referentie</i>			
Vrouw	-0,110	0,006	-0,121	-0,098
Leeftijd	0,031	0,002	0,027	0,036
Leeftijd Kwadraat	0,000	0,000	0,000	0,000
Dienstjaren	0,005	0,001	0,003	0,007
Dienstjaren Kwadraat	-0,000	0,000	-0,000	-0,000
Voltijd	<i>referentie</i>			
Deeltijd, 12 uur of meer	0,063			
Deeltijd, minder dan 12 uur	0,065			
Permanent contract	0,086	0,007	0,072	0,100
Uitzendkracht	-0,045	0,013	-0,070	-0,020
ISCED 0-1 (Primair)	-0,219	0,036	-0,289	-0,148
ISCED 2 (Lager secundair)	-0,060	0,009	-0,078	-0,042
ISCED 3 (Hoger secundair)	0,005	0,012	-0,017	0,028
ISCED 4 (Post-secundair)	<i>referentie</i>			
ISCED 5-6 (Tertiair)	0,228			
Opleiding onbekend	0,170			
Elementaire beroepen	-0,310	0,037	-0,383	-0,238
Lagere beroepen	-0,235	0,024	-0,281	-0,189
Middelbare beroepen	-0,154	0,022	-0,197	-0,112

vervolg Tabel 6.3 Basisloonschatting voor de overheidssector, OLS, o.b.v. Loonwijzer

	Coëfficiënt*	Standaardfout	95% B.I	
Hogere beroepen	-0,047	0,023	-0,091	-0,003
Wetenschappelijke beroepen	0,007	0,024	-0,040	0,054
Onbekend	<i>referentie</i>			
Technische functie	0,026	0,022	-0,018	0,070
Administratieve functie	0,076	0,022	0,034	0,118
Automatiseringsfunctie	0,073	0,026	0,022	0,124
Commerciële functie	0,160	0,026	0,109	0,211
Dienstverlenende functie	0,077	0,022	0,034	0,120
Docent functie	0,033	0,023	-0,011	0,078
Bestuurlijke functie	0,061	0,023	0,015	0,107
Leidinggevende functie	0,137	0,032	0,074	0,200
Onderwijs	-0,118	0,007	-0,131	-0,104
Rechterlijke Macht	-0,056	0,037	-0,128	0,016
Defensie	-0,008	0,011	-0,031	0,015
Overheid overig	<i>referentie</i>			
Bedrijfsgrootte: 0-10	-0,059	0,013	-0,084	-0,034
Bedrijfsgrootte: 10-100	-0,054	0,007	-0,068	-0,041
Bedrijfsgrootte: 100-500	-0,014	0,006	-0,026	-0,001
Bedrijfsgrootte: 500 plus	<i>referentie</i>			
Bedrijfsgrootte: onbekend	0,023	0,086	-0,146	0,192
Regio Noord	-0,031	0,009	-0,048	-0,014
Regio Oost	-0,024	0,007	-0,037	-0,010
Regio West	<i>referentie</i>			
Regio Zuid	-0,022	0,007	-0,035	-0,008
Regio Onbekend	-0,068	0,035	-0,137	0,001
Autochtoon	<i>referentie</i>			
Allochtoon	n.s.			
Etniciteit onbekend	n.s.			
CAO aanwezig	0,052	0,009	0,036	0,069
2004	-0,032	0,009	-0,049	-0,014
2005	-0,059	0,006	-0,071	-0,048
2006	<i>referentie</i>			
Constante	1,917	0,045	1,828	2,006
Aantal observaties	9365			
Aangepaste R-kwadraat	0,472			

*De afkorting n.s. (niet significant) geeft aan dat het kenmerk geen significante invloed heeft op het uurloon.

Het loonverschil m-v kunnen we ook bepalen door gebruik te maken van de eerder genoemde matching-methoden. Op basis van de standaardfout van de verschillende matching-methoden is in Hoofdstuk 4 geconcludeerd dat Radius en Kernel matching betrouwbare en efficiënte schattingen van het loonverschil m-v opleveren. Daarom zullen we in Hoofdstuk 1 alleen deze twee matching-methoden toepassen. De variabelen die gebruikt worden bij het matchen zijn dezelfde als die gebruikt worden in bovenstaande basisloonschattingen.

Tabel 6.4 Basisloonschatting, propensity score matching, o.b.v. Loonwijzer

Matching Methode	Loonverschil	Standaardfout	t-waarde
<i>Bedrijfsleven:</i>			
Radius matching (r = 0.05)	-0,132	0,005	27,9
Kernel matching (Epanechnikov)	-0,131	0,005	27,8
<i>Overheidssector</i>			
Radius matching (r = 0.05)	-0,120	0,010	12,0
Kernel matching (Epanechnikov)	-0,120	0,010	12,0

Het voordeel van de matching-methode boven OLS is gelijk reeds in de eerdere paragrafen besproken. Het geschatte loonverschil m-v valt voor het bedrijfsleven met 13.2 procent iets lager uit dan de eerder gevonden 14.5 procent bij schatting via OLS. Voor de overheidssector is het geschatte loonverschil juist iets hoger: 12 procent om 11 procent. Welke matching-methode precies gebruikt wordt maakt voor de resultaten niets uit. Echt schokkend zijn de verschillen tussen matching en OLS niet te noemen. In de volgende paragraaf zullen we de bijdrage van de aanvullende variabelen aan de verklaring van de loonkloof bepalen met behulp van de matching-resultaten.

6.3 Aanvullende verklaringen voor verschillen in loon

In de Loonwijzer-dataset is informatie over diverse kenmerken beschikbaar die wellicht een aanvullend deel van de variatie in het uurloon kunnen verklaren. Het geschatte loonverschil m-v kan hierdoor ook worden beïnvloed. Om dit te analyseren breiden we de basisloonschatting uit de vorige paragraaf uit met vier categorieën potentieel verklarende variabelen: (1) huishoudkenmerken; (2) subjectieve baankenmerken; (3) kenmerken die informatie geven met betrekking tot de man/vrouw verhouding in een type baan of type sector; en (4) overige relevante kenmerken.

Deze aanvullende variabelen zijn:

Huishoudkenmerken

- Er is een thuiswonend kind in het huishouden.
- Respondent heeft een partner.
- De combinatie werk-familie wordt als moeilijk beschouwd.
- De respondent heeft een koophuis.
- Er is een thuiswonend kind in de leeftijd van 0 tot 5 jaar.

Subjectieve baankenmerken

- Tevredenheid met de reistijd.
- Tevredenheid met de werk/vrije tijd verhouding.

- De respondent wil meer/minder uren werken.

Man/vrouw verhouding

- De meeste mannen in een soortgelijke functie zijn man
- Verhouding man/vrouw per sector

Overige kenmerken

- Ploegendienst/onregelmatige werkuren.
- Respondent vindt huidige baan saai.
- Respondent vindt huidige baan gevaarlijk.
- Respondent vindt huidige baan ongezond.
- Respondent is lid van een vakbond.

Een overzicht van de karakteristieken van deze variabelen (gemiddelde, spreiding, minimum, maximum en aantal waarnemingen) is opgenomen in Bijlage F, apart voor bedrijfsleven en overheid.

6.3.1 Wordt het loonverschil man-vrouw beter verklaard?

Het pure feit dat we meer verklarende variabelen beschouwen, wil niet automatisch zeggen dat het loonverschil tussen mannen en vrouwen kleiner wordt. We krijgen met een beter model een beter inzicht in welke factoren belangrijk zijn in de loonvorming en welke minder belangrijk, maar het verschil tussen mannen en vrouwen wordt daarmee niet direct verklaard. In de literatuur, besproken in het vorige hoofdstuk, worden echter hypothesen geformuleerd die aangeven dat vrouwen minder waarde zouden hechten aan ‘harde’ factoren zoals salaris en meer waarde aan ‘zachte’ factoren zoals baankenmerken. Ook is er een belangrijke hypothese die zegt dat vrouwen vaak minder verdienen omdat ze kiezen voor typische ‘vrouwenberoepen’, en dat die nu eenmaal lager betaald worden. Waarom die vrouwenberoepen lager betaald zijn blijft onduidelijk: zijn deze verschillen terecht, zijn ze onbewust traditioneel ontstaan, of is er sprake van bewuste discriminatie? De antwoorden op deze vragen vallen buiten het bestek van dit onderzoek, maar we kunnen wel vaststellen dat bovenstaande hypothesen inderdaad een bijdrage leveren aan de verklaring van de loonkloof. Wanneer we namelijk ons basismodel uitbreiden met bovengenoemde variabelen, zien we dat de resulterende loonkloof kleiner is geworden:

Tabel 6.5 Loonverschillen bedrijfsleven matching, o.b.v. Loonwijzer

Matching Methode	Loonverschil	Standaardfout	t-waarde
<i>Bedrijfsleven:</i>			
Radius matching ($r = 0.05$)	-0,105	0,006	16,6
Kernel matching (Epanechnikov)	-0,105	0,006	16,5
<i>Overheidssector</i>			
Radius matching ($r = 0.05$)	-0,098	0,013	7,6
Kernel matching (Epanechnikov)	-0,098	0,013	7,6

Dus 13,2% (in het bedrijfsleven) en 12,0% (bij de overheid) wordt respectievelijk 10,5% en 9,8%. Dit betekent dat het gecorrigeerde loonverschil man-vrouw dus grofweg 20% kleiner wordt wanneer we ‘corrigeren’ voor looncompenserende voorkeuren, huishoudvariabelen en

vrouwenberoepen.²⁹ Dit zegt niets over de wenselijkheid of redelijkheid van dergelijke beloningsverschillen, maar levert wel een bijdrage aan de vraag hoe de loonkloof ontstaat. Het geeft aan dat het loonverschil inderdaad beter ‘verklaard’ wordt met behulp van deze variabelen. Maar welke variabelen zijn nou de grootste oorzaak van deze ‘verkleining’ van de loonkloof? Dat bespreken we in de volgende subparagraaf aan de hand van de uitkomsten van een regressieanalyse, met dezelfde variabelen en dezelfde dataset.

6.3.2 Welke variabelen spelen een rol?

De loonschattingen met aanvullende variabelen, analoog aan de matching-analyse in paragraaf 6.3.1, worden weergegeven in Tabel 6.6 en

Tabel 6.7. We merken op dat we de schattingsresultaten behorende bij deze variabelen omwille van de duidelijkheid met grijs aangeven. De coëfficiënten van variabelen die ook in de basisloonschatting zijn opgenomen blijven in grote lijnen gelijk, de aanvullende variabelen zijn gearceerd. Merk op dat de uitkomsten van twee modelvarianten vermeld zijn. In de eerste variant zijn aparte dummy-variabelen opgenomen voor elke subsector. In de tweede variant nemen we man/vrouw-ratio binnen elke sector als verklarende variabele, deze pakt in feite het sectoreffect op maar biedt een andere interpretatie.

De aanvullende variabelen zijn nagenoeg allemaal significant, met uitzondering van de variabele die aangeeft of personen tevreden zijn met de combinatie werk en vrije tijd.³⁰ Merk op dat we weliswaar de *sterkte* van bepaalde verbanden in ons model kunnen bepalen, dat betekent niet dat we daarmee harde uitspraken kunnen doen over de *oorzaken* van deze verbanden. Wel proberen we indien mogelijk, met behulp van de literatuurstudie uit het vorige hoofdstuk, enkele waarschijnlijke hypothesen benoemen.

Kijken we naar de huishoudkenmerken dan blijkt dat de aanwezigheid van een thuiswonend kind een negatief effect heeft op het loon, behalve als het hele jonge kinderen betreft. Dit lijkt vreemd, maar is het niet als rekening gehouden wordt met een belangrijk selectie-effect: we kijken alleen naar personen die werken, en houden dus bv. geen rekening met het feit dat sommige moeders³¹ besluiten om niet te werken. Het positieve verband tussen jonge kinderen en het uurloon is dan begrijpelijk. Vrouwen zullen immers slechts dan gaan werken als het uurloon hoger is dan de prijs van de kinderopvang per uur, anders levert het financieel niets op om te gaan werken. M.a.w. mensen met hele jonge kinderen en een baan zijn te vinden in de relatief hogere salarisklassen. Het positieve verband is daarmee geen causaal verband in de zin dat ‘het hebben van jonge kinderen een hoger salaris oplevert’, maar een ondersteuning voor de hypothese dat vrouwen met jonge kinderen een hoger reserveringsloon zullen hebben dan vrouwen zonder jonge kinderen. Als er alléén oudere kinderen in het gezin aanwezig zijn is het uurloon gemiddeld lager. Dit effect weerspiegelt waarschijnlijk dat sommige vrouwen veelal weer (parttime) gaan werken als de kinderen eenmaal verplicht naar school gaan overdag. Als het hier vaker vrouwen betreft met een

²⁹ Hoewel we dit effect hier alleen demonstreren met behulp van matching, wordt de uitkomst bevestigd door OLS-analyse. In de volgende subparagraaf blijkt het loonverschil ook bij OLS na uitbreiding substantieel kleiner te worden: 14,5% wordt 12,4% en 11,0% wordt 9,5%.

³⁰ Het idee om combinatie tussen werk en vrije tijd op te nemen in de loonvergelijking komt voort uit veel gebruikte economische modellen waarin een werknemer wordt voorgesteld als een rationeel persoon die de optimale verdeling kan kiezen tussen vrije tijd en werk. Als personen ontevreden zijn over de combinatie werk en vrije tijd kan dit duiden op een onevenwichtige vrije tijd/werk-verhouding wat mogelijk tot uiting komt in het geobserveerde uurloon.

³¹ Uit nadere analyse is gebleken dat beide genoemde effecten vooral sterk van toepassing zijn op vrouwen.

gemiddeld lagere ‘verdiencapaciteit’ (waardoor ze aanvankelijk niet werkten toen het kind nog jonger was omdat het financieel niet loonde) treed eenzelfde selectie-effect op als hierboven, maar dan andersom. Aanvullende hypothese is dat de verdiencapaciteit van mensen die enkele jaren uit het arbeidsproces zijn geweest is afgenomen t.o.v. mensen die in die jaren zijn blijven werken. De laatsten hebben immers werkervaring opgedaan, cursussen gevolgd etc. en op die manier hun productiviteit verhoogd.

Ook de aanwezigheid van een partner hangt positief samenhang met het loon. Dit effect is empirisch moeilijk te scheiden van het bezitten van een koopwoning. Tweepersonshuishoudens hebben veel vaker een eigen woning hetgeen over het algemeen hogere lasten met zich meebrengt dan een huurwoning, en derhalve een relatief hoger inkomen noodzakelijk is. Waarschijnlijk observeren we dus vaker een koophuis bij die respondenten die de lasten kunnen opbrengen doordat ze een hoger uurloon verdienen. Het betreft dan geen beleidsmatig interessant causaal effect maar slechts een selectie effect.

De combinatie van werk en familie lijkt (in het bedrijfsleven althans) een looncompenserende factor te zijn: mensen met een lager loon zijn meer tevreden over de mogelijkheden familietaken met het werk te kunnen combineren.

Met betrekking tot de subjectieve baankenmerken zien we dat mensen die meer tevreden zijn met de reistijd een lager uurloon hebben. Zonder dit onomstotelijk te kunnen bewijzen, suggereert het wel dat mensen bereid zijn een iets lager loon te accepteren als daar minder reistijd tegenover staat. Wel moet bedacht worden dat dit effect weliswaar significant is, maar in omvang vrij beperkt. De variabelen die aangeven of personen meer of juist minder willen werken suggereert een duidelijk inkomenseffect: als het uurloon lager (hoger) is wil men meer werken om daarmee het totale arbeidsinkomen te verhogen (verlagen).

De kenmerken genoemd onder overige kenmerken en die betrekking hebben op het werk zijn allen negatief. Enerzijds is de hypothese dat gevaarlijk, ongezond, saai werk en onregelmatige arbeidstijden in loon gecompenseerd zou moeten worden door de werkgever. Dat blijkt dus niet het geval. De alternatieve hypothese van de ‘segmented labour market’ lijkt waarschijnlijker. Deze beweert namelijk dat gevaarlijk & ongezond werk samenhangt met de ‘slechtere’ banen aan de onderkant van de arbeidsmarkt die vervuld worden door werknemers met weinig verdiencapaciteit. In dit soort banen gaat laag loon vaak samen met slechtere arbeidsomstandigheden, simpelweg omdat de werknemers in deze functies weinig te eisen hebben. Vanwege deze samenhang wordt een gedeelte van het looneffect wel opgepikt door de variabelen opleiding en functieniveau, maar niet de gehele samenhang. Wat resteert is dus een statistisch artefact.

Het verband tussen het uurloon en lidmaatschap van een vakbond is positief. De achterliggende (indirecte) relatie zou kunnen zijn dat vakbonden vaker de belangen representeren van werknemers bij grotere bedrijven en in de industrie, en vakbonden hier hogere cao-lonen afdwingen. Het betreft hier dus niet het directe effect van lid worden van een vakbond: de voordelen gelden immers dankzij algemeen verbindend verklaring voor alle werknemers die onder de cao vallen. Wel is er een samenhang tussen vakbondslidmaatschap en de variabelen bedrijfsgrootte en sector: ze kunnen daarom niet zomaar los van elkaar geïnterpreteerd worden. Het lijkt erop dat vakbondslidmaatschap van de drie genoemde variabelen het minste verklarende kracht toevoegt, gelet op de grootte van de coëfficiënten.

De laatste twee variabelen geven informatie over de man/vrouw verhouding in vergelijkbare functies en de man/vrouw verhouding per sector. Met name als in een bepaalde functie de meeste personen man zijn is het uurloon gemiddeld hoger. De omvang van dit effect is aanzienlijk, zowel bij de overheid als bij het bedrijfsleven. Dit duidt niet zozeer op discriminatie, wel op een duidelijk verschil in beloning van typische vrouwenfuncties en typische mannenfuncties.

We merken op dat we in dit onderzoek niet tot doel hebben om de alle genoemde mogelijke verklaringen empirisch te toetsen. We willen slechts aangeven welke verklaringen ten grondslag liggen aan de a priori verwachting dat een variabele een significant effect heeft het uurloon, voortbouwend op de literatuurstudie uit Hoofdstuk 5. Het feit dat een variabele significant is toont de relevantie van de variabele aan, ongeacht de achterliggende verklaring.

Tabel 6.6 Uitgebreide loonschatting voor het bedrijfsleven, o.b.v. Loonwijzer

	Coëfficiënt	t-waarde	Coëfficiënt	t-waarde
Man	<i>referentie</i>			
Vrouw	-0,124	-41,5	-0,123	-41,0
Leeftijd	0,040	35,6	0,041	35,7
Leeftijd kwadraat	-0,000	-25,1	0,000	-25,1
Dienstjaren	0,004	9,6	0,004	9,4
Dienstjaren kwadraat	-0,000	-11,2	0,000	-11,2
Voltijd	<i>referentie</i>			
Deeltijd, 12 uur of meer	0,077	24,2	0,079	25,1
Deeltijd, minder dan 12 uur	0,139	10,7	0,137	10,5
Permanent contract	0,075	23,7	0,075	23,8
Uitzendkracht	-0,049	-9,7	-0,048	-9,5
ISCED 0-1 (Primair)	-0,095	-11,0	-0,100	-11,5
ISCED 2 (Lager secundair)	-0,071	-22,2	-0,074	-23,2
ISCED 3 (Hoger secundair)	0,010	2,0	0,007	1,5
ISCED 4 (Post-secundair)	<i>referentie</i>			
ISCED 5-6 (Tertiair)	0,161	55,9	0,164	57,1
Opleiding onbekend	0,022	1,2	0,023	1,3
Elementaire beroepen	-0,175	-13,2	-0,178	-13,5
Lagere beroepen	-0,142	-11,9	-0,140	-11,8
Middelbare beroepen	-0,057	-4,9	-0,047	-4,1
Hogere beroepen	0,036	3,0	0,047	3,9
Wetenschappelijke beroepen	0,092	7,0	0,102	7,8
Beroep onbekend	<i>referentie</i>			

Vervolg Tabel 6.6 Uitgebreide loonschatting voor het bedrijfsleven, o.b.v. Loonwijzer

	Coëfficiënt	t-waarde	Coëfficiënt	t-waarde
Technische functie	0,005	0,4	0,000	0,0
Administratieve functie	0,035	3,0	0,028	2,4
Automatiseringsfunctie	0,007	0,6	0,001	0,1
Commerciële functie	0,060	5,1	0,045	3,8
Dienstverlenende functie	0,030	2,5	0,029	2,4
Docent functie	0,030	1,2	0,026	1,0
Bestuurlijke functie	0,077	4,9	0,070	4,5
Leidinggevende functie	-0,057	-1,8	-0,068	-2,2
Landbouw en visserij	-0,034	-3,9		
Industrie	0,022	6,0		
Bouwnijverheid	0,006	1,3		
Handel	-0,040	-9,8		
Horeca	-0,040	-6,5		
Transport	-0,028	-6,0		
Zakelijke dienstverlening	referentie			
Zorg	0,009	2,1		
Overige diensten	-0,002	-0,4		
Bedrijfsgrootte: 1-9	-0,142	-35,4	-0,150	-38,0
Bedrijfsgrootte: 10-99	-0,083	-25,4	-0,087	-27,1
Bedrijfsgrootte: 100-499	-0,038	-10,8	-0,038	-10,9
Bedrijfsgrootte: 500 of meer	referentie			
Bedrijfsgrootte: onbekend	-0,075	-2,3	-0,078	-2,4
Regio Noord	-0,067	-16,7	-0,065	-16,2
Regio Oost	-0,046	-15,2	-0,044	-14,6
Regio West	referentie			
Regio Zuid	-0,040	-13,9	-0,039	-13,4
Regio Onbekend	-0,073	-4,7	-0,071	-4,6
Autochtoon	referentie			
Allochtoon	0,004	1,2	0,005	1,3
Etniciteit onbekend	-0,006	-0,5	-0,006	-0,4
CAO aanwezig	-0,009	-3,2	-0,010	-3,9
2004	-0,028	-7,5	-0,028	-7,4
2005	-0,043	-16,4	-0,042	-16,2
2006	referentie			
Huishoudkenmerken				
Thuiswonend kind	-0,012	-4,4	-0,012	-4,3
Respondent heeft een partner	0,011	4,0	0,010	3,8
Combinatie familie-werk is moeilijk	-0,012	-4,7	-0,011	-4,4
Respondent heeft een koophuis	0,075	29,6	0,075	29,7
Kind van 0 – 5 in huishouden	0,036	8,6	0,036	8,6

vervolg Tabel 6.6 Uitgebreide loonschatting voor het bedrijfsleven, o.b.v. Loonwijzer

	Coëfficiënt	t-waarde	Coëfficiënt	t-waarde
Subjectieve baankenmerken				
Tevredenheid met reistijd	-0,006	-5,9	-0,006	-5,8
Tevredenheid met combinatie werk-vrije tijd	n,s	-0,6	n,s	-0,2
Wil minder uren werken	0,015	5,5	0,014	5,2
Wil meer uren werken	-0,049	-11,9	-0,049	-11,7
Overige kenmerken				
Ploegendienst/onregelmatige werkuren	-0,014	-4,8	-0,017	-5,8
Respondent vindt huidige baan saai	-0,016	-13,5	-0,007	-5,4
Respondent vindt huidige baan gevaarlijk	-0,008	-6,2	-0,016	-13,4
Respondent vindt huidige baan ongezond	-0,006	-5,8	-0,006	-5,7
Lid van een vakbond	0,005	1,9	0,007	2,4
Man / vrouw verhouding				
De meeste personen in een soortgelijke functie zijn man	0,049	17,3	0,047	16,5
Ratio man/vrouw per sector			0,007	4,5
Constante	1,777	77,5	1,760	79,0
Aantal observaties	52625		52625	
Aangepaste R-kwadraat	0,462		0,462	

Tabel 6.7 uitgebreide loonschatting voor de overheidssector, o.b.v. Loonwijzer

	Coëfficiënt	t-waarde	Coëfficiënt	t-waarde
Man	<i>referentie</i>			
Vrouw	-0,095	-14,1	-0,094	-14,0
Leeftijd	0,025	9,1	0,025	9,2
Leeftijd kwadraat	0,000	-5,2	0,000	-5,2
Dienstjaren	0,005	5,4	0,005	5,5
Dienstjaren kwadraat	0,000	-5,5	0,000	-5,6
Voltijd	<i>referentie</i>			
Deeltijd, 12 uur of meer	0,084	11,7	0,084	11,7
Deeltijd, minder dan 12 uur	0,075	1,8	0,074	1,8
Permanent contract	0,059	7,3	0,059	7,4
Uitzendkracht	-0,032	-2,4	-0,031	-2,3
ISCED 0-1 (Primair)	-0,165	-3,9	-0,164	-3,9
ISCED 2 (Lager secundair)	-0,046	-4,7	-0,047	-4,8
ISCED 3 (Hoger secundair)	0,003	0,3	0,003	0,2
ISCED 4 (Post-secundair)	<i>referentie</i>			
ISCED 5-6 (Tertiair)	0,204	27,3	0,205	27,5
Opleiding onbekend	0,192	4,6	0,194	4,7

vervolg Tabel 6.7 uitgebreide loonschatting voor de overheidssector, o.b.v. Loonwijzer

	Coëfficiënt	t-waarde	Coëfficiënt	t-waarde
Elementaire beroepen	-0,293	-7,5	-0,289	-7,4
Lagere beroepen	-0,214	-8,5	-0,213	-8,4
Middelbare beroepen	-0,142	-6,0	-0,140	-6,0
Hogere beroepen	-0,043	-1,8	-0,042	-1,7
Wetenschappelijke beroepen	0,006	0,2	0,007	0,3
Onbekend	referentie			
Technische functie	0,023	1,0	0,023	0,9
Administratieve functie	0,079	3,4	0,078	3,4
Automatiseringsfunctie	0,058	2,1	0,058	2,1
Commerciële functie	0,157	5,6	0,156	5,5
Dienstverlenende functie	0,099	4,2	0,095	4,0
Docent functie	0,051	2,1	0,050	2,0
Bestuurlijke functie	0,070	2,8	0,070	2,8
Leidinggevende functie	0,131	3,9	0,127	3,8
Onderwijs	-0,107	-14,2		
Rechterlijke Macht	-0,035	-0,9		
Defensie	-0,022	-1,8		
Overheid overig	referentie			
Bedrijfs grootte: 1-9	-0,041	-3,0	-0,040	-2,9
Bedrijfs grootte: 10-99	-0,036	-4,9	-0,035	-4,8
Bedrijfs grootte: 100-499	-0,007	-0,9	-0,006	-0,8
Bedrijfs grootte: 500 of meer	referentie			
Bedrijfs grootte: onbekend	-0,007	-0,1	-0,008	-0,1
Regio Noord	-0,048	-5,2	-0,048	-5,2
Regio Oost	-0,026	-3,4	-0,026	-3,5
Regio West	referentie			
Regio Zuid	-0,030	-4,0	-0,030	-4,1
Regio Onbekend	-0,069	-1,9	-0,068	-1,9
Autochtoon	referentie			
Allochtoon	0,007	0,8	0,008	0,9
Etniciteit onbekend	-0,009	-0,3	-0,009	-0,3
CAO aanwezig	0,046	4,9	0,046	4,9
2004	-0,013	-1,4	-0,013	-1,4
2005	-0,036	-5,5	-0,036	-5,5
2006	referentie			
Huishoudkenmerken				
Thuiswonend kind	-0,029	-4,2	-0,029	-4,2
Respondent heeft een partner	0,016	2,5	0,016	2,5
Combinatie familie-werk is moeilijk	0,000	0,0	0,000	0,0

vervolg Tabel 6.7 uitgebreide loonschatting voor de overheidssector, o.b.v. Loonwijzer

	Coëfficiënt	t-waarde	Coëfficiënt	t-waarde
Respondent heeft een koophuis	0,071	11,2	0,071	11,1
Kind van 0 – 5 in huishouden	0,026	2,4	0,026	2,4
Subjectieve baankenmerken				
Tevredenheid met reistijd	-0,004	-1,7	-0,004	-1,6
Tevredenheid met combinatie werk-vrije tijd	0,005	1,8	0,005	1,8
Wil minder uren werken	0,006	0,8	0,006	0,8
Wil meer uren werken	-0,056	-5,6	-0,056	-5,6
Overige kenmerken				
Ploegendienst/onregelmatige werkuren	-0,020	-2,6	-0,018	-2,4
Respondent vindt huidige baan saai	-0,017	-5,5	-0,018	-5,7
Respondent vindt huidige baan gevaarlijk	-0,006	-2,0	-0,006	-2,0
Respondent vindt huidige baan ongezond	-0,006	-2,2	-0,006	-2,2
Lid van een vakbond	0,022	3,6	0,021	3,5
Man / vrouw verhouding				
De meeste personen in een soortgelijke functie zijn man	0,067	10,6	0,066	10,5
Ratio man/vrouw per sector			0,304	14,0
constante	2,001	36,6	1,769	31,9
Aantal observaties	7365		7365	
Aangepaste R-kwadraat	0,494		0,494	

6.4 Conclusie alternatieve verklaringen

In Hoofdstuk 6 bekeken we in hoeverre het aanvankelijke gecorrigeerde loonverschil nog verder ontleed kan worden met behulp van aanvullende, vaak subjectieve, variabelen. Nemen bijvoorbeeld vrouwen in functies met afwisselend werk en weinig reistijd genoeg met minder salaris, omdat de compenserende waarde van het afwisselende werk of de geringe reistijd voor hen belangrijker is? Dergelijke vragen zijn niet te beantwoorden met het AVO-bestand, zodat we gebruik moeten maken van de Loonwijzer dataset. De focus is daarom niet meer op het absolute percentage loonverschil man-vrouw maar op het gedeelte onverklaard loonverschil dat alsnog verklaard kan worden, en welke nieuwe variabelen hier een rol in spelen.

Het uitgangspunt is wederom een basisreplicatie, zo goed mogelijk vergelijkbaar gemaakt met de basisreplicatie op de AVO-data. Het blijkt dat de toevoeging van huishoudenkenmerken, subjectieve baankenmerken, man/vrouw-ratios per functie en enkele overige kenmerken een extra verklaring geeft voor het loonverschil tussen mannen en vrouwen. Al deze variabelen samen verklaren grofweg 20 procent van het aanvankelijke gecorrigeerde loonverschil uit de basisreplicatie. Er spelen dus meer factoren een rol bij loonverschillen tussen mannen en vrouwen dan de standaard-variabelen die normaal gesproken beschikbaar zijn in datasets die gebaseerd zijn op registraties.

Hoewel het onderzoek niet is bedoeld of opgezet om allerlei afgeleide hypothesen te toetsen, biedt de empirische verkenning in dit hoofdstuk wel degelijk aanknopingspunten bij enkele theorieën uit de wetenschappelijke literatuur over loonverschillen. Tevredenheid met reistijd blijkt inderdaad een looncompenserende factor te zijn (zij het beperkt in omvang) bij zowel

bedrijfsleven als overheid; een goede combinatie werk-vrije tijd alleen bij de overheid en de combinatie familie-werk alleen in het bedrijfsleven. Daarentegen worden negatieve baankenmerken als saai, gevaarlijk en ongezond werk niet gecompenseerd met een hoger loon, maar blijken het indicatoren van een slechte arbeidsmarktpositie.

Onder de huishoudensvariabelen blijken enkele selectie-effecten zichtbaar, die deels van invloed zijn op het gemiddelde loon maar waarschijnlijk minstens zozeer op de participatiebeslissing. Onze studie is niet toereikend om op basis van enkel deze huishoudensvariabelen concrete beleidsaanbevelingen te kunnen formuleren. Zij biedt echter wel enkele handvaten voor verdiepend onderzoek dat meer specifiek zou moeten kijken naar de effecten van beleid op het reserveringsloon, en de invloed daarvan op de participatiebeslissing.

Interessant is dat een in Hoofdstuk 3 geformuleerde conclusie, over grotere loonverschillen in typische mannensectoren, hier bijval krijgt. Hoe meer mannen in een bepaalde sector werken, hoe hoger het loon; dit geldt nog sterker voor de man/vrouw-ratio per beroep. Er blijken typische mannen- en typische vrouwenberoepen te zijn en juist die vrouwenberoepen zijn vaak de lagerbetaalde beroepen, ook binnen sectoren.

Bovenstaande suggereert dat de traditionele ‘gecorrigeerde beloningsverschillen’ in ieder geval niet zonder meer wijzen op directe discriminatie. Een gedeelte van het verschil zal altijd bestaan vanwege beperkingen aan databronnen en methoden, terwijl de interpretatie van de verschillen deels ook rekening moet houden met secundaire beloning, subjectieve factoren en zelfselectie van vrouwen in relatief minder betalende banen.

7 Kanttekeningen & aanbevelingen

Tot slot vermelden we in dit laatste hoofdstuk nog enkele overwegingen. In paragraaf 7.1 staan de kanttekeningen die gemaakt kunnen worden bij onze onderzoeksresultaten. Deze kanttekeningen zijn van belang bij de interpretatie van de resultaten, en de eventuele vertaling naar beleid. Voor hen die geïnteresseerd zijn in de technische achtergronden geven we in paragraaf 7.2 een opsomming van aanbevelingen die relevant zijn voor toekomstig onderzoek naar beloningsverschillen tussen mannen en vrouwen.

7.1 Kanttekeningen bij de onderzoeksresultaten

- De gegevens van de overheid zijn in de AVO-dataset van veel mindere kwaliteit dan de gegevens van de marktsector. Niet alleen ontbreken belangrijke variabelen (opleidingsniveau) maar ook belangrijke subsectoren (gemeente, provincie, politie). Dit betekent dat de uitkomsten voor de overheid onvergelykbaar zijn met de uitkomsten voor het bedrijfsleven. Er moet dus terughoudendheid betracht worden met de interpretatie; in elk geval zijn de resultaten niet te veralgemeniseren naar ‘de overheid in zijn geheel’.
- Binnen de sector overheid is onderwijs de belangrijkste subsector. Dit betreft echter een heterogene subsector die in totaal zes verschillende CAO’s omvat. Nadere uitsplitsing zou daarom wenselijk zijn, maar helaas is dat met de AVO-data niet mogelijk.
- Het voordeel van de Loonwijzer dataset is dat zij geen registratie maar een enquête is en daardoor ook subjectieve variabelen bevat. Dit voordeel heeft echter ook zijn nadeel: waarschijnlijk is sprake van zelfselectie, aangezien het geen aselechte steekproef uit een vastomlijnd steekproefkader is. Daarom kunnen we het absolute loonverschil uit deze databron niet als geldig voor heel Nederland zien, daarvoor zijn de AVO-data (wat bedrijfsleven betreft) meer geschikt.
- Werknemers in de zorgsector hebben wat betreft beloning, deeltijdvoorkeuren, aandeel vrouwen etcetera alle karakteristieken van de overheid, niet van de marktsector. Een analyse van loonverschillen wint aan nuance als de zorgsector als aparte sector wordt opgenomen, of bij de overheidsector wordt ingedeeld. Dit geldt in het bijzonder voor wat betreft beloningsverschillen tussen mannen en vrouwen (zie bijvoorbeeld Tabel 3.7).
- Op basis van ons onderzoek kunnen geen uitspraken worden gedaan over directe discriminatie. Het onverklaarde beloningsverschil uit ons type onderzoek bevat altijd meerdere factoren, één daarvan kan discriminatie zijn. Verder zal er altijd sprake blijven van meetfouten en het ontbreken van belangrijke variabelen. Om goed onderzoek te doen naar discriminatie is een ander type onderzoek nodig op een ander soort type data (bijvoorbeeld een experiment of een vignettenanalyse).
- Doel van dit onderzoek was enerzijds het verfijnen van de onderzoeksstrategie en anderzijds het onderzoeken van aanvullende verklaringen. De interessantste conclusie is wellicht dat loonverschillen deels samenhangen met selectie (zie paragraaf 6.3.2). In ieder geval wat betreft

beroepskeuze (selectie van vrouwen in functies met een lager salaris) maar waarschijnlijk ook wat betreft participatie (vrouwen met jonge kinderen blijven alleen werken als het hen genoeg oplevert). Dit laatste kan verklaren waarom de uitval onder laagopgeleide vrouwen hoger is dan onder hoogopgeleide vrouwen. Maar de constatering dat selectie een rol speelt vertelt nog niet *waarom* dit zo is; zo iets vraagt om een ander type onderzoek. Om te weten in hoeverre er specifiek beleid gemaakt kan worden om deze groep vrouwen niet te laten uittreden, is dus nader onderzoek nodig naar de oorzaken van uittreding/herintreding en het effect op beloningsverschillen.

7.2 Aanbevelingen voor vervolgonderzoek

- Beperk de analyse niet onnodig: alle relevante (beschikbare) factoren dienen onderzocht te worden. Voor onderzoek naar beloningsverschillen in de Mincer-traditie hoort daar in ieder geval het opleidingsniveau of het functieniveau bij, van alle werknemers.
- Momenteel beschikt het CBS helaas nog niet over integrale bestanden met zowel loongegevens als onderwijs- of functieniveau. Alleen van personen in de Enquête Beroepsbevolking is het onderwijs- en functieniveau bekend; sinds enkele jaren is er een onderwijsregistratie maar die betreft recente cohorten. Dit is een bekend probleem waaraan hard gewerkt wordt, maar wel met mogelijk gevolgen voor de algemene geldigheid van delen van het onderzoek. Wellicht komt het CBS in de toekomst met een oplossing voor dit probleem, maar ook selectiviteit ligt op de loer (zie volgende punt).
- Let op voor selectiviteit, die beperkt uiteindelijk de geldigheid van de resultaten. Als bijvoorbeeld voor bepaalde sectoren geen gegevens beschikbaar zijn of variabelen zijn slechts voor een beperkte groep gemeten, dan is geen sprake meer van algemene conclusies. Ook representativiteit van tijdelijke werknemers en werknemers uit het buitenland is vaak een probleem wanneer het survey-data betreft.
- Denk goed na over welke categorieën apart onderscheiden moeten worden en welke samengenomen kunnen worden. ‘Categorie onbekend’ dient bij voorkeur apart te worden behandeld, tenzij deze categorie verwaarloosbaar klein is en dus geen invloed heeft op de andere resultaten. In het geval van etniciteit in de AVO-data is aangetoond dat de groep ‘onbekend’ wezenlijk verschilt van beide andere groepen.
- Uitzendwerk is zeker een aparte dummy-variabele waard. Vaak zijn er meerdere variabelen die hiervoor geschikt zijn. Soms is zelfs mogelijk is mogelijk om het de uitzendbanen te scheiden van de vaste staffuncties bij uitzendbureaus (bv. in het Sociaal Statistisch Bestand (SSB) m.b.v. de variabelen *srtbed* & *sbi5*). Dit verdient zeker de voorkeur.
- Bij grotere datasets loont het de moeite om interactietermen te onderzoeken. Zijn de effecten op het loon gelijk voor mannen en vrouwen? Dergelijke uitbreidingen kunnen veel extra inzicht verschaffen in de precieze vorm van de relatie tussen verklarende variabelen onderling en tussen verklarende variabelen en de afhankelijk variabele. Indien wel interacties onderzocht zijn maar geen significantie gevonden wordt, is ook dat vaak vermeldenswaardig.

- Vaak is het mogelijk om een variabele aan te maken die aangeeft hoe de concentratie vrouwen in een bepaald stuk arbeidsmarkt is: bijvoorbeeld het %vrouwen per functie of anders in ieder geval per sector. Toevoeging van een dergelijke variabele aan de loonanalyse is vernieuwend en waarschijnlijk ook significant. In het SSB Banen zou dit zelfs op vrij gedetailleerd niveau kunnen met behulp van de variabele *sbi5*. Welk detailniveau het optimale is kan niet op voorhand bepaald worden, dat hangt van de analyse af.
- Subjectieve variabelen zijn belangrijk, zo is gebleken uit het laatste deel van ons onderzoek. Zijn die niet aanwezig in de dataset, houd daar dan rekening mee met interpretatie.
- Vergelijking van marktsector met overheid volgens de traditionele onderzoeklijnen (inclusief die in ons rapport) blijft altijd het theoretische probleem van zelfselectie houden. Bepaalde type werknemers kiezen nu eenmaal eerder voor de marktsector vanwege allerlei onbekende factoren ('ondernemerszin' ofzo); er is in feite voor dit soort personen geen echte vergelijkbare persoon te vinden bij de overheid. Om dit theoretische probleem te overwinnen is geavanceerdere econometrie nodig, zoals endogene switching regression-technieken waarmee eerst de kans om in een bepaalde sector te werken bepaald wordt. Dat stelt weer extra eisen aan de data: er moet worden gezocht naar factoren die wel de keuze van de sector bepalen, maar niet de hoogte van het loon. Dit is in praktijk heel erg lastig.
- Het CBS zou in haar toekomstige analyses gebruik moeten maken van het dynamische karakter van de SSB, men beschikt immers over loongegevens van personen over meerdere jaren. Het benutten van deze extra informatie maakt het mogelijk om rekening te houden met eventuele cohort-effecten en tijdvariërende variabelen zoals het krijgen van een kind. Bijvoorbeeld om een beter inzicht te krijgen in het effect van kortstondige uittrekking gevolgd door herintreding op de loonkloof van vrouwen biedt een dergelijke onderzoeksopzet veel perspectief.
- Analyses van loonverschillen lenen zich voor de toepassing van propensity score matching als schattingsmethode (zie Hoofdstuk 4). Matching heeft een aantal theoretische voordelen boven loglineaire regressies in de Mincer-traditie. Met name bij de analyse van loonverschillen onder ouderen en in de uithoeken van de populatieverdeling blijkt de functionele vorm van een Mincer-vergelijking regelmatig té bepalend. De resultaten van ons onderzoek lijken er op te wijzen dat ook propensity score matching tot relevante uitkomsten leidt.

Bijlage A Definities AVO dataset

In dit onderzoek wordt gebruik gemaakt van de gegevensbestanden van het onderzoek naar de arbeidsvoorwaardenontwikkelingen (AVO-onderzoek) van de Arbeidsinspectie. Het AVO-onderzoek wordt periodiek gehouden om gegevens te verzamelen over de stand van zaken en ontwikkeling van de belangrijkste arbeidsvoorwaarden. Informatie met betrekking tot het bedrijfsleven (markt- en zorgsector) zijn gegevens op basis van een steekproef verzameld door de inspecteurs van de Arbeidsinspectie. Naast de uit de loonadministraties afkomstige beloningsgegevens worden ook gegevens verzameld over een groot aantal persoons-, functie- en bedrijfskenmerken. Voor de vier overheidssectoren: onderwijs, rijk, defensie en rechterlijke macht zijn gegevensbestanden beschikbaar gesteld met soortgelijke informatie als voor het bedrijfsleven, waarbij we opmerken dat er geen gegevens beschikbaar zijn met betrekking tot opleidingsniveau, etnische afkomst, type arbeidsovereenkomst en functiesoort.

Hieronder lichten we toe hoe verschillende variabelen gedefinieerd zijn.

Persoonlijke toelage

Het salaris dat boven de voor een werknemer geldende normale schaal uitgaat, zoals een bindingspremie. Een toeslag wegens bijzondere bekwaamheden valt hier ook onder.

Tarief of provisie

Een, vaak, gedeeltelijke beloning naar inzet die afhankelijk kan zijn van de omzet. De belangrijkste categorie werknemers die op deze basis wordt beloond zijn handelsvertegenwoordigers.

Ploegentoeslag of een toeslag voor onregelmatige diensten

De extra uitbetalingen als gevolg voor wisselende arbeidstijden, werk in de avond, 's nachts of in het weekeinde.

Inconveniënten

Vergoedingen die gegeven worden voor vuil en onaangenaam werk

Loon in natura

De waarde van niet in geld uitgekeerd loon dat op de loonspecificatie van werknemers als fiscale bijtelling staat vermeld, zoals kost en inwoning

Overig loon

Niet eerder genoemde periodiek uitbetaalde loonbestanddelen zoals reiskostenvergoedingen en waarnemingstoelagen.

Winstafhankelijke uitkering

Een (jaarlijkse) eenmalige extra uitkering afhankelijk van het bedrijfsresultaat. Het al dan niet uitbetalen en de hoogte van de uitbetaling kan bijvoorbeeld gerelateerd zijn aan het bedrijfsresultaat, de winst voor of na belasting of de hoogte van het dividend.

Niet-winstafhankelijke uitkering

Onder deze uitbetalingen vallen arbeidsmarkttoeslagen, functioneringstoeslagen die doorgaans alleen aan bepaalde categorieën personeelsleden worden uitgekeerd en een 13^e/14^e maand. Voorts worden onder deze uitkering ook uitkeringen gerekend die een algemeen geldend karakter hebben en meestal bij cao zijn geregeld. Kenmerkend voor deze uitkeringen is dat zij niet structureel doorwerken op het beloningsniveau en niet afhankelijk zijn van het behaalde bedrijfsresultaat. In cao's worden vaak afspraken over de uitkeringen vastgelegd ter compensatie van een lagere, structureel op het beloningsniveau doorwerkende, generieke loonsverhoging.

Functieniveau

De functieniveaus zijn gebaseerd op de zogenaamde LTD-functieniveaus en zijn zodanig gekozen dat ze zo goed mogelijk aansluiten bij de ni vele cao's voorkomende niveaugroepen of combinaties hiervan. Deze functieniveaus zijn gedefinieerd door de Loontechnische Dienst, de toenmalige Arbeidsinspectie. Er wordt onderscheid gemaakt tussen de volgende functieniveaus:

Functieniveau I.

Zeer eenvoudige werkzaamheden die zich steeds herhalen, waarvoor geen scholing en slechts een geringe ervaring is vereist en die onder directe leiding worden verricht.

Functieniveau II.

Eenvoudige tot vrij eenvoudige werkzaamheden, die zich als regel herhalen, waarvoor (enige) lagere administratieve of technische kennis en enige ervaring vereist is, die enige zelfstandigheid vereisen en die doorgaans onder leiding worden verricht.

Functieniveau III - laag

Minder eenvoudige werkzaamheden met een overwegend herhalend karakter, waarvoor een lagere administratieve of technische kennis en ervaring is vereist en die tot op zekere hoogte zelfstandig worden verricht.

Functieniveau III- hoog

Minder eenvoudige werkzaamheden met een overwegend afwisselend karakter, waarvoor een lagere administratieve of technische kennis, gecompleteerd met een aanvullende beroepsopleiding in een specifieke procedure en/of techniek, en ervaring is vereist en die tot op zekere hoogte zelfstandig worden verricht.

Functieniveau IV

Moeilijke werkzaamheden van wisselende aard waarvoor een middelbaar administratief of technisch kennisniveau en daarbij behorende ervaring is vereist en die in beginsel zelfstandig worden verricht.

Functieniveau V

Samengestelde werkzaamheden binnen een bepaald vakgebied die een hoger kennisniveau en/of de nodige aanzienlijke ervaring vereisen en die zelfstandig worden verricht.

Functieniveau VI

Leidinggevende of beleidsvoorbereidende werkzaamheden van analyserende, creatieve of contactuele aard, die op grond van onafhankelijke oordeelsvorming worden verricht en die een academisch of gelijkwaardig kennisniveau vereisen.

Functieniveau VII

Leiding geven aan middelgrote ondernemingen of gelijkwaardige eenheden, diensten of staforganen, gepaard aan deelneming aan de beleidsvoorbereiding of beleidsvorming.

Functieniveau VIII

Leiding geven aan grote ondernemingen of gelijkwaardige concerndelen of instellingen.

Etnische afkomst

De etnische afkomst van een persoon is gebaseerd op basis van de Wet Samen. Volgens deze wet wordt een persoon gekenmerkt als allochtoon als de persoon zelf of een van de ouders van deze persoon geboren is in een van de volgende landen:

- Turkije;
- landen in Zuid- of Midden-Amerika (w.o. Suriname, Nederlandse Antillen, Aruba);
- landen in Afrika (w.o. Marokko);
- landen in Azië (m.u.v. Japan en voormalig Nederlands-Indië);
- voormalig Joegoslavië;
- Molukkers.

Functiesoort

Iedere arbeidsorganisatie met meerdere medewerkers bestaat uit functies die zich door bepaalde karakteristieken van elkaar onderscheiden. Op basis van de essentiële inhoud van de te verrichten werkzaamheden wordt onderscheidt gemaakt tussen 7 functiesoorten. Bij de indeling naar functiesoort is de tijdsbesteding bepalend indien tussen twee van de functiesoorten gekozen moet worden (meer dan 50%).

Productie, technische en/of handarbeidfuncties

Als zodanig worden al die functies aangemerkt die in de productiesfeer liggen evenals functies op een technisch vakgebied.

Administratieve functies

Kenmerkend voor de hieronder te brengen functies is het administratieve karakter van de te verrichten werkzaamheden voor zover die niet in de sfeer van de automatisering liggen. Dus ook een puur administratieve kracht op een verkoop-/inkoopafdeling en een typiste tekstverwerking, maar geen onstypiste of data-entry-typiste).

Automatiseringsfuncties

Alle functies waarin werkzaamheden ten behoeve van de input en/of output van een computer worden verricht en werknemers die zich bezighouden met programmering.

Commerciële functies

Alle functies waarin de te verrichten werkzaamheden zich kenmerken door klantgerichte bezigheden in de verkoopsfeer, dan wel functies in de inkoop sfeer die zich richten op activiteiten die als doel hebben de eigen organisatie toegang te verschaffen op de juiste inkoopmarkt.

Verzorgende en/of dienstverlenende functies

Binnen deze functiesoort kenmerken de werkzaamheden zich door verzorging van of dienstverlening aan anderen in de eigen organisatie of daarbuiten, waarbij die werkzaamheden niet tot een technisch vakgebied gerekend kunnen worden, noch administratief of commercieel van aard zijn.

Creatieve functies

Naast kunstenaars, mode-ontwerp(st)ers, etaleurs, architecten e.d. zijn hier ook werknemers ondergebracht die vanuit hun functie of vakgebied initiërend bezig zijn aan de ontwikkeling van procédés, processen, projecten of onderzoek, zoals een researchmedewerker (maar geen analist).

Leidinggevende of bestuursfuncties

Als zodanig worden die leidinggevende functies aangemerkt waarin het besturen op “managen” van een afdeling of een project meer dan 50% van de arbeidstijd vergt. Indien een functionaris vanuit zijn/haar vakgebied meer dan 50% van de tijd (technisch) inhoudelijk bezig is met de eigenlijke tot de functie of afdeling te rekenen uitvoerende werkzaamheden

Opleidingsniveau*1. Basisonderwijs (BO)*

Basisonderwijs, Buitengewoon Lager Onderwijs

2. Voorgezet Algemeen Onderwijs (VO)

MAVO, LAVO, IVO, Buitengewoon Voortgezet Onderwijs, Middenschool, onderbouw, HAVO/VWO, ULO, MULO.

3. Lager beroepsonderwijs (LBO)

LTS, LHNO, LEAO, LMO, LDS, VBO, Lagere Land-, Tuinen Bosbouwschool, School voor Visserij en Scheepvaart, CVV, ITO, Praktijkdiploma Boekhouden, primair leerlingwezen en Lagere Beroeps Opleidingen voor o.a.: Bus-, Taxi en Vrachtwagenchauffeur, Matroos, Machinist, Scheepskok, Postbode, Ziekenverzorgende, Verpleegassistent, Kraamhulp, Dierenartsassistent, Gezins-, Bejaarden- en Kinderverzorging, Beveiliging en Bewaking.

4. Middelbaar algemeen onderwijs (MO)

Volledig HAVO/VWO, Atheneum, Gymnasium, VHBO, HBS, MMS.

5. Middelbaar beroepsonderwijs (MBO)

MTS, MEAO, MLO, MMO, MDGO, Middelbare Land- Tuin, Bosbouw en Cultuurtechnische school, MBA, SPD-1 als ook vele richtingen van het MBO, KMBO en voortgezet leerlingwezen.

6. Hoger beroepsonderwijs (HBO)

HTS, HEAO, SPD-2, Hogere Land-, Tuin-, Bosbouw en Cultuurtechnische school, akte MO-A, N-akte, als ook vele hogere scholen en academies voor vakgebieden als onderwijs, talen, theologie, statistiek, scheepvaart, luchtvaart, verpleging, therapie, informatica, politie en defensie.

7. *Wetenschappelijk Onderwijs (WO)*

wetenschappelijke opleiding aan (technische) universiteit in een van de vele studierichtingen en akte MO-B

Wanneer er gekozen wordt voor een opleidingsverdeling: laag, middelbaar en hoog geldt de volgende verdeling:

- Laag opleidingsniveau: BO, VO en LBO
- Middelbaar opleidingsniveau: MO en MBO
- Hoog opleidingsniveau: HBO en WO

Regioindeling

- Noord: Provincies Groningen, Friesland en Drenthe
- Oost: Provincies Flevoland, Overijssel en Gelderland
- West: Provincies Noord- en Zuid-Holland, Utrecht
- Zuid: Provincies Zeeland, Noord-Brabant en Limburg

Arbeidsduur

Voltijders

Werknemers die evenveel uren werken als de voor het bedrijf geldende normale arbeidsduur (een volledige dienstbetrekking).

Grote deeltijders

Werknemers die 12 uur of meer werken, maar minder dan de voor het bedrijf geldende normale arbeidsduurwerken.

Kleine deeltijders

Werknemers die minder dan 12 uur werken.

Economische sector

Sectoren waartussen onderscheidt wordt gemaakt voor het bedrijfsleven:

- Landbouw en Visserij
- Industrie
- Bouwnijverheid
- Reparatie consumentenartikelen en handel
- Horeca
- Vervoer, opslag en communicatie
- Zakelijke dienstverlening
- Gezondheidszorg- en welzijnszorg
- Overige dienstverlening

Sectoren waartussen onderscheidt wordt gemaakt voor de overheidssector:

- Rijk
- Onderwijs
- Rechterlijke macht
- Defensie

Type werknemer

- BedrijfstakCAO
- Ondernemings CAO
- AVV
- Niet CAO'er in CAO bedrijf
- Niet CAO'er niet in CAO bedrijf
- Directie leden

Bijlage B Beschrijvende statistieken AVO

In deze bijlage worden de beschrijvende statistieken getoond voor het bedrijfsleven en de overheidssector.

Tabel B 1 Gewogen frequentietabellen bedrijfsleven

Kenmerk		Percentage	Kenmerk	Percentage		
Geslacht	Man	55%	Functieniveau	i	4%	
	Vrouw	45%		ii	14%	
Etnische afkomst	Autochtoon	94%	iii-laag	23%		
	Allochtoon	6%	iii-hoog	18%		
Arbeidsduur			iv	24%		
	Voltijd	55%	v	11%		
	Deeltijd >12 uur	33%	vi	3%		
	Deeltijd <12 uur	12%	vii en viii	0%		
Arbeidscontract			onbepaald	3%		
	Reg/uitz onb.	79%	Functiesoort	Technisch	22%	
	Reg/uitz bep	13%		Administratief	13%	
	Flex onb.	5%		Automatisering	2%	
Flex bep.	3%	Commercieel		14%		
Leeftijd			Verzorgend	37%		
	15	0%	Creatief	2%		
	16	1%	Bestuursfuncties	8%		
	17	1%	Onbepaald	3%		
	18	1%	sector	Landbouw	2%	
	19	2%		Industrie	16%	
	20	2%		Bouw	7%	
	21	2%		Handel	20%	
	22	2%		Horeca	4%	
	23 t/m 34	29%		Transport	7%	
	35 t/m 44	27%		Zakelijke	23%	
	45 t/m 54	22%		Gezondheidssector	17%	
55 plus	10%	Overige		5%		
Arbeidsverleden				Grootteklasse	Kleinbedrijf	19%
	0	15%			Middenbedrijf	28%
	1	11%			Grootbedrijf	53%
	2 t/m 5	34%				
	6 t/m 10	14%				
	11 t/m 20	15%				
Opleidingsniveau	21 plus	10%	Type werknemer	BedrijfstakCAO	58%	
				OndernemingsCAO	16%	
	Basisonderwijs	4%		AVV	5%	
	Lbo	36%		Niet-CAO'er in CAO bedrijf	1%	
	Mbo	29%		Niet-CAO'er Niet in CAO bedrijf	17%	
	Hbo	12%		Directie-leden	3%	
Wo	3%					
Onbekend	16%					
Regio			Salarisschaal?	Geen	42%	
	Noord	6%		Wel	58%	
	Oost	35%				
	West	34%				
	Zuid	25%				
Onbekend	0%					

Tabel B 2 Gewogen frequentie tabellen overheidssector

Kenmerk		Percentage	Kenmerk		Percentage
Geslacht	Man	47%	Functieniveau	i	2%
	Vrouw	53%		ii	1%
Arbeidsduur	Voltijd	55%		iii-laag	6%
	Deeltijd >12 uur	40%		iii-hoog	8%
	Deeltijd <12 uur	5%		iv	50%
Leeftijd	tot 23 jaar	3%		v	16%
	23 t/m 34	23%		vi	1%
	35 t/m 44	25%		vii en viii	17%
	45 t/m 54	33			
	55 plus	16%			
Arbeidsverleden	0	21%	Sector	Rijk	19%
	1	11%		Onderwijs	76%
	2 t/m 5	34%		Rechterlijke macht	1%
	6 t/m 10	17%		Defensie	4%
	11 t/m 20	10%			
	21 plus	7%			

Bijlage C Betrouwbaarheidsintervallen

In deze bijlage worden de geschatte loonverschillen en de bijbehorende betrouwbaarheidsintervallen getoond voor de bedrijfssector en de overheidssector.

Tabel C 1 Replicatie analyse arbeidsinspectierapport voor het bedrijfsleven: betrouwbaarheidsintervallen en loonverschillen

		Loonverschil	Stand. Fout..	95% B.I.-interval	
Gehele bedrijfsleven		-0,074	0,003	-0,080	-0,069
Etnische afkomst	Autochtoon	-0,077	0,003	-0,083	-0,071
	Allochtoon	-0,046	0,008	-0,062	-0,030
Arbeidsduur	Voltyd	-0,093	0,004	-0,101	-0,086
	Deeltijd meer dan 12 uur	-0,047	0,004	-0,056	-0,039
	Deeltijd minder dan 12 uur	-0,026	0,008	-0,042	-0,011
Arbeidscontract	Regulier/uitzend bepaalde tijd	-0,085	0,003	-0,091	-0,079
	Regulier/uitzend onbepaalde tijd	-0,061	0,006	-0,073	-0,048
	Flexibel onbepaalde tijd	-0,033	0,010	-0,051	-0,014
	Flexibel bepaalde tijd	0,010	0,011	-0,012	0,031
Leeftijd	Tot en met 22 jaar	-0,031	0,007	-0,044	-0,017
	23 tot 35 jaar	-0,050	0,004	-0,058	-0,042
	35 tot 45 jaar	-0,088	0,006	-0,099	-0,078
	45 tot 55 jaar	-0,101	0,007	-0,114	-0,087
	55 jaar en ouder	-0,053	0,011	-0,075	-0,032
Grootteklasse	Kleinbedrijf (1-9)	-0,102	0,010	-0,121	-0,084
	Middenbedrijf (10-99)	-0,078	0,004	-0,086	-0,069
	Grootbedrijf (100 of meer)	-0,064	0,004	-0,072	-0,057
Functieniveau	Functieniveau laag	-0,031	0,005	-0,040	-0,022
	Functieniveau midden	-0,081	0,003	-0,087	-0,075
	Functieniveau hoog	-0,105	0,009	-0,123	-0,086
	Functieniveau onbekend	0,012	0,017	-0,021	0,046
Economische sector	Landbouw en visserij	-0,091	0,017	-0,126	-0,057
	Industrie	-0,079	0,005	-0,090	-0,068
	Bouw	-0,114	0,019	-0,151	-0,078
	Handel	-0,086	0,005	-0,096	-0,075
	Horeca	-0,047	0,010	-0,067	-0,027
	Transport	-0,102	0,011	-0,124	-0,080
	Zakelijke dienstverlening	-0,063	0,006	-0,075	-0,051
	Gezondheids- en welzijnszorg	-0,060	0,009	-0,078	-0,041
	Overige dienstverlening	-0,052	0,009	-0,070	-0,034

Tabel C 2 Replicatie analyse arbeidsinspectierapport voor de overheidssector:
betrouwbaarheidsintervallen en loonverschillen

		Loonverschil	s.e. coeff.	95% B.I.-interval	
Gehele overheidssector		-0,040	0,001	-0,042	-0,037
Arbeidsduur	Voltijd	-0,038	0,002	-0,041	-0,035
	Deeltijd meer dan 12 uur	-0,032	0,003	-0,037	-0,026
	Deeltijd minder dan 12 uur	-0,087	0,011	-0,108	-0,066
Leeftijd	Tot en met 22 jaar	-0,002	0,010	-0,022	0,019
	23 tot 35 jaar	-0,019	0,002	-0,024	-0,014
	35 tot 45 jaar	-0,030	0,003	-0,035	-0,025
	45 tot 55 jaar	-0,045	0,003	-0,050	-0,040
	55 jaar en ouder	-0,055	0,004	-0,063	-0,048
Functieniveau	Functieniveau laag	-0,022	0,006	-0,033	-0,010
	Functieniveau midden	-0,011	0,002	-0,014	-0,008
	Functieniveau hoog	-0,027	0,001	-0,030	-0,024
	Functieniveau onbekend	-0,132	0,010	-0,151	-0,113

Bijlage D Loonschattingen voor bedrijfsleven en overheidssector

In deze bijlage presenteren we de schattingresultaten voor respectievelijk het bedrijfsleven en de overheidssector. Eerst in Tabel D.1 en Tabel D.2. de replicatie van het onderzoek van de Arbeidsinspectie: het bijbehorende loonverschil komt terug in de eerste regel van Tabel 3.1. Vervolgens geven we in Tabel D.3 en Tabel D.4 de uitgebreide schattingsresultaten van onze uitbreiding, waarvan een gedeelte is opgenomen in respectievelijk Tabel 3.3 en Tabel 3.4.

Tabel D 1 Loonschatting voor het bedrijfsleven (replicatie)

	Coëfficiënt	Standaardfout	95% B.I.-interval	
Man	<i>Referentie</i>			
Vrouw	-0,074	0,003	-0,080	-0,069
Autochtoon	<i>Referentie</i>			
Allochtoon	-0,048	0,004	-0,057	-0,040
Uitbetaald naar salarisschaal	-0,033	0,003	-0,038	-0,028
Persoon ontvangt extra loon	0,116	0,003	0,111	0,122
Voltijd	<i>Referentie</i>			
Deeltijd groter dan 12 uur	-0,044	0,003	-0,050	-0,038
Deeltijd kleiner dan 12 uur	-0,024	0,005	-0,033	-0,015
Regulier/uitzend bepaalde tijd	<i>Referentie</i>			
Regulier/uitzend onbepaalde tijd	-0,061	0,004	-0,069	-0,053
Flexibel onbepaalde tijd	-0,054	0,005	-0,065	-0,044
Flexibel bepaalde tijd	-0,076	0,007	-0,090	-0,062
15 jaar	-1,058	0,016	-1,089	-1,026
16 jaar	-0,908	0,012	-0,931	-0,885
17 jaar	-0,782	0,009	-0,800	-0,763
18 jaar	-0,622	0,009	-0,639	-0,604
19 jaar	-0,472	0,009	-0,489	-0,455
20 jaar	-0,349	0,008	-0,365	-0,334
21 jaar	-0,237	0,008	-0,252	-0,221
22 jaar	-0,155	0,008	-0,170	-0,140
23 tot 35 jaar	<i>Referentie</i>			
35 tot 45 jaar	0,102	0,003	0,096	0,107
45 tot 55 jaar	0,111	0,003	0,105	0,118
55 jaar en ouder	0,111	0,004	0,103	0,119
0 jaar arbeidsverleden	<i>Referentie</i>			
1 jaar arbeidsverleden	-0,020	0,004	-0,028	-0,011
2-6 jaar arbeidsverleden	n.s.	0,004	-0,014	0,002
7-11 jaar arbeidsverleden	0,035	0,005	0,026	0,044

Vervolg Tabel D.1 Loonschatting voor het bedrijfsleven (replicatie)

	Coëfficiënt	Standaardfout	95% B.I.-interval I	
11-21 jaar arbeidsverleden	0,050	0,005	0,041	0,059
21 of meer jaar arbeidsverleden	0,071	0,005	0,060	0,081
basisonderwijs	<i>Referentie</i>			
lbo	0,024	0,005	0,013	0,034
mbo	0,056	0,006	0,045	0,068
hbo	0,147	0,007	0,133	0,160
wo	0,259	0,009	0,241	0,276
onbekend	0,044	0,006	0,032	0,056
functieniveau 1	<i>Referentie</i>			
functieniveau 2	0,052	0,006	0,041	0,063
functieniveau 3 laag	0,113	0,006	0,102	0,124
functieniveau 3 hoog	0,183	0,006	0,171	0,194
functieniveau 4	0,310	0,006	0,298	0,322
functieniveau 5	0,544	0,007	0,530	0,559
functieniveau 6	0,754	0,010	0,735	0,773
functieniveau 7 en 8	1,192	0,014	1,165	1,219
functieniveau onbekend	0,302	0,017	0,268	0,336
Functiecat.: productie	<i>Referentie</i>			
Functiecat.: administratief	0,043	0,004	0,036	0,051
Functiecat.: automatisering	0,015	0,008	-0,001	0,030
Functiecat.: commercieel	n.s.	0,004	-0,013	0,003
Functiecat.: dienstverlenend	-0,017	0,003	-0,024	-0,010
Functiecat.: creatief	0,033	0,008	0,016	0,050
Functiecat.: leidinggevend	0,113	0,005	0,103	0,123
Functiecat.: onbekend	-0,134	0,016	-0,166	-0,102
Sector: Landbouw en Visserij	-0,076	0,008	-0,091	-0,060
Sector: Industrie	-0,063	0,005	-0,073	-0,053
Sector: Bouwnijverheid	<i>Referentie</i>			
Sector: Reparatie en handel	-0,112	0,005	-0,123	-0,102
Sector: Horeca	-0,085	0,007	-0,099	-0,071
Sector: Vervoer, opslag en communicatie	-0,102	0,006	-0,114	-0,089
Sector: Zakelijke dienstverlening	-0,104	0,006	-0,115	-0,093
Sector: Gezondheidszorg- en welzijnszorg	-0,012	0,007	-0,025	0,001
Sector: Overige dienstverlening	-0,080	0,006	-0,093	-0,067
Kleinbedrijf (1-9)	<i>Referentie</i>			
Middenbedrijf (10-99)	0,046	0,004	0,039	0,053
Grootbedrijf (100 of meer)	0,075	0,004	0,067	0,083
BedrijfstakCAO	<i>Referentie</i>			
Ondernemings CAO	0,105	0,004	0,098	0,113
AVV	-0,016	0,005	-0,026	-0,005

Vervolg Tabel D.1 Loonschatting voor het bedrijfsleven (replicatie)

	Coëfficiënt*	Standaardfout	95% B.I.-interval	
Niet CAO'er in CAO bedrijf	0,100	0,006	0,088	0,112
Niet CAO'er niet in CAO bedrijf	-0,018	0,003	-0,024	-0,012
Directie leden	0,140	0,008	0,125	0,156
Regio Noord	-0,044	0,004	-0,052	-0,035
Regio Oost	n.s.	0,003	-0,001	0,009
Regio Zuid	<i>Referentie</i>			
Regio West	-0,012	0,003	-0,017	-0,006
Constante	2,290	0,010	2,271	2,308
Aantal observaties	52208			
Aangepaste R-kwadraat	0,787			

* De afkorting n.s. (niet significant) geeft aan dat het kenmerk dat opgenomen is geen significante invloed heeft op het uurloon.

Tabel D 2 Loonschatting voor de overheidssector (replicatie)

	Coëfficiënt*	Standaardfout	95% B.I.-interval	
Man	<i>Referentie</i>			
Vrouw	-0,040	0,001	-0,042	-0,037
Persoon ontvangt extra loon	0,071	0,002	0,068	0,075
Voltijd	<i>Referentie</i>			
Deeltijd groter dan 12 uur	-0,010	0,001	-0,013	-0,007
Deeltijd kleiner dan 12 uur	-0,061	0,003	-0,068	-0,054
17 jaar	-0,985	0,146	-1,271	-0,698
18 jaar	-0,455	0,042	-0,537	-0,372
19 jaar	-0,353	0,027	-0,405	-0,301
20 jaar	-0,267	0,015	-0,296	-0,238
21 jaar	-0,180	0,010	-0,200	-0,160
22 jaar	-0,123	0,008	-0,138	-0,108
23 tot 35 jaar	<i>Referentie</i>			
35 tot 45 jaar	0,117	0,002	0,113	0,120
45 tot 55 jaar	0,166	0,002	0,163	0,170
55 jaar en ouder	0,186	0,002	0,182	0,191
0 jaar arbeidsverleden	<i>Referentie</i>			
1 jaar arbeidsverleden	0,007	0,003	0,002	0,013
2-6 jaar arbeidsverleden	0,014	0,002	0,010	0,018

Vervolg Tabel D.2 Loonschatting voor de overheidssector (replicatie)

	Coëfficiënt*	Standaardfout	95% B.I.-interval	
7-11 jaar arbeidsverleden	0,047	0,002	0,042	0,052
11-21 jaar arbeidsverleden	0,058	0,003	0,053	0,063
21 of meer jaar arbeidsverleden	0,016	0,003	0,011	0,021
Functieniveau 1	<i>Referentie</i>			
Functieniveau 2	0,131	0,009	0,113	0,150
Functieniveau 3 laag	0,201	0,009	0,184	0,218
Functieniveau 4	0,319	0,009	0,302	0,336
Functieniveau 5	0,569	0,009	0,552	0,586
Functieniveau 6	0,893	0,009	0,877	0,910
Functieniveau 7 en 8	1,267	0,010	1,248	1,286
Functieniveau 9	0,305	0,009	0,288	0,322
Functieniveau onbekend	0,131	0,009	0,113	0,150
Sector: Rijk	<i>Referentie</i>			
Sector: onderwijs	-0,044	0,002	-0,048	-0,041
Sector: rechterlijke macht	0,019	0,005	0,008	0,029
Sector: defensie	-0,059	0,002	-0,063	-0,055
Constante	2,250	0,009	2,233	2,268
Aantal observaties	65381			
Aangepaste R-kwadraat	0,810			

* De afkorting n.s. (niet significant) geeft aan dat het kenmerk dat opgenomen is geen significante invloed heeft op het uurloon.

Tabel D 3 Loonschatting voor het bedrijfsleven (uitbreiding)

	Coëfficiënt*	Standaardfout	95% B.I.-interval	
Man	<i>Referentie</i>			
Vrouw	-0,076	0,003	-0,082	-0,071
Autochtoon	<i>Referentie</i>			
Allochtoon	-0,049	0,004	-0,057	-0,04
Etnische afkomst onbekend	-0,016	0,006	-0,028	-0,004
Uitbetaald naar salarisschaal	-0,04	0,003	-0,041	-0,031
Persoon ontvangt extra loon	0,12	0,003	0,110	0,120
Voltijd	<i>Referentie</i>			
Deeltijd groter dan 12 uur	-0,04	0,003	-0,048	-0,036
Deeltijd kleiner dan 12 uur	-0,01	0,005	-0,022	-0,004
Regulier/uitzend bepaalde tijd	<i>Referentie</i>			

Vervolg Tabel D.3 Loonschatting voor het bedrijfsleven (uitbreiding)

	Coëfficiënt	Standaardfout	95% B.I.-interval	
Regulier bepaalde tijd	-0,05	0,004	-0,058	-0,041
Flexibel onbepaalde tijd	-0,053	0,005	-0,063	-0,042
Flexibel bepaalde tijd	-0,073	0,007	-0,087	-0,059
Uitzend onbepaalde tijd	-0,46	0,022	-0,504	-0,416
Uitzend bepaalde tijd	-0,357	0,031	-0,418	-0,297
15 jaar	-1,05	0,016	-1,086	-1,022
16 jaar	-0,91	0,012	-0,930	-0,882
17 jaar	-0,78	0,010	-0,798	-0,760
18 jaar	-0,62	0,009	-0,637	-0,601
19 jaar	-0,47	0,009	-0,488	-0,453
20 jaar	-0,35	0,008	-0,361	-0,330
21 jaar	-0,23	0,008	-0,249	-0,218
22 jaar	-0,15	0,008	-0,166	-0,136
23 tot 35 jaar	<i>Referentie</i>			
35 tot 45 jaar	0,10	0,003	0,094	0,106
45 tot 55 jaar	0,11	0,003	0,103	0,116
55 jaar en ouder	0,11	0,004	0,099	0,115
0 jaar arbeidsverleden	<i>Referentie</i>			
1 jaar arbeidsverleden	-0,02	0,004	-0,029	-0,011
2-6 jaar arbeidsverleden	<i>n,s</i> ,	0,004	-0,013	0,003
7-11 jaar arbeidsverleden	0,04	0,005	0,026	0,045
11-21 jaar arbeidsverleden	0,05	0,005	0,043	0,062
21 of meer jaar arbeidsverleden	0,07	0,005	0,061	0,083
basisonderwijs	<i>Referentie</i>			
lbo	0,02	0,006	0,010	0,031
mbo	0,05	0,006	0,041	0,064
hbo	0,14	0,007	0,130	0,158
wo	0,25	0,009	0,229	0,264
onbekend	0,04	0,006	0,030	0,054
functieniveau 1	<i>Referentie</i>			
functieniveau 2	0,06	0,006	0,049	0,071
functieniveau 3 laag	0,12	0,006	0,108	0,131
functieniveau 3 hoog	0,19	0,006	0,177	0,201
functieniveau 4	0,32	0,006	0,303	0,328
functieniveau 5	0,55	0,008	0,539	0,568
functieniveau 6	0,78	0,010	0,756	0,795
functieniveau 7 en 8	1,26	0,014	1,230	1,286
functieniveau onbekend	0,53	0,021	0,489	0,572
Functiecat.: productie	<i>Referentie</i>			
Functiecat.: administratief	0,04	0,004	0,037	0,053
Functiecat.: automatisering	<i>n.s.</i>	0,008	-0,006	0,026
Functiecat.: commercieel	<i>n.s.</i>	0,004	-0,010	0,006

Vervolg Tabel D.3 Loonschatting voor het bedrijfsleven (uitbreiding)

	Coëfficiënt*	Standaardfout	95% B.I.-interval	
Functiecat.: dienstverlenend	-0,02	0,003	-0,024	-0,011
Functiecat.: creatief	0,02	0,009	0,007	0,041
Functiecat.: leidinggevend	0,12	0,005	0,110	0,130
Functiecat.: onbekend	n.s.	0,018	-0,036	0,036
Sector: Landbouw en Visserij	-0,06	0,008	-0,078	-0,047
Sector: Industrie	-0,06	0,005	-0,075	-0,055
Sector: Bouwnijverheid	<i>Referentie</i>			
Sector: Reparatie en handel	-0,11	0,005	-0,118	-0,096
Sector: Horeca	-0,07	0,007	-0,088	-0,060
Sector: Vervoer, opslag en communicatie	-0,10	0,006	-0,113	-0,088
Sector: Zakelijke dienstverlening	-0,09	0,006	-0,104	-0,082
Sector: Gezondheidszorg- en welzijnszorg	<i>n,s,</i>	0,007	-0,017	0,010
Sector: Overige dienstverlening	-0,06	0,007	-0,077	-0,052
Kleinbedrijf (1-9)	<i>Referentie</i>			
Middenbedrijf (10-99)	0,04	0,004	0,034	0,049
Grootbedrijf (100 of meer)	0,07	0,004	0,060	0,076
BedrijfstakCAO	<i>Referentie</i>			
Ondernemings CAO	0,10	0,004	0,089	0,104
AVV	-0,01	0,006	-0,024	-0,002
Niet CAO'er in CAO bedrijf	0,11	0,006	0,099	0,124
Niet CAO'er niet in CAO bedrijf	-0,01	0,003	-0,021	-0,008
Directie leden	0,16	0,008	0,143	0,174
Regio Noord	-0,049	0,004	-0,058	-0,041
Regio Oost	n.s.	0,003	-0,002	0,008
Regio Zuid	<i>Referentie</i>			
Regio West	-0,014	0,003	-0,02	-0,009
Regio Onbekend	n.s.	0,016	-0,048	0,015
Dummy winstafhankelijke uitkering	0,133	0,003	0,126	0,14
Dummy niet-winstafhankelijke uitkering	0,055	0,002	0,051	0,06
Constante	2,29	0,01	2,271	2,308
Aantal observaties	52208			
Aangepaste R-kwadraat	0,787			

* De afkorting n.s. (niet significant) geeft aan dat het kenmerk dat opgenomen is geen significante invloed heeft op het uurloon.

Tabel D 4 Loonschatting voor de overheidssector (uitbreiding)

	Coëfficiënt*	Standaardfout	95% B.I.-interval	
Man	<i>Referentie</i>			
Vrouw	-0,038	0,001	-0,041	-0,036
Persoon ontvangt extra loon	0,070	0,002	0,066	0,073
Voltijd	<i>Referentie</i>			
Deeltijd groter dan 12 uur	-0,011	0,001	-0,014	-0,008
Deeltijd kleiner dan 12 uur	-0,060	0,003	-0,067	-0,053
17 jaar	-0,979	0,144	-1,262	-0,696
18 jaar	-0,453	0,042	-0,535	-0,371
19 jaar	-0,350	0,026	-0,402	-0,298
20 jaar	-0,264	0,015	-0,292	-0,235
21 jaar	-0,184	0,010	-0,204	-0,164
22 jaar	-0,124	0,008	-0,139	-0,109
23 tot 35 jaar	<i>Referentie</i>			
35 tot 45 jaar	0,117	0,002	0,113	0,120
45 tot 55 jaar	0,167	0,002	0,164	0,171
55 jaar en ouder	0,187	0,002	0,183	0,191
0 jaar arbeidsverleden	<i>Referentie</i>			
1 jaar arbeidsverleden	0,020	0,003	0,015	0,026
2-6 jaar arbeidsverleden	0,026	0,002	0,022	0,030
7-11 jaar arbeidsverleden	0,060	0,002	0,055	0,064
11-21 jaar arbeidsverleden	0,070	0,002	0,065	0,075
21 of meer jaar arbeidsverleden	0,028	0,003	0,023	0,033
Functieniveau 1	<i>Referentie</i>			
Functieniveau 2	0,106	0,009	0,088	0,124
Functieniveau 3 laag	0,179	0,009	0,162	0,196
Functieniveau 4	0,296	0,009	0,279	0,312
Functieniveau 5	0,542	0,008	0,525	0,558
Functieniveau 6	0,863	0,009	0,846	0,879
Functieniveau 7 en 8	1,232	0,009	1,213	1,250
Functieniveau onbekend	0,293	0,009	0,276	0,310
Sector: Rijk	<i>Referentie</i>			
Sector: onderwijs	-0,021	0,002	-0,025	-0,018
Sector: rechterlijke macht	0,018	0,005	0,008	0,028
Sector: defensie	-0,054	0,002	-0,058	-0,050

Vervolg Tabel D.4 Loonschatting voor de overheidssector (uitbreiding)

	Coëfficiënt*	Standaardfout	95% B.I.-interval	
Dummy niet-winstafhankelijke uitkering	0,103	0,048	0,009	0,197
Constante	2,184	0,049	2,088	2,280
Aantal observaties	65381			
Aangepaste R-kwadraat	0,81			

* De afkorting n.s. (niet significant) geeft aan dat het kenmerk dat opgenomen is geen significante invloed heeft op het uurloon.

Bijlage E Definities Loonwijzer-dataset

In deze bijlage wordt aangegeven hoe de verscheidene analysevariabelen uit de Loonwijzer data gedefinieerd zijn.

Opleidingsniveau

Het gaat hier om het hoogst behaalde opleidingsniveau waarbij de classificatie plaatsvindt op basis van een ISCED-code. ISCED staat voor “International Standard Classification of Education”. De indeling is als volgt:

- Preprimair onderwijs;
- Primair onderwijs;
- Lager secundair onderwijs;
- Hoger secundair onderwijs;
- Postsecundair onderwijs of tertiair onderwijs;

Arbeidsduur

Voltijders

Werknemers die evenveel uren werken als de voor het bedrijf geldende normale arbeidsduur (een volledige dienstbetrekking).

Grote deeltijders

Werknemers die 12 uur of meer werken, maar minder dan de voor het bedrijf geldende normale arbeidsduurwerken.

Kleine deeltijders

Werknemers die minder dan 12 uur werken.

Uitzendkracht

Een persoon wordt gekenmerkt als uitzendkracht als hij of zij in dienst is via een uitzendorganisatie.

Economische sector

Sectoren waartussen onderscheidt wordt gemaakt voor het bedrijfsleven:

- Landbouw en Visserij
- Industrie
- Bouwnijverheid
- Reparatie consumentenartikelen en handel
- Horeca
- Transport
- Zakelijke dienstverlening
- Gezondheidszorg- en welzijnszorg
- Overige sectoren

Sectoren waartussen onderscheidt wordt gemaakt voor de overheidssector:

- Rechterlijke macht
- Onderwijs
- Defensie
- Overige sectoren

Standaard bedrijfsclassificatie

De standaard bedrijfsclassificatie is gemaakt aan de hand van de SBC'93 code. We onderscheiden de volgende categorieën:

- Elementaire beroepen
- Lagere beroepen
- Middelbare beroepen
- Hogere beroepen
- Wetenschappelijke beroepen

Functiesoort

- Technische, Productie en/of handarbeidfuncties
- Administratieve functies
- Automatiseringsfuncties
- Commerciële functies
- Verzorgende en/of dienstverlenende functies
- Docent
- Bestuurlijke functies
- Leidinggevende functies

Bedrijfsgrootte

- Grootteklasse 0 tot 10
- Grootteklasse 10 tot 100
- Grootteklasse 100 tot 500

- Grootteklasse 500 plus

Bedrijfsregio

Vestiging van het bedrijf waar de respondent werkzaam is waarbij onderscheid gemaakt wordt tussen Noord, Oost, Zuid en West.

Etnische afkomst

In de AVO-data was de etnische afkomst van een persoon is gebaseerd op basis van de Wet Samen. In de Loonwijzer beschouwen de respondent als allochtoon als de moeder of vader van de respondent geboren is buiten Nederland.

Bijlage F Aanvullende variabelen Loonwijzer

Tabel F 1 Karakteristieken aanvullende variabelen Loonwijzer, bedrijfsleven

	gem.	s.d.	min.	max.	N
Thuiswonend kind	0,43	0,49	0	1	82626
Respondent heeft een partner	0,70	0,46	0	1	82626
Combinatie familie-werk is moeilijk	0,33	0,47	0	1	82626
Respondent heeft koophuis	0,61	0,49	0	1	82626
Kind van 0 – 5 in huishouden	0,09	0,29	0	1	82626
Tevredenheid met reistijd	3,83	1,16	1	5	82626
Tevredenheid met combinatie werk-vrije tijd	3,67	1,05	1	5	72185
Wil minder uren werken	0,25	0,43	0	1	82626
Wil meer uren werken	0,08	0,28	0	1	82626
Ploegendienst/onregelmatige werkuren	0,21	0,41	0	1	82626
Respondent vindt huidige baan saai	1,86	1,13	1	5	67854
Respondent vindt huidige baan gevaarlijk	2,29	0,95	1	5	82626
Respondent vindt huidige baan ongezond	2,44	1,25	1	5	68172
Lid van een vakbond	0,20	0,40	0	1	82626
Meeste personen in een functie zijn man	0,53	0,50	0	1	82626
Ratio man/vrouw per sector	1,19	0,91	0.19	3.70	82626

Tabel F 2 Karakteristieken aanvullende variabelen Loonwijzer, overheid

	gem.	s.d.	min.	max.	N
Thuiswonend kind	0,43	0,50	0	1	10708
Respondent heeft een partner	0,69	0,46	0	1	10708
Combinatie familie-werk is moeilijk	0,35	0,48	0	1	10708
Respondent heeft koophuis	0,64	0,48	0	1	10708
Kind van 0 – 5 in huishouden	0,07	0,26	0	1	10708
Tevredenheid met reistijd	3,81	1,19	1	5	10708
Tevredenheid met combinatie werk-vrije tijd	3,73	1,02	1	5	9753
Wil minder uren werken	0,23	0,42	0	1	10708
Wil meer uren werken	0,09	0,28	0	1	10708
Ploegendienst/onregelmatige werkuren	0,20	0,40	0	1	10708
Respondent vindt huidige baan saai	1,84	1,20	1	5	8922
Respondent vindt huidige baan gevaarlijk	2,25	0,93	1	5	10708
Respondent vindt huidige baan ongezond	2,30	1,20	1	5	8979
Lid van een vakbond	0,34	0,47	0	1	10708
Meeste personen in een functie zijn man	0,46	0,50	0	1	10708
Ratio man/vrouw per sector	0,61	0,17	0,40	0,74	10708

Bijlage G Nederlandse onderzoeken, opgenomen in het overzicht van studies naar de loonverschillen man-vrouw

Auteur	Publicatiejaar	Titel en uitgever	Instantie	Populatie	Bestand	Analyses	Schattingen deelpop.
Bakker, B.F.M., K.G. Tijdens, J.W. Winkels	1999	Gender, occupational segregation and wages in the Netherlands, Netherlands Official Statistics, 14, Winter, 36-41	CBS	werknemers	SEP	OLS + decompositie	m-v
Wetzels, C.M.M.P. & K.G. Tijdens	2002	Women's career break due to motherhood and the effects on wages. Cahiers Economiques de Bruxelles 45 (1), 169-189	TNO	vrouwelijke werknemers	Vrouwenloonwijzer	OLS	doorwerkers/h erintreedsters
Ruijter, J. de	2002	Occupational Wage differences: A Gender Approach. Diss., Rotterdam: Erasmus University Rotterdam	EUR	werknemers	Structure of Earnings Survey 1997	Multilevel analyses + regressies	n.v.t.
Veldman, A.G., S. Schalkwijk, m.m.v. A. van Doorne-Huiskes	2002	Gelijke beloning naar sekse, ras en arbeidsduur op ondernemingsniveau Een methode voor onderzoek naar de naleving van de Nederlandse gelijkebeloningswetgeving. Onderzoek verricht in opdracht van het Ministerie van Sociale Zaken en Werkgelegenheid door de Universiteit Utrecht, sectie Arbeidsrecht en Sociaal Beleid	MinSZW	41 paren van werknemers in 3 bedrijven	Personeelsbestanden + aanvullende info bedrijven	bivariat	n.v.t.
Hu, Y.	2005	Essays on Labour Economics: Empirical Studies on Wage Differentials across Categories of Working Hours, Employment Contracts, Gender and Cohorts. Diss., Tinbergen Institute / University of Amsterdam	UvA	vrouwen 16-65	ECHP 1998, restricted to individuals 16-65	ordered probit model + OLS + decompositie	arbtijd
Hoeben, J., P.M. Venema	2002	De arbeidsmarktpositie van werknemers in 2002. Een onderzoek naar de verschillen in beloning en mobiliteit tussen groepen werknemers. Den Haag: Arbeidsinspectie	Arbeidsinspectie	werknemers in marktsector plus zorgsector + werknemers bij overheid (rijk, onderwijs, rechterlijke macht, defensie)	AVO 2002	OLS + decompositie	m-v, lft, a/a, arbtijd, contract

Erdem, O., M.A. Bos, J. Hoeben	2006	De arbeidsmarktpositie van werknemers in 2004. Verschillen in beloning en mobiliteit. Den Haag: Arbeidsinspectie	Arbeidsinspectie	werknemers in marktsector plus gepremieerde en gesubsidieerde sector + werknemers bij overheid (rijk, onderwijs, rechterlijke macht, defensie)	AVO-2004	OLS + decompositie	m-v, lft, a/a, arbtijd, contract
Rienstra, M., K. Hagoort	2006	Beloningsverschillen verklaard? Verschillen in uurloon bij de overheid, 2004. Voorburg/Heerlen, CBS, Centrum voor Beleidsstatistiek 06001	CBS	werknemers in overheidssector (rijksoverheid, onderwijs, defensie, politie, rechterlijke macht, gemeenten, provincies en waterschappen)	Enquête Werkgelegenheid en Lonen (EWL) + persoonsgegevens uit de Enquête Beroepsbevolking (EBB) + persoonsgegevens uit de Gemeentelijke Basisadministraties (GBA)	OLS	m-v
Meer, P. van der	2007	Is the Gender Pay Gap declining in the Netherlands?	RUG	werknemers	OSA aanbod panel	OLS + decompositie	XX
Kooreman, P.	2007	The Early Inception of Labor Market Gender Differences. Tilburg University, revised paper	UTilburg	scholieren	Nationaal Scholierenonderzoek 1984-2001	OLS	m-v
Price Waterhouse Consulting (PWC)	2002	Gelijke monniken (m-v), gelijke kappen? Resultaten van een kwantitatief onderzoek naar gelijke beloning van mannen en vrouwen bij variabele beloning. Almere: PWC	PWC	630 werkgevers in Voeding; Groothandel; ICT; Finance en verzekering en 10 ministeries	No name	bivariaat	n.v.t.
Van Klaveren, M., W. Sprenger, K.G. Tijdens	2007	Dicht de loonkloof!	FNV	7 sectoren	statline data per sector	bivariaat	n.v.t.
Fouarge, D., R. Grim, M. Kerkhofs, A. Roman, T. Wiltshagen	2004	<i>Trendrapport Aanbod van arbeid 2003</i> . Tilburg, OSA	OSA	werknemers	OSA aanbod panel	kwartiel analyse	n.v.t.
Silfhout, R. van	2000	Inequality in pay within organizations: normative and instrumental perspectives. Tilburg University, dissertation	UTilburg	werknemers	werknemers in 10 organisaties	Lisrel, OLS regressies	n.v.t.
Tijdens, K.G., A. Dragstra, D. Dragstra, M. van Klaveren, P. Osse, C.M.M.P. Wetzels,	2002	Loonwijzers 2001/2002. Werk, lonen en beroepen van mannen en vrouwen in Nederland. Amsterdam University of Amsterdam, AIAS Research Report RR02/10	UvA/AIAS	werknemers	Loonwijzer 2001/02	OLS + decompositie	m-v, a/a

A. Zorlu							
Berkhout, E., A. Heyma, W. Salverda	2006	<i>Beloningsverschillen tussen de marktsector en collectieve sector in 2004.</i> Amsterdam: SEO Economisch Onderzoek, SEO rapport nr. 889	UvA/SEO	NL banen & werknemers	EWL & EBB	OLS + decompositie	markt-overheid
Berkhout, E., M. Biermans, W. Salverda, K. Tijdens	2007a	<i>Internationale beloningsverschillen van wetenschappelijk personeel.</i> Amsterdam: SEO Economisch Onderzoek, SEO-rapport nr. 981	UvA/SEO	intern. wetenschappelijk personeel universiteiten	eigen verzameling	n.v.t.	n.v.t.
Berkhout, E.E., P.H.G. Berkhout, M.L. Biermans	2006	<i>Studie & Werk 2006.</i> Amsterdam: SEO Economisch Onderzoek, 2006. SEO rapport nr. 902	SEO	recent afgestudeerde academici & hbo'ers	eigen enquête Studie & Werk	OLS	n.v.t.
Berkhout, E.E., S.G. van der Werff	2008	<i>Studie & Werk 2008.</i> Amsterdam: SEO Economisch Onderzoek, 2007. SEO rapport nr. 2008-28.	SEO	recent afgestudeerde academici & hbo'ers	eigen enquête Studie & Werk	OLS	n.v.t.
Heyma, A., E. Berkhout, W. Salverda, M. Biermans	2004	Beloningsverschillen tussen de marktsector en collectieve sector in 2001. Amsterdam: SEO Economisch Onderzoek	UvA/SEO	NL banen & werknemers	EWL & EBB	OLS + decompositie	markt-overheid
Heyma, A., K. Tijdens, E. Berkhout, L. Janssens	2005	<i>De aantrekkelijkheid van de collectieve sector als werkgever: een vergelijking van beloningsprofielen tussen collectieve sector en marktsector.</i> Amsterdam: SEO Economisch Onderzoek, SEO rapport nr. 797	UvA/SEO	NL banen & werknemers	EWL & EBB	OLS + decompositie	markt-overheid



seo economisch onderzoek

Roetersstraat 29 . 1018 WB Amsterdam . T (+31) 20 525 16 30 . F (+31) 20 525 16 86 . www.seo.nl