

Den uforklarede del af forskellen mellem kvinders og mænds timeløn

Mona Larsen, Mette Verner og Christian Højgaard Mikkelsen

Den uforklarede del af forskellen mellem kvinders og mænds timeløn

© VIVE og forfatterne, 2020

e-ISBN: e-ISBN-nr.

Forsidefoto: Fotografens navn/VIVE

Projekt: 301436

VIVE – Viden til Velfærd

Det Nationale Forsknings- og Analysecenter for Velfærd

Herluf Trolles Gade 11, 1052 København K

www.vive.dk

VIVEs publikationer kan frit citeres med tydelig kildeangivelse.

Forord

Kvinder har en lavere indtjening over livsforløbet end mænd. Det skyldes bl.a., at de i gennemsnit har en lavere timeløn. Tidligere undersøgelser viser, at halvdelen af denne lønforskel hænger sammen med, at kvinder og mænd er forskelligt fordelt på målbare karakteristika, herunder især, at de befinder sig på forskellige dele af arbejdsmarkedet.

Formålet med denne undersøgelse er at danne grundlag for at få mere viden om den resterende halvdel af lønforskellen (den 'uforklarede' del). Undersøgelsens resultater giver dermed et bedre udgangspunkt for at drøfte, hvad der skal til for at reducere forskellen mellem kvinders og mænds timeløn.

Til undersøgelsen har været knyttet en følgegruppe, der takkes for konstruktive kommentarer og diskussioner undervejs. Rapporten er blevet læst og kommenteret af to eksterne reviewere, der takkes for grundige, indsigtsfulde og konstruktive kommentarer.

Rapporten er udarbejdet af professor Mette Verner, analytiker Christian Højgaard Mikkelsen og seniorforsker Mona Larsen, der også har været projektleder på undersøgelsen.

Beskæftigelsesministeriet har bedt VIVE – Det Nationale Forsknings- og Analysecenter for Velfærd om at gennemføre undersøgelsen, ligesom Beskæftigelsesministeriet har finansieret undersøgelsen.

Lisbeth Pedersen

Forsknings- og analysechef for VIVE Arbejde og Ældre
2020

Indhold

Sammenfatning.....	6
Ny viden om lønforskellen.....	7
Et forbedret sæt af kontrolvariabler.....	9
1 Undersøgelsens formål og baggrund.....	10
1.1 Formål.....	10
1.2 Baggrund.....	10
2 Hvad kan reducere den 'uforklarede' lønforskel?.....	14
2.1 Segregering på arbejdsmarkedet.....	14
2.2 Firmaspecifikke effekter.....	15
2.3 Effekter af jobskifte.....	16
2.4 Effekter af nedsat arbejdstid.....	16
2.5 Effekter af fravær.....	17
2.6 Effekter af børn.....	17
2.7 Personlighedstræk.....	18
2.8 Ubevidst bias.....	18
2.9 Hvad tager vi højde for?.....	19
2.10 Hvad tager vi ikke højde for?.....	20
3 Datagrundlag og analysedesign.....	21
3.1 Datagrundlag.....	21
3.2 Analysedesign.....	23
4 Valg af ændret sæt af kontrolvariabler.....	29
4.1 Fase 1: Kønsforskel i afkast af karakteristika.....	29
4.2 Fase 2: Forbedret opgørelse af allerede målte karakteristika.....	31
4.3 Fase 3: Tilføjelse af nye karakteristika.....	32
4.4 Lønfordelingen 2016.....	36
4.5 Sektorniveau.....	37
4.6 Andre år.....	39
5 Ny viden om lønforskellen mellem kvinder og mænd.....	41
5.1 Hele arbejdsmarkedet 2016.....	41
5.2 Den private sektor 2016.....	44
5.3 Udviklingen i perioden 2010-2017.....	46
Bilag 1 Blinder-Oaxaca-dekomponeringsanalyse.....	49
Bilag 2 Beskrivende statistik og resultater af regressionsanalyser.....	51
Bilag 3 Fraværsvariabler.....	54
Bilag 4 Resultater for sektorer for 2016.....	56
Bilag 5 Resultater for hele arbejdsmarkedet for udvalgte år.....	57

Litteratur.....	59
-----------------	----

Sammenfatning

Kvinder har en lavere timeløn end mænd. I 2017 var forskellen på 14,4 pct. Cirka halvdelen af denne lønforskel hænger ifølge tidligere SFI-/VIVE-studier sammen med, at kvinder og mænd er forskelligt fordelt på en række karakteristika. Det spiller især en rolle, at kvinder og mænd befinder sig på forskellige dele af arbejdsmarkedet. Lønforskellen ville have været endnu større, hvis kvinder ikke havde overhalet mænd med hensyn til uddannelseslængde. Den del af lønforskellen, der kan tilskrives forskelle i de to køns fordeling på målbare karakteristika, betegnes den 'forklarede' del af lønforskellen. Den resterende halvdel af lønforskellen er 'uforklaret'.

Formålet med denne undersøgelse er at danne grundlag for at få mere viden om forskellen mellem kvinders og mænds timeløn. Mere specifikt er formålet at forbedre det sæt af kontrolvariabler, der har været anvendt i de tidligere SFI-/VIVE-undersøgelser af lønforskellen, for dermed at få mere viden om den del af forskellen, der hidtil har været 'uforklaret'.

Den 'uforklarede' del af lønforskellen er et statistisk begreb, hvis størrelse afhænger af de målbare forhold, der indgår i den gennemførte analyse. Denne del af lønforskellen hænger sammen med:

- Forskellig aflønning af kvinder og mænd med 'samme' karakteristika, fx samme uddannelse eller befinder sig på den samme del af arbejdsmarkedet, og
- uobserverede forhold, som analysen ikke har taget højde for.

I analysen har vi reduceret størrelsen af den 'uforklarede' del ved at forbedre det anvendte sæt af kontrolvariabler på to måder: Dels ved at forbedre målingen af nogle af de karakteristika, der hidtil er taget højde for og dels ved at tilføje flere karakteristika, der hænger sammen med lønforskellen, men som ikke har været inddraget i de tidligere analyser. I forlængelse heraf har vi beskrevet den nye viden om lønforskellen, som vi kommer frem til ved brug af det ændrede sæt af kontrolvariabler. Datagrundlaget for undersøgelsen er registerdata fra Danmarks Statistik.

Det er indledningsvis vigtigt at understrege, at den 'uforklarede' lønforskel *ikke er et mål for*, om der finder forskelsbehandling sted på det danske arbejdsmarked. Rapportens resultater giver med andre ord ikke noget svar på, hvor stor en del af bruttolønforskellen, der er 'retfærdig' eller 'uretfærdig', herunder spørgsmålet om potentiel diskrimination. 'Retfærdige' lønforskelle, det vil her sige lønforskelle, som skyldes forhold, som alle er enige om har betydning for timelønnen, kan for så vidt findes både i den 'forklarede' og den 'uforklarede' del. Og tilsvarende kan 'uretfærdige' elementer af lønforskellen optræde begge steder.

Tilsvarende siger rapportens resultater ikke noget om, hvorvidt kvinder og mænd får samme timeløn for samme arbejde eller for arbejde af samme værdi, dvs. at resultaterne ikke giver et billede af, om der er ligeløn på det danske arbejdsmarked i juridisk forstand. Resultaterne *afspejler* heller *ikke de bagvedliggende årsager* til lønforskellen, men kan alene *danne udgangspunkt* for en drøftelse af, hvad der forklarer forskellen mellem de to køns timeløn.

Ny viden om lønforskellen

Ved at bruge et ændret sæt af kontrolvariabler får vi mere viden om lønforskellen på arbejdsmarkedet som helhed og i den private sektor. Desuden får vi mere viden om faktorer, der har betydning for de senere års fald i lønforskellen.

I det følgende er nogle af de forhold, der hænger sammen med, at kvinder i gennemsnit får en lavere timeløn end mænd skitseret. De resultater, som er nye i forhold til resultaterne af de tidligere SFI-/VIVE-studier, er markeret med "nyt".

Arbejdsmarkedet som helhed

85 pct. (mod tidligere 52 pct.) af lønforskellen mellem kvinder og mænd på arbejdsmarkedet som helhed hænger ifølge de nye analyser sammen med, at en række karakteristika er skævt fordelt på de to køn. De vigtigste resultater baseret på tal for 2016 er her:

- Lønforskellen hænger især sammen med, at arbejdsmarkedet er kønsopdelt – betydningen heraf er større end hidtil antaget. Kønsopdelingen fører til lønforskel, fordi:
 - Der er systematiske forskelle mellem lønnen i de kvinde- og de mandedomerede job: Jo flere af de ansatte i ens jobcelle (dvs. arbejdsfunktion på arbejdssted), der er kvinder, jo lavere er lønnen (nyt).
 - Mænd er placeret højere oppe i stillingshierarkiet end kvinder (nyt)
- Lønforskellen hænger også sammen med, at:
 - mænd har mere erhvervserfaring end kvinder
 - kvinder har mere fravær end mænd (nyt)
- Lønforskellen mellem kvinder og mænd ville have været større:
 - hvis kvinder i gennemsnit havde haft samme (kortere) uddannelseslængde som mænd.

Der er en række mulige forklaringer på, at *jo flere kvinder, der er i en given jobcelle, jo lavere er lønnen*. Disse forklaringer er forskellige for den offentlige og den private sektor.

I den offentlige sektor kan en forklaring på den negative sammenhæng mellem kvindeandel og løn for det første være, at lønningerne i den offentlige sektor ved en reform af lønsystemet i den offentlige sektor for år tilbage blev fastlåst i et bestemt forhold, hvor kvindefagene fik tildelt relativt lavere lønninger end mandefagene. Efterfølgende overenskomstforhandlinger skulle så have reproduceret disse fastlåste lønrelationer. For det andet kan forklaringen være, at nogle kvindefag i perioder i større omfang end mandefag har forhandlet sig til goder frem for løn. For det tredje kan forklaringen være, at der i kvindefagene er én køber af fagets arbejdskraft, som har stor mulighed for at påvirke lønnen (i nedadgående retning).

For den private sektor kan forklaringen på den negative sammenhæng mellem kvindeandel og løn for det første være, at løndannelsen her foregår mere decentralt. Kvinder er mindre tilbøjelige end mænd til at forhandle løn og er mindre konkurrenceorienterede, hvorfor de får lavere lønstigninger. For det andet kan forklaringen være, at kvinder er mere risikoaverse end mænd. Dette kan betyde, at kvinder i større omfang vælger fag i den private sektor med en stabil, men typisk også en lavere indtjening i gennemsnit.

Mere fravær kan give lavere løn, hvis humankapitel afskrives under fraværperioder og hvis længere (frivillige) fraværperioder er et signal om, at man er mindre dedikeret til sit job.

Hvis ikke kvinder i stigende omfang havde taget *mere uddannelse* end mænd, ville lønforskellen mellem de to køn have været endnu større. Kvinder tager i dag i gennemsnit flere års *uddannelse* end mænd – denne forskel er forøget over tid. Udviklingen kan illustreres ved, at nogle af de kvinder, der tidligere fx valgte at blive sygeplejerske eller at tage en kontoruddannelse, i dag vælger at uddanne sig til fx læge eller jurist.

De resterende 15 pct. af lønforskellen på arbejdsmarkedet som helhed er ifølge de nye analyser 'uforklaret'. Det vil sige, at 15 pct. af lønforskellen kan tilskrives enten, at kvinder og mænd med 'samme' karakteristika bliver aflønnet forskelligt eller uobserverede forhold, som analysen ikke har taget højde for. Viden om indholdet af den 'uforklarede' del er ny i forhold til de tidligere SFI-/VIVE-studier.

De vigtigste resultater, når det gælder forskellig aflønning af kvinder og mænd med 'samme' karakteristika, er:

- Kvinder får et mindre lønafkast end mænd, når de to køn har samme uddannelseslængde
- Mænd får et mindre lønafkast, jo flere kvinder, der er i den sektor, hvor de er ansat – en negativ sammenhæng, der ikke viser sig blandt kvinder
- Mænd får et mindre lønafkast end kvinder blandt ansatte, der er placeret "lavt" i jobhierarkiet, dvs. ansatte, der hverken er ledere eller mellemledere eller har et særligt ansvar
- Mænd bliver "straffet" mere lønmæssigt for langvarigt fravær end kvinder.

Af uobserverede forhold, der også kan have betydning for lønforskellen, kan fx nævnes kønsforskelle mht. ikke-registreret overarbejde, præferencer og personlighedstræk. Desuden kan ubevidst bias spille en rolle, det vil her sige, at mennesker klassificeres ud fra deres køn. Forekomsten af ubevidst bias kan både direkte og indirekte bidrage til, at der er lønsforskelle mellem kvinder og mænd.

Den private sektor

78 pct. (mod tidligere 19 pct.) af lønforskellen i den private sektor kan tilskrives, at kvinder og mænd er forskelligt fordelt på målbare karakteristika. De vigtigste resultater for den private sektor baseret på tal for 2016 er:

- En væsentlig del af lønforskellen hænger sammen med, at arbejdsmarkedet er kønsopdelt – kønsopdelingen spiller en større rolle i den private sektor end hidtil antaget. Det spiller især en rolle, at:
 - Jo flere af de ansatte i ens jobcelle (dvs. arbejdsfunktion på arbejdssted), der er kvinder, jo lavere er lønnen (nyt).
- Lønforskellen hænger også sammen med, at:
 - mænd har mere erhvervs erfaring end kvinder
 - kvinder har mere fravær end mænd. Især kønsforskelle i barselsfravær spiller en rolle (nyt)
- Lønforskellen mellem kvinder og mænd ville have været større:
 - hvis kvinder i gennemsnit havde haft samme (kortere) uddannelseslængde som mænd.

De resterende 22 pct. af lønforskellen i den private sektor er 'uforklaret'.

Fald i lønforskellen fra 2010 til 2016

Lønforskellen på arbejdsmarkedet som helhed er faldet med 2,5 procentpoint fra 2010 til 2017. Faldet i lønforskellen hænger især sammen med, at:

- Kvinders forspring mht. uddannelse er steget i perioden.
- Den "lønstraf", der er forbundet med barselsfravær er blevet mindre i perioden (nyt)

En af forklaringerne på, at "lønstraffen" forbundet med barselsfravær er blevet mindre fra 2010 til 2017 kan være den gunstige konjunkturudvikling i perioden. Faldet i "lønstraffen" viser sig i den private sektor. Dette kan tale for, at "lønstraffen" for barselsfravær er mindre i private virksomheder, når der er højkonjunktur.

Et forbedret sæt af kontrolvariabler

Vi har forbedret det sæt af kontrolvariabler, der har været anvendt i de tidligere SFI-/VIVE-undersøgelser. I de tidligere undersøgelser blev der taget højde for følgende karakteristika: Uddannelsesniveau, erhvervs erfaring, sektor, branche, arbejdsfunktion, arbejdstid, familieforhold og bopæl. Sektor, branche og arbejdsfunktion er samlet set et mål for kønsopdelingen på arbejdsmarkedet.

De vigtigste ændringer i sættet af kontrolvariabler består i:

- Uddannelsesniveau er erstattet af uddannelseslængde og -retning,
- Sektor og arbejdsfunktion er erstattet af kvindeandel inden for hhv. sektor og arbejdsfunktion på arbejdssted (kaldet 'jobcelle'),
- Jobstatus, dvs. om man er elev/lærling, leder/mellemlider, medarbejder med særligt ansvar eller øvrig medarbejder, er tilføjet,
- Antal fuldtidsansatte på arbejdsstedet er tilføjet, og
- Fravær i året i den aktuelle ansættelse og hhv. syge- og barselsfravær i de fem forudgående år er tilføjet.

Brugen af det ændrede sæt af kontrolvariabler giver os mere viden om lønforskellen mellem kvinder og mænd på arbejdsmarkedet som helhed. Vi får mere viden om lønforskellen i hele lønfordelingen. Selv i toppen af lønfordelingen, hvor lønforskellen er størst, kan hele 60 pct. af lønforskellen i 2016 (mod tidligere 36 pct.) tilskrives forskelle i kvinders og mænds karakteristika.

Det ændrede sæt af kontrolvariabler er især velegnet til analyser for den private sektor. Det skyldes tilføjelsen af oplysninger på arbejdsstedsniveau, hvilket gør det muligt i højere grad at indfange betydningen af, at decentral løndannelse er udbredt i den private sektor. Jo mindre overenskomstdækket et område er, jo mere individuel er løndannelsen, og jo vigtigere er det, der foregår på arbejdspladsen for lønforskellen mellem kvinder og mænd. Tilføjelsen af oplysninger om fravær spiller også en rolle. "Lønstraf" for fravær forekommer især i den private sektor.

1 Undersøgelsens formål og baggrund

I dette kapitel beskriver vi undersøgelsens formål. Som baggrund for undersøgelsen forklarer vi dernæst, hvad den 'uforklarede' lønforskel består i og hvor stor den er.

1.1 Formål

Kvinder har en lavere indtjening over livsforløbet end mænd, hvilket bl.a. hænger sammen med, at de i gennemsnit får en lavere timeløn. Forskellen mellem to køns timeløn var ifølge Larsen & Larsen (2019) på 14,5 pct. i 2016. Larsen & Larsen (2018) viser også, at ca. halvdelen af denne lønforskel hænger sammen med forskelle i de to køns karakteristika, herunder især, at kvinder og mænd befinder sig på forskellige dele af arbejdsmarkedet. Lønforskellen ville have været endnu større, hvis kvinder ikke havde overhalet mænd med hensyn til uddannelseslængde. Den resterende halvdel af lønforskellen er 'uforklaret'.

Formålet med denne undersøgelse er at forbedre det sæt af kontrolvariabler, der har været anvendt i en række tidligere SFI-/VIVE-undersøgelser af forskellen mellem kvinders og mænds timeløn (Deding & Wong, 2004; Deding & Larsen, 2008; Larsen & Houlberg, 2013a; Larsen & Larsen, 2018). Med et forbedret sæt af kontrolvariabler kan vi forøge vores viden om den del af forskellen mellem kvinders og mænds timeløn, der hidtil har været 'uforklaret'. Dette giver et bedre udgangspunkt for at drøfte, hvad der skal til for at reducere forskellen mellem kvinders og mænds timeløn yderligere.

Således er udgangspunktet for denne undersøgelse det sæt af kontrolvariabler og datagrundlag, der har været anvendt i de tidligere SFI-/VIVE-undersøgelser. Hovedformålet har i de tidligere undersøgelser været at opgøre størrelsen af lønforskellen over en årrække og på at belyse, hvordan lønforskellen i de enkelte år hænger sammen med de to køns forskellige fordeling på en række målbare forhold som fx uddannelsesniveau, branche og arbejdsfunktion. Det ændrede sæt af kontrolvariabler vil potentielt kunne anvendes i tilsvarende, fremtidige undersøgelser af lønforskellen.

Vi giver med denne rapport vores bud på, hvordan et ændret sæt af kontrolvariabler kan se ud. Det er i denne forbindelse vigtigt at understrege, at der ikke er ét sæt af kontrolvariabler, der er entydigt bedre end alle andre. I udviklingen af det ændrede sæt af kontrolvariabler har vi løbende foretaget nogle valg. Andre valg vil formentlig have givet et andet resultat. Dertil kommer, at inddragelse af faktorer af betydning for lønforskellen ud over dem vi har valgt at medtage her, potentielt vil kunne bidrage til yderligere at forøge vores viden om lønforskellen.

1.2 Baggrund

I dette afsnit redegør vi først for, hvad den 'uforklarede' lønforskel er. Dernæst beskriver vi, hvor stor den 'uforklarede' lønforskel ifølge tidligere SFI-/VIVE-undersøgelser er på forskellige dele af arbejdsmarkedet.

Hvad er den 'uforklarede' lønforskel?

Når vi opgør forskellen mellem kvinders og mænds timeløn skelner vi mellem bruttolønforskellen¹, den 'forklarede' del af lønforskellen og den 'uforklarede' del² af lønforskellen. *Bruttolønforskellen* er den 'rå' forskel mellem den gennemsnitlige timeløn for kvinder og den gennemsnitlige timeløn for mænd, dvs. lønforskellen mellem de to køn, når vi ikke har taget højde for de to køns karakteristika. Den 'forklarede' del er den del af lønforskellen, der kan tilskrives, at kvinder og mænd er forskelligt fordelt på en række målbare forhold, fx uddannelse, erhvervs-erfaring, sektor, branche og arbejdsfunktion. Ved målbare forhold forstår vi her forhold, som kan måles vha. registerdata. Den 'uforklarede' del er differencen mellem bruttolønforskellen og den 'forklarede' del af lønforskellen.

Vi anvender en statistisk model (dekomponeringsanalyse) til at beregne, hvor stor en del af bruttolønforskellen, der hænger sammen med, at de to køn er forskelligt fordelt på målbare forhold, altså størrelsen af den såkaldt 'forklarede' del. På trods af brugen af ordet 'forklaret' giver resultaterne alene et billede af, om der er en *statistisk sammenhæng* mellem på den ene side de to køns fordeling på bl.a. uddannelse, erhvervs-erfaring, sektor, branche og arbejdsfunktion og på den anden side forskellen mellem kvinders og mænds timeløn.

Den 'uforklarede' del er et statistisk begreb, hvis størrelse afhænger af de målbare forhold, der indgår i analysen. Denne del af lønforskellen hænger sammen med:

- at kvinder og mænd med 'samme' karakteristika bliver aflønnet forskelligt, fx har samme uddannelse eller arbejdsfunktion, og
- uobserverede forhold, som analysen ikke har taget højde for.

Som eksempel på forskellig aflønning trods 'samme' karakteristika kan nævnes, at sammenhængen mellem timeløn og det at være småbørnsforælder kan være forskellig for kvinder og mænd. Simonsen & Skipper (2008) viser således, at forældreskab reducerer timelønnen for mødre og forøger timelønnen for fædre. En yderligere årsag er af mere 'teknisk' karakter: To personer, der er placeret i samme kategori for en given variabel, kan reelt have forskellige karakteristika. For eksempel kan to personer, der i data er kategoriseret i samme arbejdsfunktion, reelt have forskelligt jobindhold, der aflønnes forskelligt. I kontrast til de tidligere SFI-/VIVE-undersøgelser har vi i denne undersøgelse opgjort graden af sammenhæng mellem på den ene side kønsforskelle mht. aflønning af 'samme' karakteristika og på den anden side forskellen mellem kvinders og mænds timeløn. Dette kan give os mere viden om, hvad den 'uforklarede' lønforskel består i.

'Uobserverede forhold' omfatter forskelle mellem kvinder og mænd, som hænger sammen med lønforskellen, men som der hidtil ikke er taget højde for. Nogle af disse forskelle er målbare, mens andre ikke kan måles med afsæt i de registerdata, som er datagrundlaget for denne undersøgelse, se også afsnit 3.1. Som eksempler på (hidtil) uobserverede, målbare forhold kan nævnes, om en person har ledelsesansvar, samt antal ansatte på arbejdspladsen. Som eksempler på ikke-målbare forhold kan nævnes ikke-registeret over-/merarbejde, uformelle kvalifikationer af betydning for lønmodtageres produktivitet og konkrete forhandlingsprocesser på den enkelte arbejdsplads. Den 'uforklarede' del af lønforskellen reduceres, hvis vi tager højde for (hidtil uobserverede) målbare forhold af betydning for lønnen, der er forskelligt fordelt på kvinder og mænd.

¹ Også kaldet den 'ukorrigerede' lønforskel.

² Også kaldet den 'korrigerede' lønforskel.

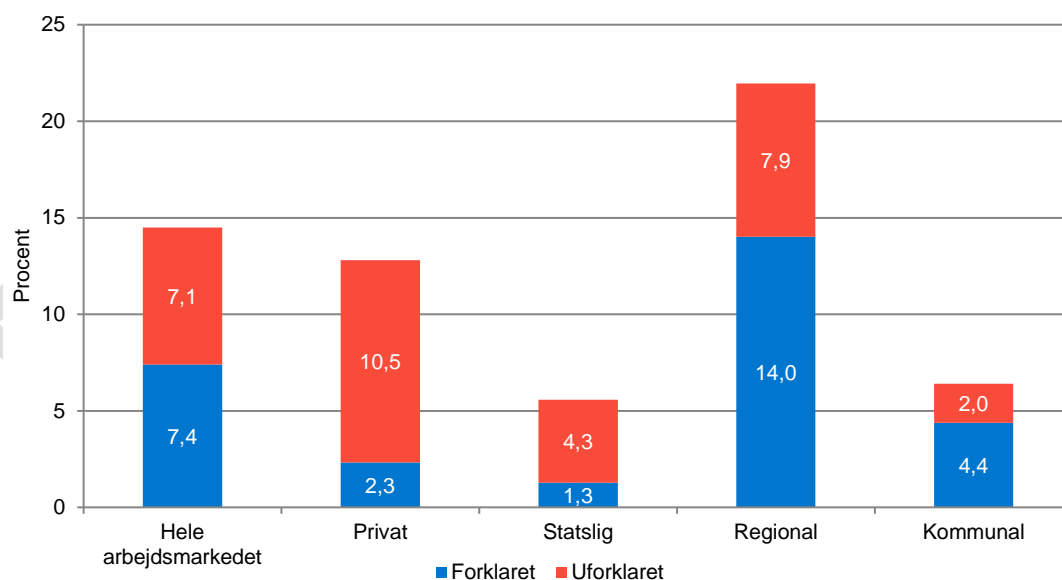
Det er vigtigt at understrege, at den 'uforklarede' lønforskel *ikke er et mål for*, om der finder forskelsbehandling sted på det danske arbejdsmarked. Resultaterne af de gennemførte analyser giver med andre ord ikke noget svar på, hvor stor en del af bruttolønforskellen, der er 'retfærdig' eller 'uretfærdig', herunder spørgsmålet om potentiel diskrimination. 'Retfærdige' lønforskelle, dvs. lønforskelle, som skyldes forhold, som alle er enige om har betydning for lønnen, kan for så vidt findes både i den 'forklarede' og den 'uforklarede' del. Og tilsvarende kan 'uretfærdige' elementer af lønforskellen optræde begge steder.

Tilsvarende siger analyseresultaterne ikke noget om, hvorvidt kvinder og mænd får samme løn for samme arbejde eller for arbejde af samme værdi, dvs. at resultaterne ikke giver et billede af, om der er ligeløn på det danske arbejdsmarked i juridisk forstand. Resultaterne *afspejler* heller *ikke de bagvedliggende årsager* til lønforskellen: Resultaterne giver os ikke en forklaring på, hvorfor de to køn fx er forskelligt placeret på arbejdsmarkedet, eller hvorfor timelønnen er forskellig på forskellige dele af arbejdsmarkedet. Resultaterne i denne rapport kan dermed alene *danne udgangspunkt* for en drøftelse af, hvad der forklarer forskellen mellem de to køns timeløn.

Størrelsen af den 'uforklarede' lønforskel

Ifølge et tidligere VIVE-studie er bruttolønforskellen på 14,5 pct. i 2016 baseret på standardberegnet timefortjeneste, se Larsen & Larsen (2018). Det fremgår også, at lige over halvdelen af lønforskellen kan tilskrives kønsforskelle i fordelingen på uddannelse, erhvervs erfaring, sektor, branche, arbejdsfunktion, arbejdstid, familieforhold og bopæl, se Figur 1.1. Det vil sige, at lige knap halvdelen af lønforskellen er 'uforklaret' ifølge de tidligere studier.

Figur 1.1 Bruttolønforskel mellem kvinder og mænd for hele arbejdsmarkedet og særsklit for sektorer, opdelt på 'forklaret' og 'uforklaret' del, korrigerede tal¹, 2016. Opgjort for standardberegnet timefortjeneste. Procent.



Note: ¹ For korrigerede tal, se boks 3.2 og Larsen & Larsen (2018).

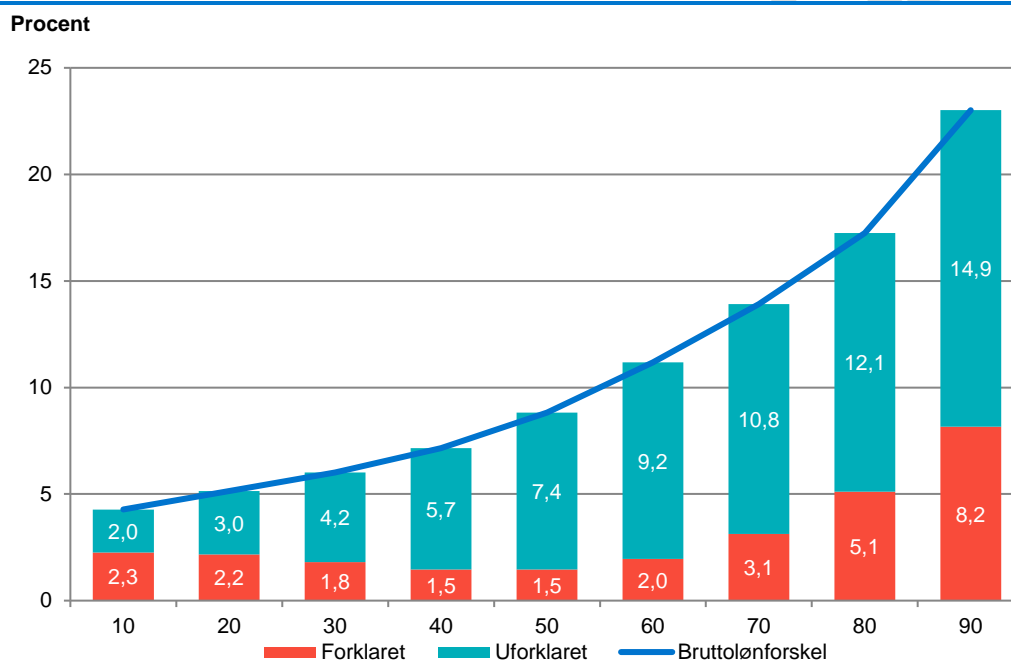
Kilde: Larsen & Larsen (2018). Baseret på tal fra figur 6.1, 7.1, 7.3, 7.5 og 8.1.

Lønforskellen er størst i regioner og mindst i stat og kommuner. Det tidligere VIVE-studie viser baseret på tal for 2016, at den 'uforklarede' del er størst i den private sektor, hvor 10,5 pct. af

en lønforskel på 12,8 pct. (svarende til ca. 80 pct. af lønforskellen) er 'uforklaret' – en sektor, der samtidig omfatter flertallet af de ansættelser, der indgår i denne undersøgelse.

Bruttolønforskellen mellem kvinder og mænd stiger ifølge samme VIVE-studie med stigende løn, se Figur 1.2. Dette studie tyder på, at den uforklarede del absolut set er størst blandt de højtlojnnede og mindst blandt lavtlønnede³. Se vi derimod på, hvor stor en andel af lønforskellen, der er 'uforklaret' i det tidligere VIVE-studie, finder vi, at 47 pct. er 'uforklaret' i bunden af lønfordelingen (ved 10-percentilen), 84 pct. 'uforklaret' i midten af lønfordelingen (ved 50-percentilen), mens 65 pct. er 'uforklaret' i toppen af lønfordelingen (ved 90-percentilen). Den største uforklarede andel finder vi med andre ord midten af lønfordelingen.

Figur 1.2 Bruttolønforskel mellem kvinder og mænd på arbejdsmarkedet som helhed, særskilt for percentiler¹, i alt og fordelt på 'forklaret' og 'uforklaret' del, 2016. Opgjort for standardberegnet timefortjeneste. Procent.



Note: ¹ Opdelingen af timelønnen i percentiler er foretaget for kvinder og mænd hver for sig. Det vil sige, at beregningen for fx 10. percentil tager afsæt i gennemsnitslønnen ved hhv. kvinders og mænds 10. percentil.

Kilde: Larsen & Larsen (2018). Tilrettet version af figur 6.2.

³ Det har dog ikke været muligt for at inddrage så detaljerede oplysninger om branche og især arbejdsfunktion i denne analyse som i den analyse, der ligger bag resultaterne i Figur 1.1 (se også Larsen & Larsen, 2018, s. 66).

2 Hvad kan reducere den 'uforklarede' lønforskel?

Litteraturen om forskelle mellem kvinders og mænds timeløn er ganske omfattende. I dette afsnit skitserer vi nogle af de faktorer, der ifølge nyere litteratur kan bidrage til at forøge vores viden om lønforskellen i Danmark. Gennemgangen giver os ideer til, hvilke variabler vi kan inkludere i vores analysedesign for at reducere størrelsen af den 'uforklarede' lønforskel. Desuden giver gennemgangen os mere viden om de faktorer af betydning for lønforskellen, som det ikke er muligt at tage højde for fx pga. begrænsninger i data. Det er vigtigt at understrege, at gennemgangen ikke skal ses som en fuldstændig oversigt over den nyeste litteratur.

2.1 Segregering på arbejdsmarkedet

Det er velkendt, at det danske arbejdsmarked er kønsopdelt (se fx Larsen et al. 2016). Det vil sige, at kvinder og mænd i stor udstrækning befinder sig på forskellige dele af arbejdsmarkedet. Segregeringen er både horisontal, dvs. at de to køn er forskelligt fordelt på sektorer og brancher, og vertikal, hvilket viser sig ved, at mænd er højere placeret i stillingshierarkiet end kvinder. Lønforskellen mellem kvinder og mænd hænger i væsentligt omfang sammen med den markante kønsmæssige segregering.

Der er flere bagvedliggende mekanismer, der kan have betydning for, at kvinder og mænd arbejder på forskellige dele af arbejdsmarkedet. Det er fx præferencer, evner, personlighedstræk, uddannelsesvalg m.v. Jo flere af disse faktorer, der kan inddrages i analysen, jo mere af lønforskellen kan vi forvente at kunne forklare.

Uddannelse

I tidligere undersøgelser er der typisk taget højde for uddannelseslængde. I følge Blau & Kahn (2017) er det dog sandsynligt, at lønforskellen er forskellig for forskellige uddannelser, selvom der er tale om uddannelser med samme længde. Det kan således spille en rolle, at kvinder og mænd traditionelt vælger forskellige fag, fx er der væsentlig færre kvinder end mænd i "STEM"-fagene (Science, Technology, Engineering, Mathematics). Derfor er det relevant at undersøge, om oplysninger om uddannelsesretning vil bidrage til at forklare en større del af lønforskellen.

I Blau & Kahn (2017) beskrives desuden, at kønsforskelle i matematik kompetencer har betydning for lønforskellen. Fx finder Joensen & Nielsen (2017) i et studie for Danmark, at kvinders løn påvirkes positivt, hvis de vælger gymnasielinjer med højniveau matematik.

Kvindeandele

Kønssammensætningen varierer imellem forskellige typer af job, sektorer, brancher og virksomheder, og således varierer andelen af kvinder i forskellige arbejdsfunktioner markant. Albæk et al. (2016, 2019) viser, at der er en negativ sammenhæng mellem lønnen og andelen af kvinder i forskellige arbejdsfunktioner – en sammenhæng, der især viser sig i den offentlige, men også i den private sektor. Det er beregnet, at et fag, der har en 10 procentpoint større kvindeandel end et andet fag, har i gennemsnit en løn, der er 4,38 kr. lavere i den offentlige sektor og 2,52 kr. lavere i den private sektor.

Der skelnes mellem tre mulige mekanismer til forklaring af denne sammenhæng:

- Forskel i kvindeandel inden for fag – altså kønssegregering,
- forskel i gennemsnitlig uddannelseslængde mellem fag, og
- forskel i kvinders og mænds gennemsnitsløn uanset fag.

Ifølge Albæk et al. (2019) er det de to førstnævnte mekanismer, der har størst betydning for den fundne lønforskel i både den offentlige og den private sektor.

Albæk et al. (2019) beskriver mulige forklaringer på at fag med høj kvindeandel er lavt lønnet. En forklaring på, at kvindefag i den offentlige sektor er lavtlønsfag er, at lønningerne i den offentlige sektor ved en reform af lønsystemet i den offentlige sektor for år tilbage blev fastlåst i et bestemt forhold, hvor kvindefagene fik tildelt relativt lavere lønninger end mandefagene. Efterfølgende overenskomstforhandlinger skulle så have reproduceret disse fastlåste lønrelationer. En anden forklaring på, at kvindefag er lavtlønsfag i den offentlige sektor, kan være, at nogle kvindefag i perioder i større omfang end mandefag har forhandlet sig til goder frem for løn, se Borchorst (2010). En tredje forklaring på kvindefagenes relativt lave lønninger i den offentlige sektor kan være eksistensen af købermonopol (også kaldet monopson) i kvindefagene. Købermonopol indebærer, at der alene er én køber af fagets arbejdskraft, og at denne ene køber dermed har stor mulighed for at påvirke lønnen (i nedadgående retning).

De anførte forklaringer på, at kvindefag i den offentlige sektor er lavtlønsfag, kan ikke umiddelbart overføres til den private sektor. Det er dog velkendt, at løndannelsen foregår mere decentralt i den private end i den offentlige sektor. En af forklaringerne på, at kvindefag er lavtlønsfag i den private sektor, kan være, at fordi løndannelsen i den private sektor foregår mere decentralt er kvinder mindre tilbøjelige til at forhandle løn, og at de desuden ikke er helt så konkurrenceorienterede som mænd, se Blau & Kahn (2017). Endvidere viser forskellige studier, at kvinder i gennemsnit er mere risikoaverse end mænd. Indtjeningen i nogle fag, specielt inden for den private sektor, er mere varierende end i andre fag, og for at kompensere for denne usikkerhed og dermed tiltrække arbejdskraft kan den gennemsnitlige aflønning være højere i disse fag end i fag med mere stabil indtjening. Tankegangen tilsiger således, at kvinder i højere grad end mænd vælger fag med stabil indtjening, men altså også med en lavere gennemsnitlig indtjening.

Præferencer

Wiswall & Zafar (2018) viser, at kvindelige universitets-/collegestuderende har stærke præferencer for at arbejde i job, der er kendetegnet ved større fleksibilitet og stabilitet. Studiet undersøger de studerendes 'willingness to pay' for disse jobkarakteristika og finder, at de kvindelige studerende er villige til at acceptere at få lavere lønstigninger i job med sådanne karakteristika. En opfølgende spørgeskemaundersøgelse viser, at denne form for kønsforskelle i præferencer forklarer mindst 25 pct. af lønforskellen 4 år efter, at de deltagende studerende dimitterer.

2.2 Firmaspecifikke effekter

Ud over segregeringen mellem sektorer og brancher viser en række relativt nye undersøgelser, at forskelle mellem aflønningen i forskellige virksomheder har betydning for lønforskellen. Sådanne analyser kan laves, når data gør det muligt at koble oplysninger om den enkelte ansatte med oplysninger om den virksomhed, hvor den pågældende er beskæftiget (kaldet 'matched employer-employee' data). Card et al. (2016) finder i et studie for Portugal, at virksomhederne

adskiller sig fra hinanden, når det gælder, om de aflønner relativt lavt eller højt – og at kvinder har større sandsynlighed for at være ansat i de virksomheder, der aflønner relativt lavt. I den enkelte "højt-løns virksomhed" er der desuden en tendens til, at de ansatte kvinder får en mindre del af "virksomhedspræmien" end mænd, dvs. at de får et lavere lønafkast af deres kvalifikationer.

Gallen et al. (2019) finder i lignende analyser for Danmark, at kvinder (især mødre) har en tendens til at være ansat i virksomheder, der er karakteriseret ved relativt lavt timetal for de ansatte. Dette reducerer deres samlede indtjening – både pga. det lavere timetal og fordi de generelt har lavere timeløn.

Den lavere aflønning af kvinder i virksomhederne tilsiger, at der er behov for nærmere undersøgelser af løndannelsen på virksomhedsniveau. Theodoropoulos et al. (2019) finder, at lønforskellen er signifikant lavere i virksomheder med en høj andel af kvindelige ledere. Dette skyldes både, at kvinders løn er relativt høj og at mænds løn er relativt lav. Denne tendens forstærkes, når de ansatte får resultatløns.

2.3 Effekter af jobskifte

Ifølge Blau & Kahn (2017) viser ældre studier, at kvinder er mere tilbøjelige end mænd til at forlade jobbet af familierelaterede årsager (eller til at overgå til ikke-beskæftigelse), mens mænd i større omfang fratræder af jobrelaterede årsager. Sådanne kønsforskelle kan bidrage til, at et jobskifte er mere gunstigt lønmæssigt for mænd end for kvinder. Der eksisterer ikke nyere studier af, om der er kønsforskelle, når det gælder årsagerne til jobskifte på det danske arbejdsmarked.

Frederiksen, Halliday & Koch (2016) viser med afsæt i danske data for den private sektor, at den type jobskifte, der giver størst lønstigning og størst kønsforskel i lønstigningen, er en forfremmelse fra ikke-topleder til topleder på tværs af firmaer. Frederiksen, Halliday & Koch (2016) belyser også dels jobskifte på samme stillingsniveau (topleder eller ikke-topleder) mellem firmaer, dels jobskifte mellem ikke-topleder og topleder inden for samme firma, dvs. interne forfremmelser/degraderinger. For begge disse typer af jobskifte gælder, at der ikke er nævneværdig forskel på de lønændringer, som kvinder og mænd oplever i forbindelse med jobskiftet. Mænd oplever tre gange oftere end kvinder forfremmelse mellem firmaer. Jobskifte på samme stillingsniveau mellem firmaer og firmainterne jobskifte mellem ikke-topleder og topleder foretages også oftere af mænd end af kvinder. Resultaterne i Frederiksen, Halliday & Koch (2016) tyder alt i alt på, at jobskifte inden for den private sektor medvirker til, at lønforskellen mellem kvinder og mænd stiger med alderen.

2.4 Effekter af nedsat arbejdstid

Kvinder har i større omfang end mænd nedsat arbejdstid. Hvis nedsat arbejdstid aflønnes lavere pr. time end fuldtidsarbejde, kan det ifølge Blau & Kahn (2017) skyldes:

1. Forskelle med hensyn til humankapital
2. Omkostninger forbundet med tidsmæssig fleksibilitet
3. Signalerings/sortering.

Ad 1. En lavere timeløn til beskæftigede med nedsat arbejdstid kan skyldes, at både lønmodtagerne og deres arbejdsgivere har mindre incitamenter til at investere i humankapitel i form af oplæring på jobbet, når det gælder job på nedsat tid.

Ad 2. Der kan i nogle tilfælde være omkostninger forbundet med tilvejebringelse af tidsmæssig fleksibilitet, og disse omkostninger er forskellige for forskellige sektorer og virksomheder (Goldin, 2014). Dette har især konsekvenser for lønnen for højtuddannede kvinder med præferencer for tidsmæssig fleksibilitet, jf. også afsnittet om præferencer ovenfor. Den negative løneffekt forbundet med tidsmæssig fleksibilitet er oftest stor i den type jobs, hvor der er krav om at opfylde deadlines, at være i kontakt med andre for at bestride jobbet, at opretholde og etablere relationer med det øvrige personale, at overholde eksisterende tidsplaner og at udføre arbejde, hvor der ikke er kollegaer, der umiddelbart kan overtage opgaverne – dette gælder fx jobs i den finansielle sektor og i advokatvirksomheder.

Ad 3. Mange arbejdstimer kan signalere en større villighed til at arbejde hårdt samt større motivation og engagement – signaler, der kan anvendes af arbejdsgivere til at sortere i forhold til disse egenskaber. Ifølge Simon, Sanroma & Ramos (2017) kan man med afsæt heri argumentere for, at den negative løneffekt forbundet med nedsat arbejdstid kan være større for mænd end for kvinder. Kønsforskellen opstår, fordi mænd i større udstrækning end kvinder bliver udsat for stigmatisering, når de – frivilligt eller ufrivilligt – har nedsat arbejdstid: Kvinders nedsatte arbejdstid betragtes oftere som forventeligt af familiære årsager, mens mænds nedsatte arbejdstid oftere betragtes som et resultat af inkompetence eller manglende motivation.

Deltidsbeskæftigelse kan dog i nogle tilfælde aflønnes *højere* end fuldtidsbeskæftigelse (se fx Hirsch, 2005). Situationen kan opstå, hvis arbejdspladser har en særlig præference for deltid-beskæftigelse (fx restauranter med spidsbelastning omkring spisetiderne), og hvis de lønmodtagere, der efterspørges, foretrækker kontinuert fuldtidsarbejde frem for korte eller opdelt skift.

2.5 Effekter af fravær

Perioder med fravær fra jobbet eller arbejdsmarkedet i det hele taget kan være årsag til faldende løn (eller lavere lønstigninger). En forklaring på, at fravær forårsager lavere løn er ifølge fx Albrecht et al. (1999) og Nielsen et al. (2003), at humankapital afskrives under fraværsperioder, og jo længere tid man er fraværende, jo større bliver den samlede afskrivning. En anden forklaring er, at længere (frivillige) fraværsperioder kan signalere mindre dedikation til sit job. Sidstnævnte argument har fx været brugt som forklaring på, at 'lønstraffen' for at tage lang barsel er større for mænd end for kvinder, se også afsnittet om effekter af nedsat arbejdstid.

Andersen (2018) analyserer effekten af fædres andel af parrets samlede afholdte barselsorlov for husholdningens lønindkomst. Resultatet viser, at lønforskellen inden for husholdningen mindskes, når faderen tager en større andel af barslen og samtidig øges husholdningens totale lønindkomst.

2.6 Effekter af børn

En række studier har igennem tiden vist en negativ sammenhæng mellem kvinders løn og børnefødsler. Dette kan skyldes, at mødre har en anden adfærd end ikke-mødre eller at mødre forventes at have en anden adfærd og evt. er mindre produktive, når de er på arbejde.

Det seneste danske studie af sammenhængen mellem løn og børnefødsler finder 'child-penalties' på ca. 9 pct. (i forhold til mænd) for danske kvinders timeløn i de første 10 år efter fødslen af det første barn (Kleven et al., 2019). Effekten er af permanent karakter, idet den ikke formindskes i løbet af de næste 10 år, men faktisk forøges yderligere. Studiet viser, at forklaringen på dette bl.a. skal findes i kønsforskelle i stillingstyper, sektor og firmatypevalg. Således findes der altså en betydelig sammenhæng mellem børn og den kønsmæssige segregering på arbejdsmarkedet. Dekomponeringen i studiet viser desuden, at mens den samlede kønsforskel i indtjening er blevet halveret i perioden 1980-2012, har den børnerelaterede kønsforskel i indtjening været konstant og dermed står den børnerelaterede kønsforskel for næsten hele kønsforskellen i indtjening.

2.7 Personlighedstræk

En række nyere studier undersøger sammenhængen mellem forskelle i personlighedstræk og lønforskellen. I følge Blau & Kahn (2017) har kønsforskelle i personlighedstræk en moderat betydning for lønforskellen. Flinn et al. (2017) finder vha. en teoretisk model, at personlighedstræk påvirker valget af fordelingen af husarbejde og ressourcefordeling. Personlighedstrækene, der måles ud fra Big Five⁴, har stor betydning for forhandlingsstyrke og mulig løn og dermed for lønforskellen.

Schäfer & Schwiebert (2018) undersøger vha. tyske data, hvordan personlighedstræk påvirker udviklingen i lønforskellen mellem kvinder og mænd. Resultaterne viser, at kønsforskelle i følelsesmæssig stabilitet og bevidsthed ('consciousness') øger lønforskellen, mens forskelle i udadvendthed indsnævrer lønforskellen.

Lesner (2019) undersøger vha. danske registerdata, hvorledes kønsforskelle i personlighedstræk, der er påkrævet i forskellige typer af stillinger, hænger sammen med lønforskellen mellem kvinder og mænd. Resultaterne viser, at der er en betydelig sammenhæng mellem sådanne jobkarakteristika og lønforskellen, specielt er der et stort afkast til mænd med job, hvor "ubehagelighed" ('disagreeableness') er påkrævet. Risse et al. (2018) finder ligeledes i et studie fra Australien, at mænds lavere frygt for fiasko og lavere niveau af 'agreeableness' bidrager til lønforskellen, mens kvinders højere niveau af bevidsthed og nærvær indskrænker lønforskellen. Collischon (2018) udfører lignende analyser for Tyskland, UK og Australien, men ved hjælp af quantile regressioner, således at effekten kan variere for forskellige lønniveauer. Resultatet viser, at personlighedstræk har størst betydning i den øverste del af lønfordelingen.

2.8 Ubevidst bias

I litteraturen er der en øget bevidsthed om begrebet "ubevidst bias" og dette som årsag til manglende kønsmæssig ligestilling på arbejdsmarkedet. Ubevidst bias kan opstå, når mennesker har automatiske associationer, der hjælper til at simplificere fordøjelsen af større mængder information ved at generere "kognitive genveje". Når sådanne "kognitive genveje" betyder, at man har tendens til at klassificere mennesker fx ud fra deres køn eller hudfarve kaldes det ofte "stereotyper". Ubevidst bias formes over tid ud fra menneskelige erfaringer og kulturelle kontekster. Et klassisk eksempel på at kønsmæssige stereotyper eksisterer, er et studie, hvor børn bedes om at tegne en forsker (Miller et al. 2018) og de langt oftere tegner mandlige end

⁴ Elementerne i 'Big-five' er ekstraversion (indadvendt-udadvendt), neuroticisme (bekymret-rolig), åbenhed (konventionel-original), samvittighedsfuldhed (upålidelig-trofast) og venlighed (irritabel-godmodig).

kvindelige forskere. På arbejdsmarkedet kan ubevidst bias således resultere i kønsmæssig segregering (både vertikalt og horisontalt) og lønforskelle. Humlum et al. (2018) finder desuden, at den kønsmæssige segregering på arbejdsmarkedet "går i arv", idet den overføres mellem generationer. Dette skyldes, at pigers uddannelsesvalg er korreleret med deres mors arbejdsmarkedsadfærd, mens drenges uddannelsesvalg er korreleret med deres fars arbejdsmarkedsadfærd.

2.9 Hvad tager vi højde for?

De karakteristika, der er blevet taget højde for i de tidligere SFI-/VIVE-undersøgelser af lønforskellen, er uddannelse, erhvervs erfaring, sektor, branche, arbejdsfunktion, arbejdstid, familieforhold og bopæl.

I de tidligere SFI-/VIVE-undersøgelser indgår oplysninger om uddannelse og erhvervs erfaring som mål for den enkeltes humankapital. Uddannelsesniveau har været anvendt som mål for uddannelse (i form af 10 dummyvariable), se også afsnit 3.2. Vi har i forlængelse af ovenstående valgt at undersøge, om sammenhængen mellem lønforskellen og uddannelse forstærkes, hvis vi i stedet tager højde for både uddannelsesretning og -længde. I de tidligere undersøgelser har opgørelsen af erhvervs erfaring alene været baseret på ATP-indbetalinger siden 1964. Som noget nyt findes der nu opgørelser af erhvervs erfaring baseret på løntimer fra og med 2008, hvilket er en mere præcis opgørelsesmetode. Vi har valgt at erstatte den tidligere opgørelse af erhvervs erfaring med en opgørelse, der er baseret på ATP-indbetalinger fra 1964 til 2007 og på løntimer fra og med 2008.

I de tidligere undersøgelser er dummyvariable for sektor, branche og arbejdsfunktion anvendt som mål for det kønsopdelte arbejdsmarked, se også afsnit 3.2. Vi har i forlængelse af ovenstående valgt at undersøge, om vi får mere markant sammenhæng mellem lønforskellen og det kønsopdelte arbejdsmarked, hvis vi erstatter dummyvariable for sektor, branche og arbejdsfunktion med oplysninger om kvindeandele inden for sektorer, brancher, arbejdsfunktioner og arbejdsfunktioner på arbejdssteder (her kaldet 'jobceller'). Den bagvedliggende hypotese er, at lønforskellen hænger mere sammen med, hvor mange af de ansatte inden for et givet område (sektor, branche, arbejdsfunktion eller jobcelle), der er kvinder, end hvilket område, der specifikt er tale om. Ved at kontrollere for aflønningen inden for jobceller tager vi desuden – i modsætning til tidligere – højde for, hvad der foregår på arbejdsstedsniveau.

Vi har desuden valgt at undersøge, om tilføjelse af oplysninger om jobstatus også kan forstærke sammenhængen mellem lønforskellen og det kønsopdelte arbejdsmarked. 'Jobstatus' er et mål for placeringen i stillingshierarkiet, da denne variabel gør det muligt at skelne mellem elever/lærlinge, ledere/mellemledere, medarbejdere med særligt ansvar og øvrige medarbejdere. Vi inddrager dermed et mål for den vertikale kønsopdeling på arbejdsstedsniveau.

Endelig har vi i lighed med en række andre studier (se fx Theodoropoulos et al., 2019) valgt at tage højde for arbejdsstedsniveauet ved at kontrollere for antallet af ansatte på arbejdsstedet. En hypotese kan være, at lønforskellen er mindre på større virksomheder, fordi løndannelsen her er mere systematisk, hvilket potentielt kan bidrage til, at (uhensigtsmæssige) lønforskelle mellem de to køn undgås.

Der har ikke været taget højde for jobskifte i de tidligere SFI-/VIVE-undersøgelser. En medvirkende årsag er givetvis, at det ikke er enkelt at opgøre jobskift i et datasæt, hvor hver person kan indgå med mere end én ansættelse i året, se afsnit 3.1. Vi har valgt at undersøge, om

lønforskellen hænger sammen med antallet af andre ansættelser i de forudgående år – en opgørelse, som samtidig siger lidt om anciennitet på det aktuelle arbejdssted.

Som mål for nedsat arbejdstid har de tidligere SFI-/VIVE-undersøgelser alene skelnet mellem, om man er ansat mindre end 30 timer pr. uge eller ej. Vi har valgt at undersøge, om inddragelse af en mere detaljeret opdeling af det ugentlige antal arbejdstimer betyder, at sammenhæng mellem lønforskellen og arbejdstid bliver mere udtalt.

Kontrol for arbejdstid er samtidig en måde at tage højde for nogle kvinders præference for tidsmæssig fleksibilitet. En anden måde er at tage højde for overarbejde. Desværre har vi alene oplysninger om betalt overarbejde.⁵ Vi har valgt at undersøge, om betalt overarbejde hænger sammen med lønforskellen.

I de tidligere SFI-/VIVE-undersøgelser har der ikke været taget højde for perioder med fravær fra jobbet eller arbejdsmarkedet. Vi har valgt at undersøge, om der er en sammenhæng mellem lønforskellen og dels fravær i det aktuelle job, dels henholdsvis sygefravær, barselsfravær og ledighed i de forudgående år.

Det fremgår af afsnit 3.2, hvordan vi har valgt at operationalisere de nye kontrolvariabler.

2.10 Hvad tager vi ikke højde for?

Der er en lang række forhold af potentiel betydning for lønforskellen mellem kvinder og mænd, som vi ikke tager højde for. I nogle tilfælde skyldes det, at der ikke findes registeroplysninger om disse forhold. I andre tilfælde findes der registeroplysninger på de pågældende forhold, men de kan ikke uden videre kobles med de oplysninger om ansættelser fra lønstatistikken, som vi arbejder med her, se afsnit 3.1.

Vi kan ikke tage højde for karakterer i matematik (eller andre skolekarakterer), fordi denne type registeroplysninger kun findes fra og med 2001. Det vil sige, at vi ikke har oplysninger om karakterer for en stor del af de personer, der indgår i vores undersøgelse, da de afsluttede deres skolegang før 2001, se afsnit 3.1.

Der findes i sagens natur ikke registeroplysninger om præferencer, evner, personlighedstræk og forekomst af ubevidst bias. Vi kan fx ikke kontrollere for kønsforskelle i præferencer for fx tidsmæssig fleksibilitet ud over det, vi kan indfange ved brug af oplysninger om den kontraktmæssigt aftalte arbejdstid og det betalte overarbejde. I denne forbindelse mangler vi også oplysninger om ubetalt overarbejde.

I lønstatistikken findes der kun et begrænset antal oplysninger på arbejdsstedsniveau. Der findes arbejdsstedsrelaterede oplysninger i andre registre, fx oplysninger om anciennitet og pendlingsafstand mellem arbejde og bopæl, som ikke (direkte) kan kobles med ansættelserne i lønstatistikken. Forekomsten af flere ansættelser i året gør det også vanskeligt at udarbejde et godt mål for jobskift ud over en opgørelse af antallet af andre ansættelser i de forudgående år.

⁵ Mænd ser ud til at have mere ubetalt overarbejde end kvinder (Danmarks Statistik, 2017). Lønforskellen mellem kvinder og mænd kan i givet fald bl.a. hænge sammen med kønsforskelle mht. ubetalt overarbejde, fordi den beregnede gennemsnitlige timeløn i så fald er for høj for især mænd i lyset af de timer de reelt arbejder.

3 Datagrundlag og analysedesign

I det følgende har vi beskrevet rapportens datagrundlag og det anvendte analysedesign.

3.1 Datagrundlag

Vores analyser er udarbejdet med afsæt i Danmarks Statistiks lønstatistik. Denne statistik er velegnet til analyser af lønforskellen mellem kvinder og mænd, fordi den indeholder ret præcise opgørelser af timeløn. Ved at bruge lønstatistikken sikrer vi kontinuitet i forhold til de tidligere SFI-/VIVE-undersøgelser af forskellen mellem kvinders og mænds timeløn. Lønstatistikken er beskrevet nærmere i det følgende.

3.1.1 Lønstatistikken

For at beregne og analysere lønforskellen mellem kvinder og mænd anvender vi registeroplysninger om løn og ansættelsesforhold fra Danmarks Statistiks lønstatistik. Oplysningerne fra lønstatistikken er koblet med registeroplysninger fra Befolkningen, IDA personer, Familieforhold, Uddannelser samt Sygedagpenge.

Vores analysepopulation består af ansættelsesforhold, som er statistikenheden i lønstatistikken. *Et ansættelsesforhold er defineret som en ansat hos en bestemt arbejdsgiver, på et bestemt arbejdssted, med en bestemt arbejdsfunktion.* At statistikenheden er det enkelte ansættelsessted indebærer i praksis, at personer med flere ansættelsesforhold i løbet af et år indgår flere gange i datamaterialet for det pågældende år. Vi arbejder med uvægtede data, dvs. at ansættelsesforhold af forskellig varighed indgår med samme vægt i vores beregninger. Med andre ord indgår timeløn fra fx en ansættelse på nedsat tid på 2 måneder med samme vægt som en fuldtidsansættelse, der varer hele året.^{6,7}

Lønstatistikken omfatter ansættelsesforhold for følgende grupper af lønmodtagere (Danmarks Statistik, 2020a, 2020b), også kaldet *målpopulationen* for Lønstatistikken:

- Lønmodtagere i den private sektor^{8,9} på nær landbrug og fiskeri, som er ansat i virksomheder med en beskæftigelse svarende til 10 eller flere fuldtidsbeskæftigede.
- Lønmodtagere i den offentlige sektor bortset fra vederlagslønnede, særligt aflønnede¹⁰, værnepligtige, ph.d.-studerende (da der ikke indberettes fravær for disse), visse time-lærere og studentermedhjælpere.

Ansættelsesforhold af meget kort varighed eller med ganske få ugentlige timer og ansættelsesforhold, hvor det ikke er muligt at afgøre antallet af mulige arbejdstimer, indgår ikke. Lønstatistikken omfatter kun lønmodtagere ansat på "normale vilkår". Ansættelsesforhold for følgende grupper af lønmodtagere indgår derfor heller ikke:

⁶ Af Larsen (2010) fremgår, at vægtning af data kun i begrænset omfang har betydning for størrelsen af de beregnede lønforskelle baseret på standardberegnet timefortjeneste (som er det lønbegreb, der anvendes her).

⁷ Lønstatistikken indeholder vægten OPRANDEL, som korrigerer for bortfaldet i den private sektor. Vi har undersøgt, om brug af denne vægt påvirker dekomponeringsresultaterne for hhv. hele arbejdsmarkedet og den private sektor, jf. kapitel 5. Det viser sig ikke at være tilfældet.

⁸ Svarende til sektoren 'virksomheder og organisationer'.

⁹ Den overvejende andel af ansatte i fleksjob på normale vilkår indgår i statistikken for denne sektor. Dette påvirker ikke timelønnen, da de netop er ansat på normale vilkår. Det er ikke muligt at udskille disse ansatte via indberetningerne til statistikken.

¹⁰ Herunder personer i fleks- og skånejob.

- Lønmodtagere der aflønnes efter en usædvanlig lav sats som følge af handicap eller lignende
- Lønmodtagere, der er rent provisionsaflønnet
- Lønmodtagere der ikke er beskattet efter de almindelige betingelser i Danmark
- Udlændinge, der arbejder i Danmark, men beskattes efter hjemlandets regler
- Udstationerede danskere, der aflønnes efter lokale regler.

Fra Danmarks Statistik har vi fået oplyst, at målpopulationen for Lønstatistikken, dvs. den del, der er forpligtet til at indberette, dækker ca. 85 pct. af lønmodtagerne på arbejdsmarkedet som helhed i 2016. Vores omtale af analyser for "hele arbejdsmarkedet" dækker med andre ord reelt ca. 85 pct. af lønmodtagerne. Stort set hele det offentlige arbejdsmarked indgår, mens private virksomheder med mindre end 10 fuldtidsansatte og virksomheder inden for landbrug og fiskeri ikke indgår, jf. ovenfor.

Lønstatistikens dækningsgrad *inden for målpopulationen* er tæt på 100 pct. for den offentlige sektor. En del af indberetningerne fra de private virksomheder har ikke kunnet anvendes pga. fejl, hvorfor dækningsgraden for den private sektor i ifølge oplysninger fra Danmark Statistik er på 90 pct. i 2016. Dækningsgraden varierer afhængigt af branche. For eksempel er dækningsgraden for branchen industri, råstofindvinding og forsyningsvirksomhed på 99 pct., mens den er på ca. 82 pct. for branchen handel og transport m.v. Danmarks Statistik har udarbejdet en vægt, der korrigerer for bortfaldet i den private sektor. Vi har undersøgt, om brug af denne vægt påvirker de endelige dekomponeringsresultater i kapitel 5 for hhv. hele arbejdsmarkedet og den private sektor (resultater ikke vist). Det viser sig ikke at være tilfældet.

For at sikre sammenlignelighed med de tidligere SFI-/VIVE-undersøgelser af forskellen mellem de to køns timeløn har vi afgrænset data for de enkelte år på følgende måde:

- Kun ansættelsesforhold for personer, der er bosat i Danmark både i starten og i slutningen af året er medtaget.
- Kun ansættelsesforhold for personer i alderen 25-59 år indgår
- Ansættelsesforhold med meget små (under 70 kr.) eller meget store timelønninger (10.000 kr. eller derover) er frasortet.
- Doubletter, det vil her sige observationer, hvor værdierne for alle variabler er de samme, er frasortet.

Den førstnævnte frasortering er medtaget for at sikre, at vi har oplysninger for både det anvendte lønbegreb og alle de forklarende variabler, der indgår i analysen. Da der indgår en række nye variabler i denne rapport, er betingelsen samtidig medvirkende til, at antallet af observationer, der indgår i analyserne for de enkelte år, afviger fra observationsantallet i de tidligere SFI-/VIVE-undersøgelser.¹¹

Af Larsen & Larsen (2018) fremgår, at de nævnte frasorteringer for hvert af årene 2012-2016 betød, at 75-79 pct. af Lønstatistikens observationer blev anvendt. Det fremgår også, at langt størstedelen af frasorteringerne kan tilskrives aldersafgrænsningen.

¹¹ Den største forskel viser sig for 2010, hvor der indgår op mod 31.000 flere observationer i denne rapport end i Larsen & Houlberg (2013). Afvigelsen kan primært tilskrives, at der i 2010-rapporten blev gennemført beregninger for lønbegrebet 'smalfortjeneste inkl. pension og personalegoder', som var uoplyst for en del observationer i lønstatistikken. For 2013 og 2016 er forskellen markant mindre: I denne rapport indgår 212 færre observationer for 2013 og 946 flere observationer for 2016 end i Larsen & Larsen (2018).

3.1.2 Lønbegreb

Opgørelserne af lønforskellen mellem kvinder og mænd er baseret på oplysninger om timeløn. En timeløn er ikke et eksakt begreb, hvilket bl.a. hænger sammen med, at løn både er en indtægt og en omkostning. Vi har i denne rapport valgt at anvende et mål for indtægt¹² som vores mål for timeløn, nemlig lønbegrebet *standardberegnet timefortjeneste*.

Standardfortjenesten er ifølge Danmarks Statistik (2011) det fortjeneste-begreb, der kommer tættest på det, der er aftalt mellem lønmodtager og arbejdsgiver. Af Boks 3.1 fremgår de lønelementer, der indgår i standardfortjenesten.

Boks 3.1 Oversigt over de lønelementer, der indgår i standardfortjenesten

Grund-, kvalifikations- og funktionsløn mv., inkl. bruttotrækordning¹³
+ Pension (pensionsbidrag, kapital-/ratepension, gruppelevspræmie, ATP mv.)
+ Personalegoder (fri bil, fri kost og logi)
+ Genebetalinger (aften-/nattillæg, nattevagt, weekendtillæg, svejsetillæg, snetillæg)
+ Ferie- og søgnehelldagsbetalinger
+ Særlig feriegodtgørelse
+ Uregelmæssige betalinger (bonus/engangsvederlag, gratiale, udbetaling af 6. ferieuge mv.)

Den standardberegnete timefortjeneste er et mål for, hvad en lønmodtager har fået i løn pr. time, lønmodtageren har aftalt at arbejde, uanset hvor mange dage, lønmodtageren er syg eller arbejder ud over det, som er aftalt. Opgørelsen af den standardberegnete timefortjeneste er dermed uafhængig af, hvor meget fravær den enkelte har.

Beregningen af den standardberegnete timefortjeneste er kompleks, fordi timetallet ikke er defineret entydigt, men afhænger af, hvordan størrelsen af den enkelte lønkomponent hænger sammen med de forskelligartede typer af timer. Grund-, kvalifikations- og funktionsløn mv. sættes således fx i forhold til betalte timer, bortset fra fraværs- og overtimer. Personalegoder og uregelmæssige betalinger divideres med betalte timer, ekskl. overtimer, mens genetillægget divideres med betalte timer. For nærmere beskrivelse af denne standardisering, se Danmarks Statistik (2011, s. 32-33).

3.2 Analysedesign

Rapportens formål er at forøge vores viden om den del af forskellen mellem kvinder og mænd, der hidtil har været 'uforklaret'. I dette afsnit har vi beskrevet den statistiske metode og den fremgangsmåde, vi har anvendt i analysearbejdet.

3.2.1 Blinder-Oaxaca-dekomponeringsanalyse

Vi har anvendt Blinder-Oaxaca-dekomponeringsanalyse som analysemetode, hvilket svarer til metoden anvendt i de tidligere SFI-/VIVE-undersøgelser af forskellen mellem kvinders og

¹² I de tidligere SFI-/VIVE-undersøgelser indgår også opgørelser baseret på fortjenesten pr. præsteret time, der er et tilnærmet mål for arbejdsgiverens lønomkostninger. Beregninger baseret på dette lønbegreb undervurderer imidlertid lønforskellen mellem kvinder og mænd, fordi refusioner til arbejdsgiveren i forbindelse med medarbejderes fravær under fx sygdom og barsel, der må formodes at være ulige fordelt på køn, ikke er modregnet i fortjenesten (se også Larsen & Larsen, 2018).

¹³ Bruttotrækningsordning er i lønstatistikken udtryk for, at en lønmodtager genforhandler en tidligere indgået aftale med arbejdsgiveren og i den nye aftale får en ændret vederlagssammensætning således, at der fremtidigt er en mindre andel af kontantløn og en større andel personalegoder.

mænds timeløn. Det skal bemærkes, at den bruttolønforskelle, som vi estimerer ved hjælp af dekomponeringsmetoden, afviger fra den bruttolønforskelle, som vi betegner som den 'sande' bruttolønforskelle, se Boks 3.2. Med tal for hele arbejdsmarkedet for 2016 som eksempel finder vi, at den 'sande' lønforskelle er 14,5 pct. mod 11,5 pct. baseret på dekomponeringsanalyse.

Boks 3.2 Bruttolønforskelle

Den 'sande' bruttolønforskelle er her¹⁴ beregnet som forskellen mellem mænds og kvinders gennemsnitlige timeløn divideret med gennemsnitstimelønnen for alle.

Bruttolønforskellen, der estimeres vha. dekomponeringsmetoden, afviger fra den 'sande' bruttolønforskelle, fordi:

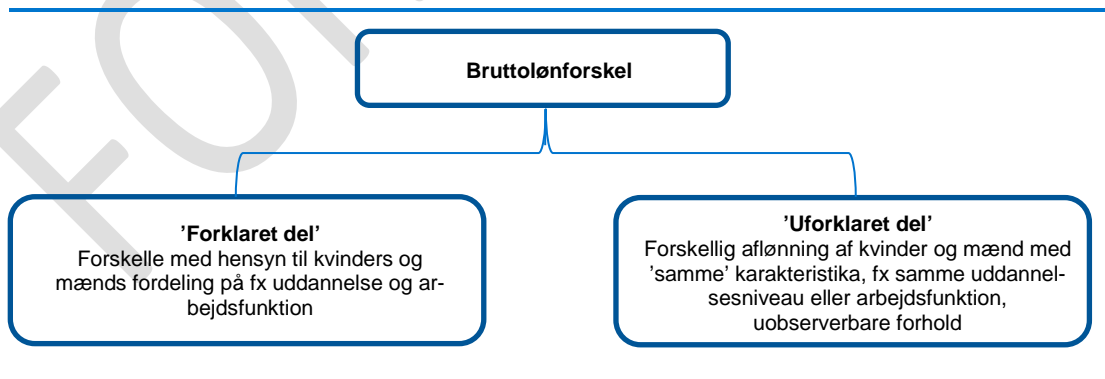
- Der er anvendt en logaritmetransformation i dekomponeringsanalysen, der desuden indebærer, at der er beregnet et geometrisk gennemsnit af mænds og kvinders løn.
- Forskellen mellem de to køns gennemsnitstimeløn er divideret med gennemsnitstimelønnen for mænd.

Med Blinder-Oaxaca-dekomponeringsanalysen er det muligt undersøge, i hvilket omfang lønforskellen hænger sammen med:

- at kvinder og mænd har forskellige målbare karakteristika ('forklaret del'),
- at kvinders og mænds lønafkast af disse karakteristika er forskellige og uobserverede forhold ('uforklaret del').

I praksis er analysen gennemført i to trin. Først foretages en regressionsanalyse for hhv. mænd og kvinder, hvoraf det ses, i hvor høj grad der er en sammenhæng mellem de enkelte forklarende variabler og de to køns timeløn. Dernæst er bruttolønforskellen mellem de to køn opløst (dekomponeret) i en 'forklaret' og en 'uforklaret' del. Opløsningen er illustreret i Figur 3.1. For nærmere beskrivelse af metoden, se bilag 1.

Figur 3.1 Dekomponeringsanalyse: Opløsning af bruttolønforskelle i 'forklaret' og 'uforklaret' del.



For fortolkning af resultater af dekomponeringsanalyse, se boks 3.3.

¹⁴ Der er forskellige metoder til at beregne bruttolønforskellen, se også afsnit 3.1 i Larsen & Larsen (2018).

'Forklaret' del: Kønsforskel i fordeling på karakteristika (dvs. forskellige X-værdier):

- Bidrag med positivt fortegn: Hvis kvinder havde haft den samme (mere gunstige) fordeling på det pågældende karakteristikum som mænd, ville lønforskellen have været mindre (reduktionen svarer til bidragets størrelse)
- Bidrag med negativt fortegn: Hvis kvinder havde haft den samme (mindre gunstige) fordeling på det pågældende karakteristikum som mænd, ville lønforskellen have været større (forøgelsen svarer til bidragets størrelse)

'Uforklaret' del: Kønsforskel i lønafkast (dvs. β -værdier) for kvinder og mænd med samme karakteristika (dvs. samme X-værdier):

- Bidrag med positivt fortegn: Hvis kvinder havde haft det samme (større) lønafkast af det pågældende karakteristikum som mænd, ville lønforskellen have været mindre (reduktionen svarer til bidragets størrelse)
- Bidrag med negativt fortegn: Hvis kvinder havde haft det samme (mindre) lønafkast af det pågældende karakteristikum som mænd, ville lønforskellen have været større (forøgelsen svarer til bidragets størrelse)

X- og β -værdier for kvinder og mænd på arbejdsmarkedet som helhed for 2016 fremgår af Bilag 2.

Med henblik på at få mere viden om lønforskellen mellem kvinder og mænd vil vi optimalt set gerne kunne 'forklare' hele lønforskellen, dvs. tilskrive hele lønforskellen forskelle i kvinders og mænds karakteristika. Imidlertid er dette givetvis ikke muligt, fordi vi ikke har data for en række karakteristika af potentiel betydning for lønforskellen, fx ubetalt overarbejde og præferencer, se også afsnit 2.10.

3.2.2 Fremgangsmåde

Udgangspunktet for de gennemførte analyser er de karakteristika, der er blevet taget højde for i de tidligere SFI-/VIVE-undersøgelser af lønforskellen mellem kvinder og mænd. Disse karakteristika og operationaliseringen heraf fremgår af Tabel 3.1. Vi har taget afsæt i, at alle disse karakteristika fra de tidligere undersøgelser skal bibeholdes (i en eller anden form) i denne undersøgelse.

Tabel 3.1 Karakteristika, der indgår i tidligere SFI-/VIVE-undersøgelser af lønforskellen mellem kvinder og mænd, og anvendt operationalisering.

Karakteristikum	Operationalisering ¹
Uddannelse	Dummyvariabler for højeste fuldførte uddannelsesniveaue (10 grupper).
Erhvervs erfaring	Kontinuer variabel for indbetalt ATP siden 1964.
Kønsopdelt arbejdsmarked	Dummyvariabler for sektor (4 grupper: privat, statslig, regional eller kommunal sektor), branche (36 grupper) og arbejdsfunktion (mere end 400 grupper baseret på DISCO-kodens fire første cifre).
Arbejdstid	Dummyvariabel for, om ens aftalte, ugentlige arbejdstid er mindre end 30 timer eller 30 timer eller derover.
Familieforhold	Dummyvariabel for, om man er enlig eller gift/samlevende; kontinuer variabel for antal børn i familien; dummyvariabel for, om yngste barn er 0-2 år, 3-6 år eller 7-17 år
Bopæl	Dummyvariabler for bopælsregion (5 grupper).

Note: ¹ Operationaliseringen af de enkelte karakteristika er nærmere beskrevet i Larsen & Larsen (2018), s. 28-32.

Vi har taget afsæt i data for "hele arbejdsmarkedet" (altså den del, der er dækket af lønstatistikken, jf. afsnit 3.1) for 2016, som er det seneste år, der belyses i Larsen & Larsen (2018). Ifølge Larsen & Larsen (2018) kan cirka halvdelen af lønforskellen mellem de to køn 'forklares', når der er taget højde for de karakteristika og den operationalisering heraf, der fremgår af Tabel 3.1, mens den resterende del er 'uforklaret', se også afsnit 1.2. Af Larsen & Larsen (2018) fremgår, hvad den 'forklarede' del af bruttolønforskellen består i, dvs. sammenhængen mellem den fundne lønforskel og kvinders og mænds forskellige fordeling på hvert af de undersøgte karakteristika. Den 'uforklarede' del fremgår derimod alene som én samlet værdi.¹⁵

Formålet med analyserne i denne undersøgelse har været at forbedre det sæt af kontrolvariabler, der hidtil har været anvendt. Sættet af kontrolvariabler anvendt i de tidligere undersøgelser kalder vi **fase 0**. Arbejdet med at finde frem til et ændret sæt af kontrolvariabler er gennemført i tre faser.

Med afsæt i de karakteristika, der fremgår af Tabel 3.1, har **fase 1** af analysearbejdet bestået i at anvende en Stata-procedure ('oaxaca'), der beregner brugbare værdier for hvert af disse karakteristika under den 'uforklarede' del, se Jann (2008).¹⁶ Dermed får vi mere viden om, hvad denne del af lønforskellen består i. De beregnede værdier angiver graden af sammenhæng mellem den fundne lønforskel og forskelle i kvinders og mænds afkast til hvert karakteristikum. Mere konkret viser en beregnet værdi, hvor stor lønforskellen ville have været, hvis lønafkastet af det pågældende karakteristikum havde været det samme for kvinder og mænd. Desuden får vi information om størrelsen af konstantleddet, dvs. hvor stor en del af lønforskellen, der kan tilskrives uobserverede forhold. For resultaterne af fase 1, se afsnit 4.1.

Det skal understreges, at de detaljerede resultater vedrørende den 'uforklarede' del skal fortolkes med varsomhed, da der kan være en del 'støj' forbundet med disse resultater, se fx Tharp et al. (2019). I vores valg af et ændret sæt af kontrolvariabler fokuserer vi på resultater under den 'forklarede' del, se kapitel 4. Tilsvarende fokuserer vi primært på den 'forklarede' del i forbindelse med vores formidling af analyseresultaterne i kapitel 5 og her vurderer vi løbende forud for en eventuel fortolkning, om de detaljerede resultater under den 'uforklarede' del forekommer meningsfulde.

Fase 2 af analysearbejdet har bestået i at forbedre opgørelsesmetoden for nogle af de karakteristika, der allerede tages højde for i analysen, se også kapitel 2. Vi har løbende taget stilling til, om inddragelse af en ny operationalisering har bidraget til at forøge vores viden om lønforskellen. Fokus har været på, om den nye operationalisering har forøget størrelsen af det pågældende karakteristikums (positive eller negative) bidrag til den 'forklarede' del af lønforskellen. Af Tabel 3.2 fremgår de karakteristika, som operationaliseringen er ændret for, ligesom den nye operationalisering er beskrevet. For resultaterne af fase 2, se afsnit 4.2.

¹⁵ I Larsen & Larsen (2018) blev Stata-procedure 'oaxaca2' anvendt. Denne procedure beregnede også værdier for den 'uforklarede' del for hver af de undersøgte karakteristika, men disse værdier er misvisende. Det skyldes, at de opgjorte værdier under den 'uforklarede' del afhænger af referencekategorien for hvert af de karakteristika, der indgår som kategorielle variable. Som eksempel kan nævnes, at hvis privat sektor er referencekategori for 'sektor', kommer vi frem til et andet resultat for 'sektor' end hvis kommunal sektor er referencekategori.

¹⁶ Når 'oaxaca'-proceduren anvendes, normaliseres de beregnede værdier, det vil sige, at værdierne udregnes som afvigelser fra gennemsnittet i det fulde datasæt. Normaliseringen sikrer, at værdierne for de kategorielle variabler er uafhængige af, hvad der er referencekategori (Jann, 2008).

Tabel 3.2 Karakteristika, der indgår i tidligere SFI-/VIVE-undersøgelser af lønforskellen mellem kvinder og mænd, og som indgår med ny operationalisering.

Karakteristikum	Ny operationalisering
Uddannelse	Kontinuert variabel for normeret tid for den højest fuldførte uddannelse. Dummyvariabler for uddannelsesretning (21 grupper ¹) ¹⁷ .
Erhvervs erfaring	Kontinuert variabel. Baseret på indbetalt ATP 1964-2007. Baseret på løntimer fra og med 2008.
Kønspodt arbejdsmarked	Kontinuert variabel for kvindeandel i sektor. Dummyvariabler for brancher (38 grupper ¹). Kontinuert variabel for kvindeandel i arbejdsfunktion på arbejdssted.
Arbejdstid	Dummyvariabler for, om ens aftalte, ugentlige arbejdstid er under 30 timer, 30-36 timer, 37 timer eller mindst 38 timer.

Note: ¹ Grupperingen af uddannelse og branche fremgår af Bilag 2.

Fase 3 i analysearbejdet har bestået i at teste, om inddragelsen af nye karakteristika, altså karakteristika, der ikke tidligere er blevet taget højde for, kan forøge vores viden om lønforskellen. Grundlaget for udvælgelsen af de nye karakteristika kan ses i kapitel 2.

Tabel 3.3 Nye karakteristika til forklaring af lønforskellen mellem kvinder og mænd, operationalisering heraf og markering af, om karakteristikum indgår i det ændrede sæt af kontrolvariabler.

Karakteristikum	Operationalisering	Indgår
Jobstatus	Dummyvariabler for, om man er a) elev eller lærling, b) leder eller mellemlider, c) medarbejder med særligt ansvar eller d) øvrig medarbejder.	X
Antal ansatte på arbejdssted	Dummyvariabler for, om der på ens arbejdssted er 0 ¹ , 1-9, 10-49, 50-249, 250-999 eller mindst 1.000 fuldtidsansatte. Arbejdsstedet er identificeret ved brug af arbejdsstedsnummer (ARBNR).	X
Jobskifte/anciennitet	Følgende har været testet: a. Kontinuert variabel for antal arbejdssteder i år t-1. b. Kontinuert variabel for antal arbejdssteder i år t-1 til t-5. c. Dummyvariabler for, om antallet af arbejdssteder i år t-1 er 0, 1, 2, 3 eller mindst 4. d. Dummyvariabler for, om antallet af arbejdssteder i år t-1 til t-5 er 0, 1-2, 3-4, 5-7 eller mindst 8. Arbejdsstedet er identificeret ved brug af arbejdsstedsnummer (ARBNR).	
Betalt overarbejde	Følgende har været testet: a. Dummyvariabler for, om antallet af timer, der er arbejdet ud over normen i den aktuelle ansættelse i året (år t) er 0, 1-24, 25-49, 50-99 eller mindst 100 timer. b. Kontinuert variabel for antallet af timer, der er arbejdet ud over normen i den aktuelle ansættelse i året divideret med antallet af betalte timer.	
Fravær ²	Kontinuert variabel for antal fraværstimer som andel af antal betalte timer i den aktuelle ansættelse i året (år t) Kontinuerte variabler for a) sygefravær og b) barselsfravær for hvert år i perioden t-1 til t-5. ³	X
Forudgående ledighed	Følgende har været testet: a. Kontinuert variabel for antal ugers bruttoledighed i år t-1 b. Kontinuert variabel for antal ugers bruttoledighed i år t-1 til t-5.	

Note: ¹ '0 fuldtidsansatte' afspejler bl.a., at militæret i nogle tilfælde ikke oplyser antal ansatte. ² De anvendte fraværsvariabler er nærmere beskrevet i bilag 3. ³ Vi har valgt at gå fem år tilbage i tid, fordi betydningen af syge- og barselsfravær ebber ud på dette tidspunkt.

¹⁷ Vi har undervejs overvejet at inddrage mere detaljerede uddannelseskategorier, der omfatter 143 uddannelsesretninger. Vi har valgt at fastholde 21-grupperingen, fordi inddragelse af de mere detaljerede kategorier ikke ændrer substantielt ved de konklusioner, der drages i kapitel 5. Resultater af at erstatte de anvendte oplysninger om uddannelsesretning med en mere detaljeret gruppering fremgår af fodnoter. Brug af 21-grupperingen giver os desuden mulighed for at præsentere en overskuelig tabel med X- og β -værdier i Bilag 2.

Af Tabel 3.3 fremgår de nye karakteristika, der er blevet testet og den anvendte operationalisering. Desuden er det markeret, hvis det pågældende karakteristikum indgår i det endelige sæt af kontrolvariabler. Et givet karakteristikum er medtaget, hvis inddragelse heraf bidrager (positivt eller negativt) til den 'forklarede' del af lønforskellen. Vi skæler desuden her til, om konstantleddet reduceres, men det tillægger vi mindre betydning. For resultaterne af fase 3, se afsnit 4.3.

For en opsummering af undersøgelsens faser, se boks 3.4.

Boks 3.4 Undersøgelsens faser

Fase 0: Sættet af kontrolvariabler svarer til dem, der er anvendt i tidligere SFI-/VIVE-undersøgelser. Ikke værdier for sættet af kontrolvariabler under den 'uforklarede' del.

Fase 1: Værdier for sættet af kontrolvariabler under den 'uforklarede' del er tilføjet til fase 0.

Fase 2: Opgørelsen af en række af de eksisterende kontrolvariabler er ændret i forhold til fase 1.

Fase 3: Nye kontrolvariabler er tilføjet i forhold til fase 2.

Det er ikke givet, at det ændrede sæt af kontrolvariabler, der er udvalgt med afsæt i data for arbejdsmarkedet som helhed for 2016, også bidrager til at forklare en større del af lønforskellen for personer, der er placeret forskellige steder i lønfordelingen. Det er heller ikke givet, at vi får mere viden om lønforskellen på sektorniveau i 2016 og for hele arbejdsmarkedet for andre år. For at undersøge dette, er det ændrede sæt af kontrolvariabler efterfølgende dels anvendt på data for personer, der er placeret forskellige steder i lønfordelingen¹⁸ i 2016, dels anvendt på data for sektorer for 2016 og på data for hele arbejdsmarkedet for 2010, 2013 og 2017.¹⁹ For resultaterne heraf, se afsnit 4.5-4.6.

I kapitel 5 redegør vi for den nye viden, vi opnår om forskellen mellem kvinders og mænds timeløn ved brug af det ændrede sæt af kontrolvariabler.

¹⁸ Det er desværre kun muligt delvist at teste betydningen at ændre kontrolvariabler i en analyse, hvor vi dekomponerer lønforskellen ved forskellige percentiler. For nærmere information om dette, se afsnit 4.5.

¹⁹ Vi går ikke længere tilbage end 2010, fordi variabelen jobstatus først findes for hele arbejdsmarkedet fra og med dette år. Man kunne argumentere for, at det er problematisk at bruge 2010 som udgangspunkt, fordi er lige efter finanskrisen. Imidlertid tyder resultater i Larsen & Larsen (2018) på, at lønforskellen stort set er faldet år for år siden 2007. Det seneste år er 2017. Da analyserne blev gennemført, var løndata for 2018 tilgængelige. Imidlertid var data for erhvervs erfaring ikke til rådighed for 2018, hvorfor det ikke var muligt for os at gennemføre analyserne for dette år.

4 Valg af ændret sæt af kontrolvariabler

I dette kapitel redegør vi først for, hvordan vi via fase 1, 2 og 3 er kommet frem til det ændrede sæt af kontrolvariabler. Resultaterne for de tre faser er baseret på tal for hele arbejdsmarkedet for 2016. Dernæst viser vi, hvor anvendeligt det ændrede sæt af kontrolvariabler er, når vi skelner mellem forskellige placeringer i lønfordelingen og når vi gennemfører analyser dels på sektorniveau, dels for andre år (2010, 2013 og 2017).

Det skal indledningsvis bemærkes, at de variabler, vi anvender til at opgøre betydningen af henholdsvis uddannelse og kønsopdelt arbejdsmarked ikke kan fortolkes uafhængigt af hinanden, fordi de korrelerer relativt meget. Dette er baggrunden for, at vi i tabellerne i dette kapitel har valgt også at vise det samlede bidrag for de karakteristika, der vedrører henholdsvis uddannelse og kønsopdelt arbejdsmarked.

Det skal samtidig bemærkes, at bidragene fra uddannelse og kønsopdelt arbejdsmarked heller ikke kan betragtes som værende uafhængige af hinanden. Det skyldes, at placeringen på arbejdsmarkedet i væsentligt omfang hænger sammen med valget af uddannelse. Bidragene er vist hver for sig – for dermed også at illustrere, at bidraget fra uddannelse af negativt, mens bidraget fra kønsopdelt arbejdsmarked er positivt – men bør fortolkes i sammenhæng.

Det skal endelig bemærkes, at vi i dette kapitel viser bidragene for hver af de forklarende variabler under både den 'forklarede' og den 'uforklarede' del, men i beskrivelsen af analyserne fokuserer vi alene på bidragene under den 'forklarede' del og konstantleddet, se også afsnit 3.2.

4.1 Fase 1: Kønsforskel i afkast af karakteristika

Formålet med fase 1 har været at få sat værdi på bidragene under den 'uforklarede' del fra dels hvert karakteristikum fra det oprindelige sæt af kontrolvariabler (fase 0), dels konstantleddet. Resultatet heraf fremgår af Tabel 4.1.

Resultaterne viser, at den beregnede bruttolønforskel er på 11,5 pct.²⁰, hvoraf de 5,5 pct. er 'uforklaret'. Det fremgår også, at konstantleddet (som udgør en del af den uforklarede del) er relativt stort, nemlig 7,1 pct. Dette kan afspejle, at en ikke-ubetydelig del af lønforskellen ifølge denne opgørelse kan tilskrives forhold, der er uobserverede i vores undersøgelse.

²⁰ Hvilket i korrigerede tal svarer til en bruttolønforskel på 14,5 pct., se også boks 3.2 og Larsen & Larsen (2018).

Tabel 4.1 Bruttolønforskel og bidrag fra karakteristika særskilt for 'forklaret' og 'uforklaret' del, 2016. Fase 1 og 2. Procent.

	Fase 1	Fase 2
Bruttolønforskel	11,5	11,5
<i>'Forklaret' del, i alt</i>	5,9	8,0
Uddannelse	-1,7	-3,9
Niveau	-1,7	
Længde		-2,5
Retning		-1,4
Erhvervs erfaring, oprindelig opgørelse	0,8	
Erfaringserfaring, ny opgørelse		0,9
Kønsopdelt arbejdsmarked	7,1	11,4
Sektor, dummyvariabler	0,6	
Branche, dummyvariabler	3,6	3,4
Arbejdsfunktion, dummyvariabler	2,9	
Sektor, kvindeandele		1,8
Jobcelle, kvindeandele		6,2
Arbejdstid, oprindelig opgørelse	0,2	
Arbejdstid, ny opgørelse		0,2
Børn	-0,5	-0,5
Enlig	0,0	0,0
Region	-0,1	-0,1
<i>'Uforklaret' del, i alt</i>	5,5	3,5
Uddannelse	0,5	14,6
Niveau	0,5	
Længde		14,1
Retning		0,6
Erhvervs erfaring, oprindelig opgørelse	-1,3	
Erfaringserfaring, ny opgørelse		-0,5
Kønsopdelt arbejdsmarked	-2,3	-8,4
Sektor, dummyvariabler	0,4	
Branche, dummyvariabler	-0,6 ^{1.S}	-1,7
Arbejdsfunktion, dummyvariabler	-2,1	
Sektor, kvindeandele		-6,9
Jobcelle, kvindeandele		0,2 ^{1.S}
Arbejdstid, oprindelig opgørelse	2,0	
Arbejdstid, ny opgørelse		2,2
Børn	-0,7	1,2
Enlig	0,2	0,6
Region	0,0	-0,1
Konstant	7,1	-6,1
Antal obs.	1.848.200	1.848.200

Anm.: ^{1.S} angiver, at bidrag fra karakteristikum ikke er statistisk signifikant på 5-pct.-niveau.

4.2 Fase 2: Forbedret opgørelse af allerede målte karakteristika

Fase 2 har bestået i at forbedre opgørelsen af nogle af de karakteristika, der allerede tages højde for i analysen. Vi har haft fokus på at forbedre opgørelsen af uddannelse, erhvervs erfaring, kønsopdelt arbejdsmarked og arbejdstid. Den oprindelige og den nye operationalisering af de fire karakteristika fremgår af Tabel 3.1 og Tabel 3.2.

Den nye operationalisering af uddannelse, erhvervs erfaring, kønsopdelt arbejdsmarked og arbejdstid har samlet set medvirket til, at vi kan forklare en større del af lønforskellen. Ændringerne har samlet set bidraget til at forøge den 'forklarede' del af lønforskellen med 2,1 procentpoint fra 5,9 pct. efter fase 1 til 8,0 pct. efter fase 2, se Tabel 4.1. Dette svarer til, at vi går fra at kunne forklare 52 pct. af lønforskellen med det oprindelige sæt af kontrolvariabler til at kunne forklare 70 pct. med det ændrede sæt af kontrolvariabler. Konstantleddet er fortsat numerisk set relativt stort, nemlig -6,1 pct. mod 7,1 pct. efter fase 1. I det følgende redegør vi for, hvilke af de ændringer, vi har foretaget i opgørelsen af hhv. uddannelse, erhvervs erfaring, kønsopdelt arbejdsmarked og arbejdstid, der primært har bidraget til at forøge den 'forklarede' del af lønforskellen fra fase 1 til fase 2.

Kønsopdelt arbejdsmarked

Ved at sammenligne resultaterne af fase 1 og 2 finder vi, at især ændringerne af opgørelsen af det kønsopdelte arbejdsmarked har gjort en forskel. Ændringen består i, at dummyvariabler for sektor er erstattet af kvindeandele inden for sektorer, mens dummyvariabler for arbejdsfunktioner er erstattet af kvindeandele inden for arbejdsfunktioner på hvert arbejdssted (her kaldet 'jobceller'), se også kapitel 2.

Udskiftningen af dummyvariabler med kvindeandele betyder, at bidraget fra sektor stiger fra 0,6 pct. i fase 1 til 1,8 pct. i fase 2, dvs. en stigning på 1,2 procentpoint. Forøgelsen, der følger af udskiftningen af dummyvariabler for arbejdsfunktioner med kvindeandele inden for jobceller, er endnu større, idet bidraget stiger med hele 3,3 procentpoint fra 2,9 pct. i fase 1 til 6,2 pct. i fase 2. Samtidig reduceres bidraget fra dummyvariabler for branche en smule fra 3,6 pct. i fase 1 til 3,4 pct. i fase 2, dvs. et fald på 0,2 procentpoint. Bidraget fra det kønsopdelte arbejdsmarked er dermed samlet set steget med 4,3 procentpoint fra 7,1 pct. i fase 1 til 11,4 pct. i fase 2.'

Ud over de nævnte ændringer har vi også testet, om kvindeandele inden for brancher (som erstatning for dummyvariabler for branche) og kvindeandele inden for arbejdsfunktioner (som supplement til kvindeandele inden for jobceller) forøger det samlede bidrag til den 'forklarede' del fra det kønsopdelte arbejdsmarked. Det viste sig ikke at være tilfældet.

Uddannelse

I de tidligere analyser blev der alene kontrolleret for niveauet for højeste fuldførte uddannelse. Kønsforskelle i uddannelsesvalg, fx mænds overrepræsentation inden for STEM-fagene, viser sig i væsentligt omfang ved, at de to køn er ansat i forskellige arbejdsfunktioner. Tidligere indgik mere end 400 dummyvariabler for arbejdsfunktion, der – ud over kønsforskellen i placeringen på arbejdsmarkedet – i høj grad også indfangede betydningen af kønsforskelle mht. uddannelsesvalg.

Når dummyvariabler for arbejdsfunktion ikke indgår, bidrager uddannelsesretning også til den 'forklarede' del af lønforskellen. Vi har derfor erstattet variabelen, der måler niveauet for højeste fuldførte uddannelse, med to nye uddannelsesvariabler, der måler henholdsvis uddannelses-

længde og -retning. Ændringen betyder, at det samlede bidrag fra uddannelse til den 'forklarede' del af lønforskellen stiger numerisk med 2,2 procentpoint fra -1,7 pct. i fase 1 til -3,9 pct. i fase 2. Bidraget ser ud til primært at stamme fra 'længde', som dog ikke kan betragtes som helt uafhængig af 'retning', jf. grupperingen af uddannelsesretning, der fremgår af Bilag 2.

Erhvervs erfaring og arbejdstid

Det fremgår, at den forbedrede opgørelse af erhvervs erfaring (jf. kapitel 2) kun bidrager marginalt mere til den 'forklarede' del af lønforskellen end den oprindelige variabel (ændring fra 0,8 pct. i fase 1 til 0,9 pct. i fase 2).

Den forbedrede opgørelse af ugentlig arbejdstid består i, at vi har erstattet en dummyvariabel for, om ens aftalte, ugentlige arbejdstid er mindre end 30 timer eller ej med en mere detaljeret kategorisering af arbejdstid. I denne kategorisering skelnes mellem, om den aftalte, ugentlige arbejdstid er under 30 timer, 30-36 timer, 37 timer eller mindst 38 timer.²¹ Den nye operationalisering ændrer ikke på bidraget fra arbejdstid til den 'forklarede' del, der er 0,2 pct. i både fase 1 og 2.

Selvom den ændrede operationalisering af hhv. erhvervs erfaring og arbejdstid ikke har forøget variablernes bidrag til den 'forklarede' del af lønforskellen nævneværdigt, har vi valgt at bibeholde erhvervs erfaring og arbejdstid med ny og – efter vores opfattelse – bedre operationalisering i det videre arbejde under fase 3.

Opsamling

Resultaterne i fase 2 viser alt i alt, at bidraget til forøgelsen af den 'forklarede' del af lønforskellen primært stammer fra ændringen i bidraget fra uddannelse og kønsopdelt arbejdsmarked. Det samlede bidrag herfra er således steget fra 5,4 pct. (-1,7 + 7,1 pct.) i fase 1 til 7,5 pct. (-3,9 + 11,4 pct.) i fase 2. Ændringen er er følge af, at vi:

- Erstatter oplysninger om uddannelsesniveau med oplysninger om uddannelsesretning og -længde
- Erstatter dummyvariabel for sektor og dummyvariabel for arbejdsfunktion med kvindeandel i sektor og kvindeandel i jobcelle, dvs. arbejdsfunktion på arbejdssted.

Det er i øvrigt værd at bemærke, at bidragene til den 'forklarede' del af lønforskellen fra børn, enlig og region er uændret fra fase 1 til fase 2.

4.3 Fase 3: Tilføjelse af nye karakteristika

Fase 3 har bestået i at teste, om inddragelse af en række nye karakteristika kan bidrage til at forøge vores viden om lønforskellen. Det er tilfældet, hvis inddragelse af det nye karakteristikum bidrager (positivt eller negativt) til den 'forklarede' del af lønforskellen. Vi skeler dog også til, om konstantleddet reduceres. De testede karakteristika kan ses i kapitel 2 og Tabel 3.3.

De karakteristika, der har vist sig at bidrage til den 'forklarede' del af lønforskellen, er:

- Jobstatus
- Antal ansatte på arbejdssted

²¹ Vi har også testet mere detaljerede kategoriseringer af arbejdstid, men det ændrede ikke resultatet.

- Fravær

Resultaterne fremgår af Tabel 4.2. Af kolonne (1) fremgår fase 2-resultaterne. Af kolonne (2)-(4) fremgår bidraget fra de nye karakteristika, når de tilføjes hver for sig efter fase 2. Af kolonne (5) fremgår resultaterne i fase 3, hvor alle nye karakteristika er tilføjet på samme tid.

I det følgende redegør vi først for resultaterne af at tilføje hvert nyt karakteristikum til fase 2. Dernæst redegør vi for de samlede resultater i fase 3. Gennemgangen illustrerer, at der er en del korrelation mellem de inkluderede variabler.

Jobstatus

Vi har indsat dummyvariabler for jobstatus, der angiver, om man er elev/lærling, leder/melleleder, medarbejder med særligt ansvar eller øvrig medarbejder. Inddragelse af denne variabel bidrager til, at vi i højere grad end tidligere kan opgøre betydningen af den vertikale kønsopdeling på arbejdsmarkedet, se også kapitel 2. Som udgangspunkt forventer vi, at bidraget fra jobstatus til den 'forklarede' del er positivt, fordi mænd er placeret højere i jobhierarkiet end kvinder, se bilagstabel 2.1 i Bilag 2. Resultaterne af at inddrage jobstatus fremgår af Tabel 4.2, kolonne (2).

Bidraget til den 'forklarede' del fra inddragelse af jobstatus er som forventet positivt og relativt stort, nemlig 2,2 pct., se kolonne (2). Tilføjelsen af jobstatus reducerer imidlertid bidraget til den 'forklarede' del for en række andre variabler i modellen, hvilket sandsynligvis afspejler, at disse variabler hidtil har "fanger" elementer af den vertikale kønsopdeling på arbejdsmarkedet (dvs. de er højt korrelerede med jobstatus).

Ser vi på de øvrige variabler, der opgør betydningen af det kønsopdelte arbejdsmarked, er det især bidraget fra kvindeandele inden for jobceller, men også fra kvindeandele inden for sektorer, der reduceres, mens bidraget fra dummyvariabler for branche er forøget en smule. Selvom bidraget fra jobstatus er relativt stort, reduceres det samlede bidrag fra kønsopdelt arbejdsmarked fra 11,4 til 10,1 pct., når jobstatus indgår. Det negative bidrag fra uddannelse reduceres også. Især bidraget fra uddannelsesretning, men også bidraget fra uddannelseslængde bliver således numerisk set mindre. Bidraget fra uddannelse til den 'forklarede' del reduceres samlet set fra -3,9 til -2,7 pct.

Inddragelse af jobstatus betyder isoleret set, at vi forklarer en *mindre* del af lønforskellen. Størrelsen af den 'forklarede' del reduceres fra 8,0 pct. efter fase 2 til 7,7 pct., se kolonne (1) og (2). Konstantleddet reduceres numerisk fra -6,1 til -5,4 pct. Vi har valgt at bibeholde jobstatus i modellen, fordi a) vi vurderer, at det er vigtigt at kontrollere for betydningen af den vertikale kønsopdeling på arbejdsmarkedet, b) bidraget fra variabelen er relativt stort og c) konstantleddet dermed reduceres.

Antal ansatte på arbejdssted

Vi har inkluderet dummyvariabler for antallet af ansatte på arbejdsstedet. Vi tester dermed, om størrelsen af arbejdsstedet har betydning for lønforskellen mellem kvinder og mænd. Som udgangspunkt forventer vi, at bidraget fra antal ansatte til den 'forklarede' del er negativt, fordi mænd i større omfang er ansat på små arbejdssteder end kvinder, se bilagstabel 2.1 i Bilag 2, og fordi vi har en hypotese om, at lønforskellen er mindre på større arbejdssteder, se også kapitel 2. Resultatet af at inddrage oplysninger om antal ansatte på arbejdsstedet fremgår af Tabel 4.2, kolonne (3).

Tabel 4.2 Bruttolønforskel og bidrag fra karakteristika særskilt for 'forklaret' og 'uforklaret' del, 2016. Fra fase 2 til fase 3. Procent.

	(1)	(2)	(3)	(4)	(5)
	Fase 2	Jobstatus	Antal ansatte	Fravær	Fase 3
Bruttolønforskel	11,5	11,5	11,5	11,5	11,5
<i>'Forklaret' del, i alt</i>	8,0	7,7	8,7	9,6	9,7
Uddannelse	-3,9	-2,7	-4,3	-3,9	-3,2
Længde	-2,5	-2,2	-2,4	-2,5	-2,0
Retning	-1,4	-0,6	-1,9	-1,4	-1,2
Erhvervs erfaring	0,9	0,7	0,9	0,9	0,7
Kønsopdelt arbejdsmarked	11,4	10,1	13,0	11,0	11,8
Sektor, kvindeandele	1,8	0,9	2,4	1,4	1,5
Branche, dummyvariabler	3,4	3,6	3,2	3,5	3,2
Jobcelle, kvindeandele	6,2	3,4	7,4	6,2	4,9
Jobstatus		2,2			2,2
Antal ansatte på arbejdssted			-0,5		-0,6
Arbejdstid	0,2	0,1	0,2	0,1	0,0 ^{1.S.}
Fravær				2,0	1,4
Fraværstimer				0,7	0,5
Sygefravær, 2011-2015				0,4	0,4
Barselsfravær, 2011-2015				0,9	0,5
Børn	-0,5	-0,4	-0,5	-0,5	-0,4
Enlig	0,0	0,0	0,0	0,0	0,0
Region	-0,1	-0,1	-0,1	-0,1	-0,1
<i>'Uforklaret' del, i alt</i>	3,5	3,7	2,7	1,8	1,8
Uddannelse	14,6	15,3	11,4	13,7	10,8
Længde	14,1	15,4	10,3	13,2	10,2
Retning	0,6	0,0 ^{1.S.}	1,1	0,6	0,6
Erhvervs erfaring	-0,5	-0,3	-0,6	-0,3	-0,1 ^{1.S.}
Kønsopdelt arbejdsmarked	-8,4	-8,4	-12,2	-7,8	-12,8
Sektor, kvindeandele	-6,9	-5,6	-8,0	-6,2	-6,8
Branche, dummyvariabler	-1,7	-2,2	-1,5	-1,9	-1,9
Jobcelle, kvindeandele	0,2 ^{1.S.}	4,0	-2,6	0,3	0,6
Jobstatus		-4,7			-4,7
Antal ansatte på arbejdssted			0,4		1,1
Arbejdstid	2,2	1,2	2,2	2,2	1,0
Fravær				-2,3	-1,5
Fraværstimer				-0,8	-0,5
Sygefravær, 2011-2015				-0,5	-0,5
Barselsfravær, 2011-2015				-0,9	-0,5
Børn	1,2	0,8	1,2	1,2	0,8
Enlig	0,6	0,5	0,6	0,5	0,5
Region	-0,1	0,0	-0,2	-0,1	-0,1
Konstant	-6,1	-5,4	0,0 ^{1.S.}	-5,4	2,1
Antal observationer	1.848.200	1.848.200	1.848.200	1.848.200	1.848.200

Anm.: ^{1.S.} angiver, at bidrag fra karakteristikum ikke er statistisk signifikant på 5-pct.-niveau.

Bidraget fra antal ansatte til den 'forklarede' del er som forventet negativt, men relativt lille, nemlig -0,5 pct. Til gengæld har inddragelse af oplysninger om antal ansatte relativ stor betydning for bidragene fra uddannelse og kønsopdelt arbejdsmarked: Det negative bidrag fra uddannelse forøges fra -3,9 til -4,3 pct., hvilket kan tilskrives en forøgelse af det negative bidrag fra uddannelsesretning. Bidraget fra kønsopdelt arbejdsmarked forøges fra 11,4 til 13,0 pct. som resultat af en forøgelse af bidragene fra kvindeandele inden for især jobceller, men også sektorer. Betydningen af disse variabler forstærkes med andre ord, når vi også tager højde for, hvor mange ansatte, der er på den enkeltes arbejdssted.

Inddragelsen af oplysninger om antal ansatte på arbejdsstedet betyder samlet set, at den 'forklarede' del forøges fra 8,0 til 8,7 pct. Desuden skal bemærkes, at konstantleddet reduceres numerisk fra -6,1 til 0,0 pct., når antal ansatte indgår i modellen.

Fravær

Vi har testet betydningen af forskellige mål for fravær. Vi har opgjort fravær med afsæt i tre variabler, der alle viste sig at bidrage til den 'forklarede' del af lønforskellen: 1) antal fraværstimer som andel af antal betalte timer i løbet af året i den aktuelle ansættelse, 2) antal dages sygefravær i hvert af de fem forudgående år og 3) antal dages barselsfravær i hvert af de fem forudgående år. Derudover har vi testet betydningen af forudgående ledighed, der viste sig ikke at bidrage til den 'forklarede' del.

Fravær påvirker lønnen negativt (se kapitel 2) og kvinder har generelt mere fravær end mænd, se bilagstabel 2.1 i Bilag 2. Vi forventer derfor, at bidraget fra fravær til den 'forklarede' del af lønforskellen er positivt. Resultaterne af at inddrage de tre fraværsvARIABLER fremgår af Tabel 4.2, kolonne (4).

Bidraget fra fraværsvARIABLENE til den 'forklarede' del er som forventet positivt og samlet set på 2,0 pct. Heraf er bidraget fra barselsfravær størst, nemlig på 0,9 pct., mens bidraget fra sygefravær er mindst, nemlig 0,5 pct. Disse forskelle afspejler, at der især er forskelle på omfanget af kvinders og mænds fravær, når det gælder barsel, se bilagstabel 2.1 i Bilag 2. Bidragene fra de øvrige variabler til den 'forklarede' del er ikke nævneværdigt påvirket af, at der tages højde for oplysninger om fravær, heller ikke bidraget fra børn. Den eneste ændring, der er værd at bemærke, er, at bidraget fra kønsopdelt arbejdsmarked reduceres lidt fra 11,4 til 11,0 pct., hvilket kan tilskrives en reduktion i bidraget fra kvindeandele inden for sektorer.

Inddragelse af oplysninger om fravær forøger det samlede bidrag fra den 'forklarede' del relativt meget, nemlig fra 8,0 til 9,6 pct. Samtidig reduceres konstantleddet numerisk fra -6,1 til -5,4 pct.

Fase 3

Inddragelse af de tre nye karakteristika – jobstatus, antal ansatte på arbejdsstedet og fravær – betyder samlet set, at bidraget fra den 'forklarede' del af lønforskellen er forøget fra 8,0 pct. i fase 2 til 9,7 pct. i fase 3, dvs. en stigning på 1,7 procentpoint. Medtager vi også ændringen fra fase 1 til fase 2 på 2,1 procentpoint, er den samlede stigning i bidraget fra den 'forklarede' del på 3,8 procentpoint. Dette svarer til, at vi med det ændrede sæt af kontrolvariabler kan forklare 85 pct. af forskellen mellem kvinders og mænds timeløn mod 52 pct. ved anvendelse af det oprindelige sæt af kontrolvariabler, se Figur 4.2 nedenfor.

Konstantleddet er numerisk set reduceret fra -6,1 pct. til 2,1 pct. fra fase 2 til fase 3, hvilket kan tale for, at de tilbageværende uobserverede forhold af betydning for lønforskellen spiller en relativt begrænset rolle. Når vi ser bort fra konstantleddet, er bidraget fra den 'uforklarede' del,

dvs. kønsforskellen i lønafkastet af de undersøgte karakteristika, samlet set tæt på 0 i fase 3, nemlig -0,3 pct.

Det ændrede sæt af kontrolvariabler, der følger efter de ændringer, vi har foretaget i fase 2 og 3, fremgår af Tabel 4.3.

Tabel 4.3 Ændret sæt af kontrolvariabler og anvendt operationalisering.

Karakteristikum	Operationalisering
Uddannelse	Kontinuert variabel for normeret tid for den højest fuldførte uddannelse. Dummyvariabler for uddannelsesretning (21 grupper).
Erhvervs erfaring	Kontinuert variabel. Baseret på indbetalt ATP 1964-2007. Baseret på løntimer fra og med 2008.
Køn opdelt arbejdsmarked	Kontinuert variabel for kvindeandel i sektor. Dummyvariabler for brancher (38 grupper ¹). Kontinuert variabel for kvindeandel i jobcelle (arbejdsfunktion på arbejdssted). Dummyvariabler for jobstatus, dvs. om man er a) elev eller lærling, b) leder eller mellemleder, c) medarbejder med særligt ansvar eller d) øvrig medarbejder.
Antal ansatte	Dummyvariabler for, om der på ens arbejdssted er 0, 1-9, 10-49, 50-249, 250-999 eller mindst 1.000 fuldtidsansatte. Arbejdsstedet er identificeret ved brug af arbejdsstedsnummer (ARBNR).
Arbejdstid	Dummyvariabler for, om ens aftalte, ugentlige arbejdstid er under 30 timer, 30-36 timer, 37 timer eller mindst 38 timer.
Fravær	Kontinuert variabel for antal fraværstimer som andel af antal betalte timer i den aktuelle ansættelse i året (år t). Kontinuerede variabler for a) sygefravær og b) barselsfravær for hvert år i perioden t-1 til t-5.
Familieforhold	Dummyvariabel for, om man er enlig eller gift/samlevende; kontinuer variabel for antal børn i familien; dummyvariabel for, om yngste barn er 0-2 år, 3-6 år eller 7-17 år
Bopæl	Dummyvariabler for bopælsregion (5 grupper).

4.4 Lønfordelingen 2016

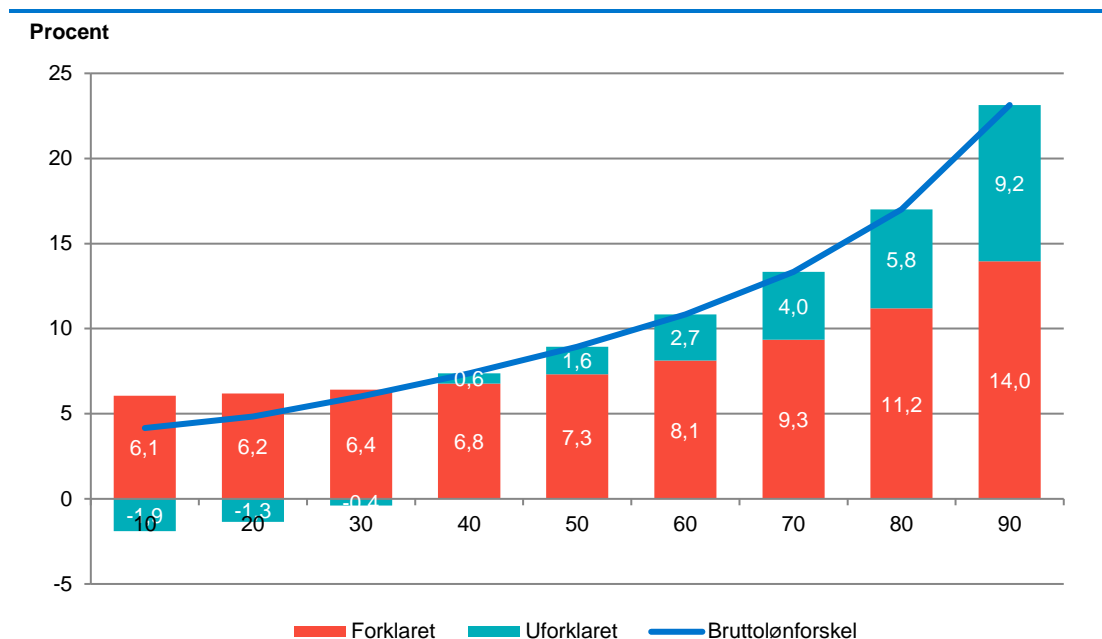
Et tidligere VIVE-studie viser, at man med (en tilpasser version af) det oprindelige sæt af kontrolvariabler kan forklare 47 pct. af lønforskellen blandt personer i bunden af lønfordelingen, 17 pct. i midten af lønfordelingen og 36 pct. i toppen af denne fordeling i 2016, se afsnit 1.2.

Vi ser i dette afsnit på, om ændringen i sættet af kontrolvariabler bidrager til, at vi kan forklare en større del af lønforskellen ved forskellige percentiler i lønfordelingen. Resultaterne er fremkommet ved brug af en såkaldt kvantil-dekomponeringsanalyse, hvor der er foretaget beregninger for hver 10. percentil i intervallet 10-90. For nærmere beskrivelse heraf, se fx Larsen & Larsen (2018). Metoden gør det desværre ikke muligt at medtage så mange og så detaljerede kontrolvariabler som indgår i fase 3. Vi har således måtte erstatte inddelingen i 36 brancher med en inddeling i 10 brancher. Vi har desuden måttet udelade dummyvariabler for enlig, region og kontinuerede variabler for hhv. syge- og barselsfravær for fem år siden (det vil her sige i 2011). De nævnte variabler er frasorteret, fordi bidragene herfra er meget små i Blinder-Oaxaca-dekomponeringsanalysen.

Vi finder, at vi med (den tilpassede version af) det ændrede sæt af kontrolvariabler kan forklare en større andel af lønforskellen i hele lønfordelingen, Figur 4.1. I bunden af lønfordelingen (ved 10-, 20- og 30-percentilen) er lønforskellen tilmed 'overforklaret'. Dette afspejler, at bidraget fra kønsforskelle i de inkluderede karakteristika i sig selv ville have ført til en større lønforskel i

denne del af lønfordelingen, men at dette imødegås af, at lønfokastet for de inkluderede karakteristika er større for kvinder end for mænd og/eller af kønsforskelle mht. uobserverede forhold til kvinders fordel.

Figur 4.1 Bruttolønforskelle mellem kvinder og mænd på arbejdsmarkedet som helhed, særskilt for percentiler¹, i alt og fordelt på 'forklaret' og 'uforklaret' del, 2016. Opgjort for standardberegnet timefortjeneste. Procent.



Anm.: Resultatet er baseret på kvantil regression. For beskrivelse heraf, se fx Larsen & Larsen (2018).

Note: ¹ Opdelingen af timelønnen i percentiler er foretaget for kvinder og mænd hver for sig. Det vil sige, at beregningen for fx 10. percentil tager afsæt i gennemsnitslønnen ved hhv. kvinders og mænds 10. percentil.

I midten af lønfordelingen (ved 50-percentilen) kan vi med (den tilpassede version af) det ændrede sæt af kontrolvariabler forklare 82 pct. af lønforskellen, mens vi kan forklare 60 pct. af lønforskellen i toppen af lønfordelingen (ved 90-percentilen). Den andel af lønforskellen, der kan forklares med det anvendte sæt af kontrolvariabler, falder med andre ord med stigende løn. Det er dog en relativ stor andel af også lønforskellen blandt de højestlønnede, der kan forklares.

Det anvendte metode gør det desværre ikke muligt at sige noget om de størrelser af bidragene ved hvert percentil for hver af de variabler, der indgår i analysen. Det er med andre ord ikke muligt at belyse, om hvor meget fx det kønsopdelte arbejdsmarked 'forklarer' af lønforskellen i forskellige dele af lønfordelingen.

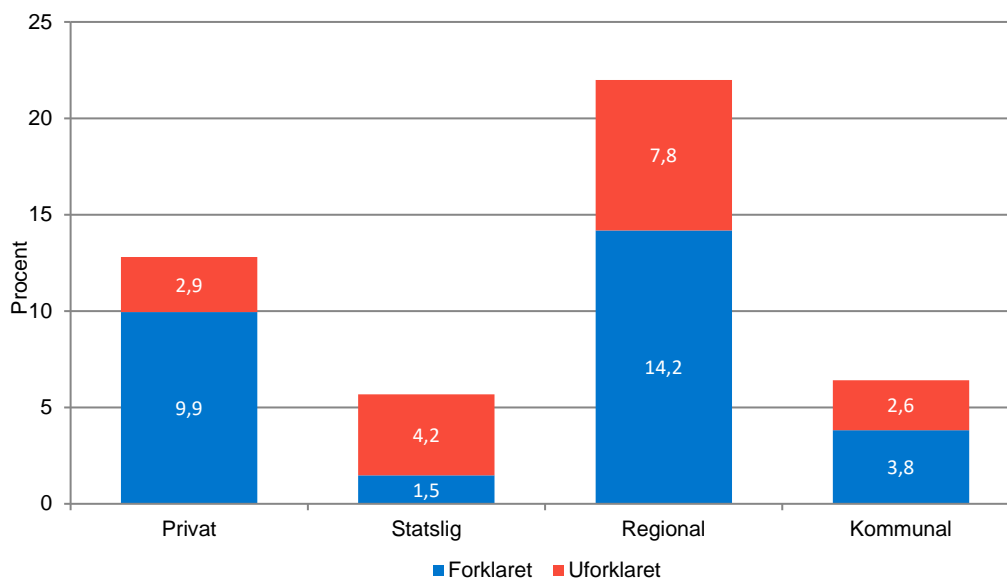
4.5 Sektorniveau

I dette afsnit ser vi, om og i hvilket omfang det ændrede sæt af kontrolvariabler for hele arbejdsmarkedet bidrager til, at vi kan forklare en større del af lønforskellen inden for de enkelte sektorer.

Vi finder, at vi kan forklare væsentligt mere af lønforskellen i den private sektor end hidtil. Med det oprindelige sæt af kontrolvariabler kunne vi kun forklare de 3,2 pct. af en lønforskel på 12,8 pct. (svarende til 19 pct. af lønforskellen) – den private sektor var dermed den af sektorerne,

hvor den 'uforklarede' del var størst, se også afsnit 1.2. Med det ændrede sæt af kontrolvariabler kan vi forklare hele 9,9 pct. af de 12,8 pct. (svarende til 78 pct.), se Figur 4.2. Den private sektor går dermed fra at være den af sektorerne, hvor vi med det oprindelige sæt af kontrolvariabler kunne forklare den mindste del af lønforskellen, til at være den sektor, hvor vi med de ændrede kontrolvariabler kan forklare den største andel.

Figur 4.2 Bruttolønforskel mellem kvinder og mænd særskilt for sektorer, opdelt på 'forklaret' og 'uforklaret' del, ændret sæt af kontrolvariabler, korrigerede tal¹, 2016. Procent.



Note: ¹ For korrigerede tal, se boks 3.2 og Larsen & Larsen (2018).

I den regionale sektor kan vi med det ændrede sæt af kontrolvariabler forklare 14,2 pct. af en lønforskel på 22 pct. (svarende til 64 pct. af lønforskellen). Dette svarer stort set til den andel, vi kunne forklare med de oprindelige kontrolvariabler. Det samme er tilfældet i den statslige sektor, hvor vi kan forklare 1,5 pct. af en lønforskel på 5,7 pct. (svarende til 25 pct. af lønforskellen). Der er dermed også med det ændrede sæt af kontrolvariabler en relativt stor andel af lønforskellen i staten, der ikke er 'forklaret'. En medvirkende årsag er, at vi ikke inddrager meget detaljerede oplysninger om uddannelsesretning. Medtager vi mere detaljerede uddannelsesoplysninger kan vi 'forklare' 39 pct. af lønforskellen i staten.²² Det er også værd at bemærke, at et relativt stort negativt bidrag fra antal ansatte bidrager til at reducere den 'forklarede' andel i staten (se Bilag 4). Endelig spiller det også en rolle, at staten er den sektor, hvor lønforskellen er mindst. Selvom vi kun 'forklarer' 25 pct. af lønforskellen i staten er den 'uforklarede' lønforskel af relativt begrænset størrelsesorden, nemlig 4,2 pct.

I den kommunale sektor kan vi forklare lidt en lidt *mindre* del af lønforskellen med det ændrede sæt af kontrolvariabler, nemlig 3,8 pct. af en lønforskel på 6,4 pct. (svarende til 60 pct. af lønforskellen) mod 4,4 pct. (svarende til 69 pct. af lønforskellen) med de oprindelige kontrolvariabler. Vi har set nærmere på baggrunden herfor. For resultater for fase 3, se bilagstabel 4.1 i bilag 4 (resultater for fase 1 er ikke vist). Vi finder, at den væsentligste årsag er, at vi kontrollerer

²² For de øvrige sektorer gælder, at vi efter inddragelse af mere detaljerede uddannelsesoplysninger også kan forklare en større del af lønforskellen i den regionale sektor (72 pct.), mens vi kan forklare en *mindre* andel af lønforskellen i privat og kommunal sektor (hvv. 70 og 47 pct.).

for fravær, hvorfra bidraget er *negativt* i denne sektor – et negativt bidrag, der strammer fra barselsfravær.²³ Det negative bidrag er uventet i lyset af, at kvinder tager væsentligt mere barselsfravær end mænd og at barselsfravær i følge den eksisterende litteratur er forbundet med 'lønstraf', se afsnit 2. Vores opgørelse tegner dermed et billede af, at der ikke er en negativ løneffekt forbundet med barselsfravær i den kommunale sektor – tværtimod. En tendens, der ifølge resultaterne er endnu mere udtalt i den regionale sektor.

En nærmere granskning af resultaterne for den regionale og den kommunale sektor viser, at "løngevinsten" ved barselsfravær i disse sektorer er større for mænd end kvinder. En umiddelbar fortolkning heraf er, at mænd belønnes mere end kvinder for at tage barsel. En forlængelse heraf kan en hypotese dog være, at den tilsyneladende løngevinst (for især mænd) ved at tage barsel i regioner og kommuner kan tilskrives, at især ressourcestærke, yngre (mænd) får lønstigninger og at det samtidig også (især) er dem, der får børn og tager barsel. For at tage højde herfor har vi forsøgsvis erstattet den anvendte oplysning om uddannelsesretning med oplysninger om mere detaljeret uddannelse og tilføjet en kontrolvariabel for alder. Dette reducerer det positive bidrag fra barselsfravær i både regioner og kommuner fra hhv. -1,9 og -1,3 pct. til ca. 1,0 pct., men det forsvinder ikke.

Resultaterne vedrørende barselsfravær i regionerne og kommuner tyder alt i alt på, at "løngevinsten" (β -værdien) ved barselsfravær her er mindre for kvinder end for mænd. Der er samlet set ikke tegn på, at der er en nævneværdig "løngevinst" (men heller ikke en "lønstraf") ved at have barselsfravær for kvinder i de offentlige sektorer.

Den positive "løngevinst" ved barselsfravær for mænd i regioner og kommuner er overraskende (se afsnit 2.5) og kan potentielt afspejle noget andet end en løngevinst ved at have haft barselsfravær *i sig selv*. Vi tager højde for fravær fem år tilbage i tid, og man kan potentielt forestille sig, at især mænd – i årene efter, at de har fået børn (og har haft fravær i den forbindelse), giver den en ekstra skalle og derved opnår en højere løn.

Resultaterne viser samlet set, at vi med det ændrede sæt af kontrolvariabler kan forklare en væsentlig større andel af lønforskellen i den private sektor, men ikke i de offentlige sektorer. En medvirkende årsag til sektorforskellen er – ud over de nævnte resultater vedrørende betydningen af at inddrage oplysninger om barselsfravær – at overenskomstforhandlingerne i de offentlige sektorer er langt mindre decentrale end det er tilfældet i den private sektor. Inddragelsen af arbejdsstedsoplysninger forøger derfor ikke i samme omfang den 'forklarede' del i de offentlige sektorer som det er tilfældet i den private sektor.

Den lavere grad af forklaring i sektorerne (25-78 pct.) sammenlignet med arbejdsmarkedet som helhed (85 pct.) kan tilskrives, at betydningen af lønforskel på tværs af sektorer ikke indgår i de sektorspecifikke opgørelser.

4.6 Andre år

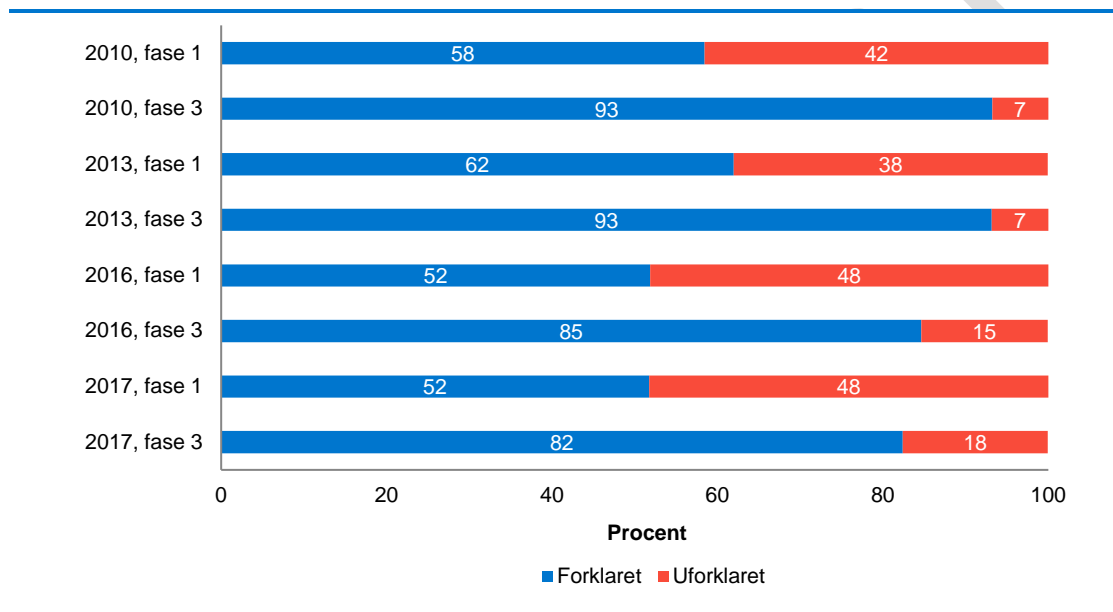
Det ændrede sæt af kontrolvariabler er udvalgt med afsæt i data for hele arbejdsmarkedet for 2016. Vi fandt som nævnt, at vi kan forklare 85 pct. af lønforskellen i 2016 med det ændrede sæt af kontrolvariabler mod 52 pct., når vi anvender de oprindelige kontrolvariabler.

²³ De inkluderede variabler for uddannelse og kønsopdelt arbejdsmarked forklarer samlet set en mindre del af lønforskellen i den kommunale sektor i fase 3 end i fase 1. Dette mere end opvejes imidlertid af, at bidraget fra arbejdstid forøges, når vi i fase 3 inkluderer mere detaljerede oplysninger om dette.

Vi ser her på, hvor anvendeligt det nye sæt af kontrolvariabler er, når vi tager afsæt i data for hele arbejdsmarkedet for andre år. Vi har til dette formål valgt at bruge data for hhv. 2010, 2013 og 2017.

Andelen af lønforskellen, der kunne forklares med det oprindelige sæt af kontrolvariabler, lå på ca. 60 pct. i 2010 og 2013 og på 52 pct. i 2017, se Figur 4.3. Den 'forklarede' andel stiger til 93 pct. i både 2010 og 2013, når det ændrede sæt af kontrolvariabler anvendes, mens den stiger til 82 pct. i 2017. Resultaterne tyder dermed på, at det ændrede sæt af kontrolvariabler ikke alene forbedrer modellen for 2016, men også med fordel kan anvendes på data for andre år.

Figur 4.3 Bruttolønforskel mellem kvinder og mænd fordelt på 'forklaret' og 'uforklaret' del, oprindeligt og ændret sæt af kontrolvariabler (fase 1 og 3), hele arbejdsmarkedet, 2010, 2013, 2016 og 2017. Procent.



5 Ny viden om lønforskellen mellem kvinder og mænd

I dette kapitel redegør vi for den nye viden, vi har opnået om forskellen mellem kvinders og mænds timeløn ved brug af det ændrede sæt af kontrolvariabler, se tabel Tabel 4.3. Vi redegør først for den nye viden, vi har fået om lønforskellen på arbejdsmarkedet som helhed i 2016. Dernæst beskriver vi den nye viden, vi har fået om lønforskellen i den private sektor i 2016. Det ændrede sæt af kontrolvariabler gør det som nævnt muligt at forklare væsentligt mere af lønforskellen i netop denne sektor. Endelig skitserer vi den nye viden, vi har fået om baggrunden for faldet i lønforskellen i de senere år. Fokus er i den forbindelse på udviklingen i perioden 2010-2017.

5.1 Hele arbejdsmarkedet 2016

I dette afsnit beskriver vi, hvilken ny viden det ændrede sæt af kontrolvariabler giver os, når vi analyserer hele arbejdsmarkedet under ét. Vi har i denne forbindelse valgt primært at fokusere på de samlede bidrag fra hvert karakteristikum, fx 'kønsopdelt arbejdsmarked' (som omfatter kvindeandel i sektor, branche, kvindeandel i jobcelle og jobstatus, se Tabel 4.3) frem for særskilte bidrag fra sektor, branche etc.²⁴ Vi skeler dog hertil i fortolkningen af resultaterne, der fremgår af Tabel 4.2.

Resultaterne for den 'forklarede' del baseret på hhv. det ændrede sæt af kontrolvariabler (fase 3) og det sæt af kontrolvariabler, der blev anvendt i Larsen & Larsen (2018) (fase 0), fremgår af Figur 5.1, mens resultaterne for den 'uforklarede' del fremgår af Figur 5.2.

'Forklaret' del

Lønforskellen mellem kvinder og mænd på arbejdsmarkedet som helhed er – ifølge de gennemførte dekomponeringsanalyser – på 11,5 pct. i 2016.²⁵ Vi kan som nævnt i afsnit 4.3 'forklare' 85 pct. af denne lønforskel i fase 3 mod 52 pct. i fase 0.

Lønforskellen ser ud til at hænge endnu mere sammen med, at de to køn foretager forskellige uddannelsesvalg og er forskelligt placeret på arbejdsmarkedet end hidtil antaget. Det samlede bidrag fra uddannelse og kønsopdelt arbejdsmarked er på 8,6 pct. ifølge fase 3 mod 5,4 pct. ifølge fase 0, se Figur 5.1.

Bidraget fra uddannelse er samlet set på -3,2 pct. Det negative bidrag kan bl.a. tilskrives, at kvinder i gennemsnit har længere uddannelse end mænd. Det vil sige, at hvis kvinder i gennemsnit havde haft samme (kortere) uddannelseslængde som mænd, ville lønforskellen have været større.

Bidraget fra det kønsopdelte arbejdsmarked er samlet set på 11,8 pct. Kønsopdelingen genererer ikke nødvendigvis i sig selv lønforskelle. Der opstår lønforskelle som følge af denne opdeling, når der er systematiske forskelle mellem lønnen i de kvinde- og de mandedominerede job (Østbakken, Barth & Schøne, 2014). Og det synes at være tilfældet her. Der ser således især ud til at være et stort positivt bidrag fra kvindeandel i jobceller, dvs. i arbejdsfunktioner på arbejdssteder, se Tabel 4.2. Jo større kvindeandelen er i jobcellen, jo lavere er lønnen for både

²⁴ De mere detaljerede bidrag fremgår af Tabel 4.1, fase 1 og Tabel 4.2, fase 3.

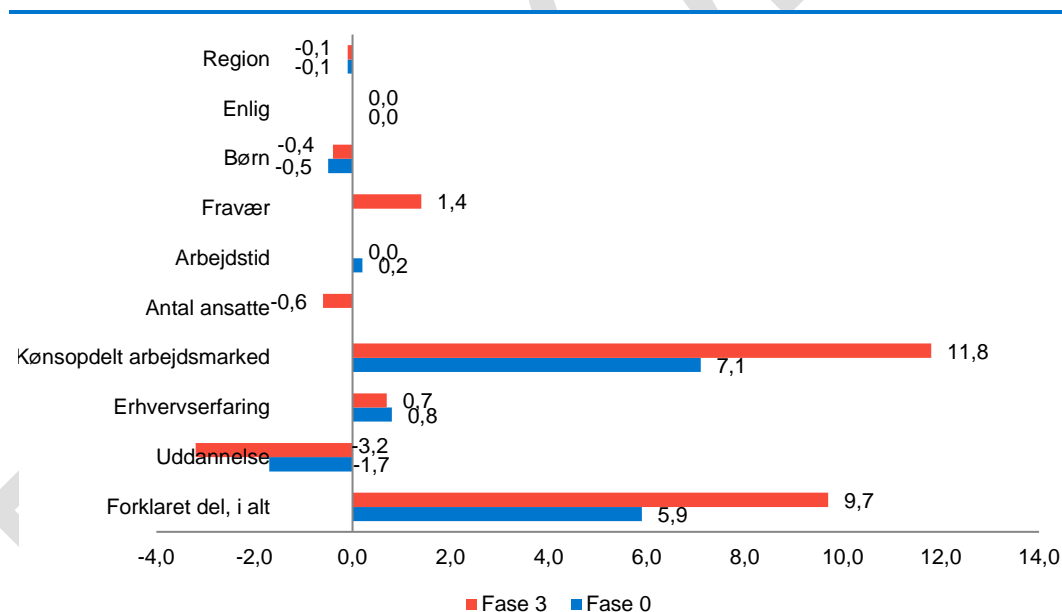
²⁵ Den bagvedliggende 'sande' bruttolønforskel er på 14,5 pct., se også boks 3.2.

kvinder og mænd, se Bilag 2. I lighed med Albæk et al. (2016, 2019) finder vi således tegn på, at der er en markant sammenhæng mellem lønnen (og dermed lønforskellen) og andelen af kvinder inden for fag.

Mulige forklaringer på sammenhængen mellem løn og kvindeandel i jobceller er for de offentlige sektors vedkommende dels historiske årsager, dels at der i kvindefagene er én køber af fagets arbejdskraft, som har stor mulighed for at påvirke lønnen (i nedadgående retning). For den private sektor kan forklaringen for det første være, at løndannelsen foregår mere decentralt i denne sektor og at kvinderne som følge heraf får lavere lønstigninger, fordi de er mindre tilbøjelige til at forhandle løn og er mindre konkurrenceorienterede end mænd. For det andet kan forklaringen være, at kvinder er mere risikoaverse end mænd og derfor i større omfang vælger fag i den private sektor med en stabil, men typisk også en lavere indtjening i gennemsnit. For nærmere beskrivelse af disse mulige forklaringer, se afsnit 2.1.

En del af bidraget fra det kønsopdelt arbejdsmarked stammer fra relativt store positive bidrag fra branche og jobstatus, se Tabel 4.2. Dette understreger, at lønforskellen hænger sammen med både den horisontale (kvinder og mænd er forskelligt fordelt på brancher) og vertikale (mænd er placeret højere i stillingshierarkiet end kvinder) kønsopdeling af arbejdsmarkedet, se Bilag 2.

Figur 5.1 'Forklaret' del, i alt og det samlede bidrag fra hvert karakteristikum til den 'forklarede' del, særskilt for oprindeligt (fase 0¹) og nyt (fase 3) sæt af kontrolvariabler, hele arbejdsmarkedet, 2016. Procent.



Note: ¹ Resultaterne for 'fase 0' er ikke præcis de samme som resultaterne for 2016 i Larsen & Larsen (2018). Det skyldes dels, at observationsantallet er lidt forskelligt i de to opgørelser, se afsnit 3.1, dels at vi har ændret en smule på, hvordan arbejdsfunktioner er kodet.

Der er også en positiv sammenhæng mellem lønforskellen og de to køns gennemsnitlige erhvervs erfaring, hvilket hænger sammen med, at mænd har mere erhvervs erfaring end kvinder, se Bilag 2. Bidraget herfra er stort set det samme uanset om vi anvender det oprindelige eller det ændrede sæt af kontrolvariabler, nemlig 0,7-0,8 pct.

Som noget nyt tyder resultaterne af fase 3 på, at lønforskellen også hænger sammen kønsforskelle mht. fravær og antal ansatte på ens arbejdssted. 1,5 pct. af lønforskellen kan tilskrives, at kvinder har mere fravær end mænd, som typisk påvirker lønnen negativt, se Bilag 2. Mulige forklaringer på den negative sammenhæng mellem løn og fravær er dels, at humankapitel afskrives under fraværperioder – jo længere fravær, jo større afskrivning, dels at længere (frivillige) fraværperioder kan være et signal om, at man er mindre dedikeret til sit job, se afsnit 0.

Antallet af ansatte på arbejdsstedet er i gennemsnit større for kvinder end for mænd og lønnen tenderer mod at være højere på større arbejdssteder, se Bilag 2. Resultaterne af fase 3 viser, at hvis kvinder havde været ansat på arbejdssteder, der i gennemsnit havde samme (mindre) størrelse som mænd, ville lønforskellen have været 0,6 pct. større. En medvirkende årsag til det negative bidrag fra antallet af ansatte på arbejdsmarkedet som helhed kan være, at vi finder de største arbejdspladser især i regioner, men også i kommuner (ikke vist).²⁶ Man kan forestille sig, at løndannelsen er mere systematisk på større (især offentlige) arbejdssteder, hvilket potentielt kan bidrage til, at (uhensigtsmæssige) lønforskelle mellem de to køn undgås.

'Uforklaret' del

Den tilbageværende lønforskel efter, at vi har taget højde for de to køns forskellige fordeling på de undersøgte karakteristika, er på kun 1,8 pct. ifølge fase 3 mod 5,5 pct. ifølge fase 0, se Figur 5.2.

Brugen af en ny fremgangsmåde til udførelse af dekomponeringsanalysen i fase 3 (for beskrivelse heraf, se afsnit 3.2) gør det muligt for os at få viden om, hvad den 'uforklarede' del består i. Resultaterne heraf (og herunder bidragenes størrelse) skal dog fortolkes med varsomhed, se også afsnit 3.2. Vi fokuserer derfor primært på bidragenes fortegn.

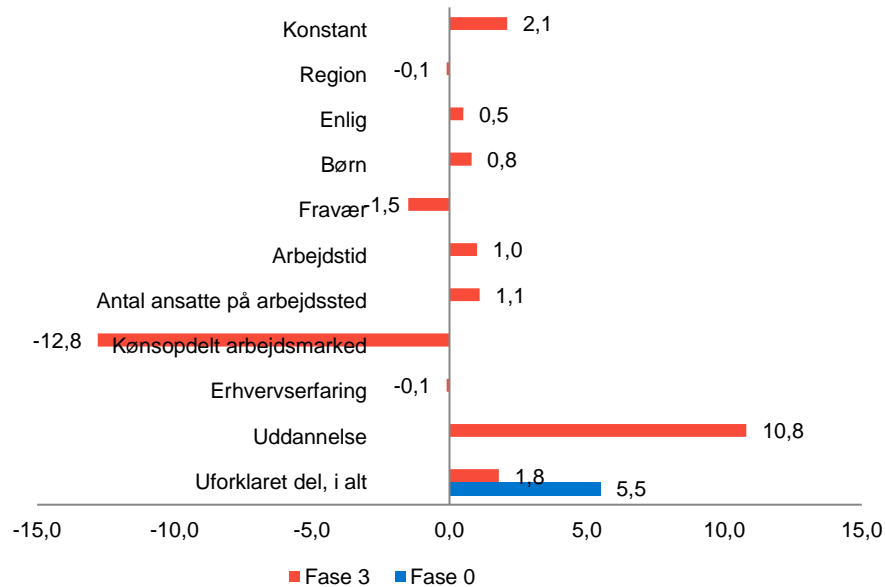
Mænd ser ud til at få et større lønafkast end kvinder, når de har 'samme' uddannelse, se Figur 5.2. De mere detaljerede fase 3-resultater (se Tabel 4.2) viser, at forskellen primært kan tilskrives, at det marginale afkast af uddannelseslængde er større for mænd end for kvinder. Dette svarer til resultatet i Larsen & Houlberg (2013b). Det samlede bidrag fra uddannelse er relativt stort, men en medvirkende årsag synes dog at være, at de to køn har forskellige uddannelser *inden for* de uddannelsesretninger, vi tager højde for her.²⁷

Samtidig får kvinder imidlertid et større lønafkast end mænd, når deres karakteristika for 'kønsopdelt arbejdsmarked' er de samme. De mere detaljerede fase 3-resultater (se Tabel 4.2) tyder på, at forskellen bl.a. kan tilskrives kønsforskelle i det marginale lønafkast af kvindeandel i den sektor, hvor man er ansat. Mænds (og ikke kvinder) får tilsyneladende lavere løn, jo større kvindeandelen i sektoren er (se Bilag 2). Forskellen kan desuden tilskrives kønsforskelle i lønafkastet af jobstatus: Det negative lønafkast, der er forbundet med at være placeret "lavt" i jobhierarkiet (at være 'øvrig medarbejder') er mindre for kvinder end for mænd (se Bilag 2).

²⁶ Vi har undersøgt, om resultatet vedrørende antal ansatte ændrer sig, når kvindeandel i sektor erstattes af dummyvariabler for sektor. Det viser sig ikke at være tilfældet.

²⁷ Vi har testet, hvad det betyder for bidraget til den 'uforklarede' del fra uddannelse, hvis de anvendte oplysninger om uddannelsesretning erstattes med mere detaljerede uddannelsesoplysninger. Resultatet er, at bidraget til den 'uforklarede' del fra uddannelse næsten halveres (bidraget fra uddannelseslængde bliver markant mindre, mens bidraget fra uddannelsesretning bliver større). Samtidig bliver det negative bidrag fra kønsopdelt arbejdsmarked større.

Figur 5.2 'Uforklaret' del, i alt, særskilt for oprindeligt (fase 0) og ændret (fase 3) sæt af kontrolvariabler og konstantled samt det samlede bidrag fra hvert karakteristikum til den uforklarede del, ændret sæt af kontrolvariabler (fase 3), hele arbejdsmarkedet, 2016. Procent.



Mænd ser ud til at blive 'straffet' lidt mere lønmæssigt for fravær end kvinder, hvilket kan hænge sammen med, at signaleffekten af langvarigt fravær er større for mænd end for kvinder, se afsnit 2. Derimod har mænd tilsyneladende et lidt større lønafkast end kvinder, når de to køn har samme karakteristika, hvad angår antal ansatte på arbejdsstedet, ugentlig arbejdstid, samlivssituation og børn. Sidstnævnte er i tråd med, at forældreskab ifølge Simonsen & Skipper (2008) reducerer timelønnen for mødre og forøger timelønnen for fædre.

Endelig er der en tilbageværende konstant på 2,1 pct., som vi tager som udtryk for, at 2,1 pct. af lønforskellen kan tilskrives forhold, som vi ikke inkluderer i modellen. Det kan fx vedrøre kønsforskelle mht. præferencer, evner, personlighedstræk og forekomsten af ubevidst bias, se afsnit 2.10.

5.2 Den private sektor 2016

I dette afsnit fokuserer vi på resultaterne for den private sektor, hvor vi med det ændrede sæt af kontrolvariabler kan forklare en væsentlig større del af lønforskellen (78 pct.) end det var muligt med de oprindelige kontrolvariabler (19 pct.). De detaljerede fase 3-resultater for 2016 for alle fire sektorer fremgår af Bilag 4.

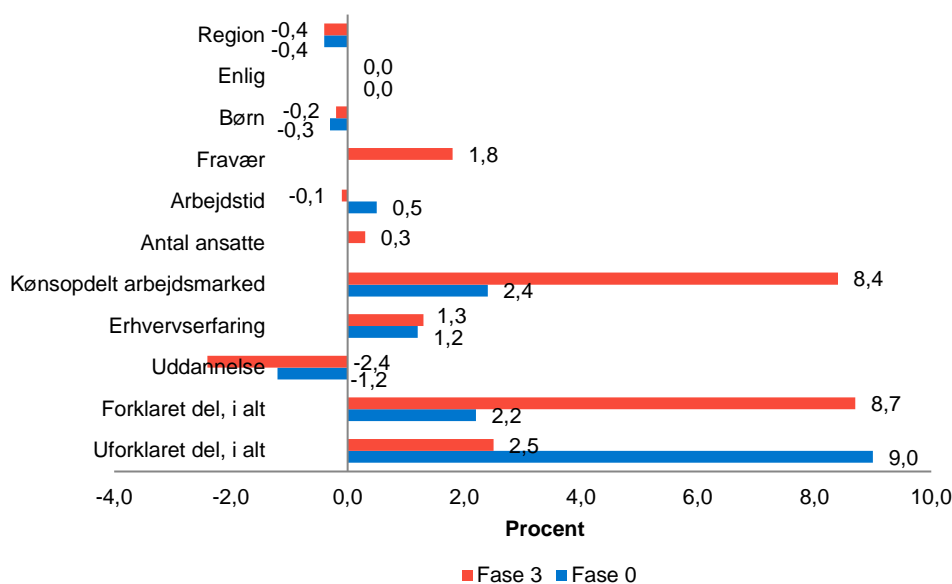
Vi har i vores gennemgang af resultaterne for den private sektor valgt alene at fokusere på de detaljerede resultater under den 'forklarede' del. Det skyldes, at konstantleddet er relativt stort, se Bilag 4, hvilket gør det vanskeligt at give en meningsfuld fortolkning af resultaterne af den uforklarede del.

Bruttolønforskellen mellem kvinder og mænd i den private sektor er – ifølge de gennemførte dekomponeringsanalyser – på 11,2 pct. i 2016.²⁸ Heraf kan vi 'forklare' de 8,7 pct. ifølge fase 3 mod kun 2,2 pct. ifølge fase 0, se Figur 5.3. Den større 'forklarede' del kan især tilskrives, at vi i fase 3 får et væsentligt større bidrag fra kønsopdelt arbejdsmarked (8,4 vs. 2,4 pct.). Ser vi på de detaljerede fase 3-resultater (se Bilag 4), finder vi, at størstedelen af dette bidrag stammer fra kvindeandel inden for jobceller, altså andelen af kvinder pr. arbejdsfunktion pr. arbejdssted. Vi er således i stand til at 'forklare' en større del af lønforskellen i den private sektor, når vi bevæger os ind på løndannelsen på arbejdsstedsniveau. Dette er i tråd med, at decentral løndannelse er udbredt i den private sektor. Med andre ord: Jo mindre overenskomstdækket et område er, jo mere betyder det, der foregår på arbejdsstedet for lønforskellen mellem kvinder og mænd.

Det negative bidrag fra uddannelse er også større i fase 3 (-2,4 pct. vs. -1,2 pct. i fase 0). Der er med andre ord en mere udtalt sammenhæng mellem lønforskel og uddannelse, når vi tager højde for både uddannelseslængde og -retning end når vi blot tager højde for uddannelsesniveau.

Det samlede bidrag fra uddannelse og kønsopdelt arbejdsmarked er samlet set væsentligt større baseret på fase 3 end på fase 0 (6,0 vs. 1,2 pct.). Det vil sige, at der i den private sektor er en væsentlig sammenhæng mellem lønforskellen og de to køns forskellige uddannelsesvalg og forskellige placering på det private arbejdsmarked.

Figur 5.3 'Forklaret' og 'uforklaret' del, i alt og det samlede bidrag fra hvert karakteristikum til den 'forklarede' del, særskilt for oprindeligt (fase 0¹) og ændret (fase 3) sæt af kontrolvariabler, privat sektor, 2016. Procent.



Note: ¹ Resultaterne for 'fase 0' er ikke præcis de samme som resultaterne for 2016 i Larsen & Larsen (2018). Det skyldes dels, at observationsantallet er lidt forskelligt i de to opgørelser, se afsnit 3.1, dels at vi ændret en smule på, hvordan vi koder arbejdsfunktioner.

Inddragelse af oplysninger om fravær bidrager også til, at vi kan 'forklare' en større del af lønforskellen i den private sektor. Lønforskellen i denne sektor hænger således også sammen

²⁸ Den bagvedliggende 'sande' bruttolønforskelle er på 12,8 pct., se også afsnit 1.2 og boks 3.2.

med, at kvinder har større fravær end mænd. Bidraget herfra er 1,8 pct. De detaljerede resultater i Bilag 4 viser, at dette bidrag især stammer fra kønsforskelle i barselsfravær. Resultaterne tyder på, at der især i den private sektor er en 'lønstraf' forbundet med barselsfravær.

5.3 Udviklingen i perioden 2010-2017

Fra slutningen af 1970'erne frem til 2007 ændrede forskellen mellem kvinders og mænds time-løn sig ikke nævneværdigt i Danmark (Rosholm & Smith, 1996; Datta Gupta m.fl., 2006; Larsen & Larsen, 2018). Imidlertid faldt denne lønforskel relativt markant i perioden 2007-2016. Larsen & Larsen (2018) viser, at faldet primært var et resultat af, at den 'forklarede' del af lønforskellen blev mindre, det vil her sige, at kvinders karakteristika udviklede sig mere gunstigt end mænds: Kvinder bevægede sig i højere grad end mænd opad i stillingshierarkiet og deres forspring med hensyn til uddannelseslængde blev forøget. Derimod ændrede størrelsen af den uforklarede lønforskel sig ikke nævneværdigt.

Vi ser i dette afsnit på, om vi med det ændrede sæt af kontrolvariabler kan få mere viden om, hvorfor lønforskellen er faldet i de senere år. Vi fokuserer på udviklingen siden 2010 med afsæt i data for 2010, 2013, 2016 og 2017. Detaljerede resultater for de fire år fremgår af bilagstabel 5.2 i Bilag 5. Af bilagstabel 5.1 i samme bilag er vist den bagvedliggende 'sande' bruttolønfor-skel i hvert af disse år, se også boks 3.2. Det fremgår, at lønforskellen er faldet fra 16,9 pct. i 2010 til 14,5 pct. i 2016, men er på stort set samme niveau i 2017 som i 2016 (14,4 pct.).²⁹ Det vil sige, at lønforskellen er faldet med 2,5 procentpoint fra 2010 til 2017.

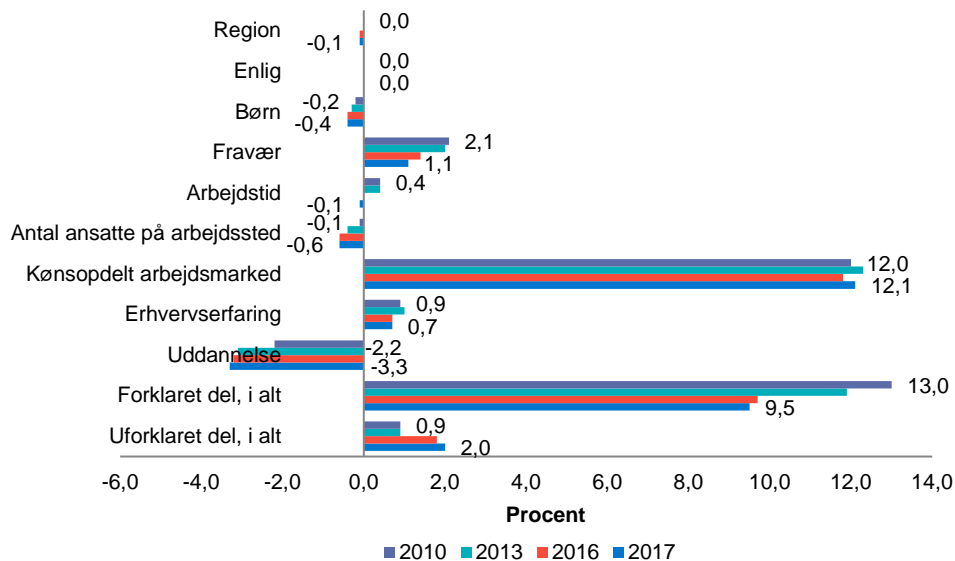
I vores beskrivelse af baggrunden for faldet i lønforskellen i de senere år har vi valgt alene at fokusere på de detaljerede resultater under den 'forklarede' del, fordi den detaljerede udvikling under den uforklarede del er vanskelig at fortolke. Det skyldes dels, at der over tid sker betydelige forskydninger i bidragene fra de enkelte variabler under denne del, som er vanskelige at fortolke, dels at konstantleddet både ændrer størrelse og fortegn, se bilagstabel 5.2 i Bilag 5.

Den del af lønforskellen, som vi kan forklare, falder over tid – fra 13,0 pct. i 2010 til 9,5 pct. i 2017, dvs. et fald på 3,5 procentpoint. Dette fald hænger i sagens natur bl.a. sammen med, at lønforskellen ligeledes falder i denne periode. Faldet kan imidlertid også tilskrives, at vi kan forklare en mindre *andel* af forskellen. Dette fremgår også af resultaterne i afsnit 4.6, der viser, at vi med det ændrede sæt af kontrolvariabler kan forklare 82 pct. af lønforskellen i 2017 mod 93 pct. i 2010.

Vores resultater viser, at det kønsopdelte arbejdsmarked spiller en stor rolle for lønforskellen i alle de undersøgte år, men at betydningen heraf samlet set ikke ændrer sig nævneværdigt fra 2010 til 2017, se Figur 5.4. Dette svarer til resultater i Albæk et al. (2019).

²⁹ De beregnede 'sande' bruttolønforskelle for især 2010, men også 2013, afviger fra de tilsvarende lønforskelle i Larsen & Houllberg (2013a) og Larsen & Larsen (2018). Det skyldes, at det anvendte observationsantal er forskelligt, se afsnit 3.1.

Figur 5.4 'Forklaret' og 'uforklaret' del, i alt og det samlede bidrag fra hvert karakteristikum til den 'forklarede' del baseret på ændret sæt af kontrolvariabler (fase 3), hele arbejdsmarkedet, 2010, 2013, 2016 og 2017.¹ Procent.



Note: ¹ For overskuelighedens skyld er der alene sat talværdier på for 2010 og 2017.

Faldet i den 'forklarede' andel kan derimod bl.a. tilskrives ændringer i bidraget fra uddannelse. Det negative bidrag fra uddannelse er steget fra -2,2 pct. i 2010 -3,3 pct. i 2017. Dette afspejler, at kvinders forspring mht. uddannelseslængde er forøget i perioden, se også de detaljerede resultater i bilagstabel 5.2 i Bilag 5.

Bidraget fra fravær har også ændret sig i perioden. Dette bidrag er faldet fra 2,1 til 1,1 pct. fra 2010 til 2017. De detaljerede resultater i bilagstabel 4.2 i Bilag 5 viser, at det er bidraget fra barselsfravær, der er faldet. Vi har kigget nærmere på de bagvedliggende resultater (ikke vist). Heraf fremgår, at kønsforskellen i omfanget af barselsfravær rent faktisk er steget fra 2010 til 2017, hvilket skulle tale for, at fravær skulle spille en større rolle i 2017 end i 2010. Betydningen heraf overgås imidlertid af, at den 'lønstraf', der er forbundet med barselsfravær i 2010 på arbejdsmarkedet som helhed, ikke er tilstede i 2017.³⁰ En medvirkende årsag er, at det positive bidrag fra barselsfravær i den private sektor er faldet i perioden (ikke vist). Dette kan tale for, at konjunktursituationen i perioden har medvirket til den fundne udvikling. Det vil sige, at "lønstraffen" for barselsfravær potentielt er mindre, når private virksomheder er i en mere gunstig økonomisk situation.

De skitserede resultater er samtidig væsentlige årsager til, at vi kan forklare en mindre andel af lønforskellen i 2017 end i 2010:

- Det negative bidrag fra uddannelse stiger som følge af kvinders stigende forspring med hensyn til uddannelseslængde.
- Det positive bidrag fra barselsfravær falder som følge af et fald i den negative løneffekt af barselsfravær, hvilket potentielt bl.a. kan være en følge af konjunkturudviklingen i perioden.

³⁰ Den negative løneffekt af barselsfravær i de fem forudgående år for mænd (hvis β -værdi, der anvendes til at beregne den forklarede del, jf. formel A7 i Bilag 1) er blevet mindre fra 2010 til 2017 og er statistisk insignifikant i 2017. For kvinder gælder, at løneffekten af barselsfravær går fra at være negativ i 2010 til samlet set at være positiv i 2017.

Det er dog vigtigt at holde fast i, at vi også for 2017 kan forklare en stor andel af lønforskellen med det ændrede sæt af kontrolvariabler, nemlig 82 pct.

FORTROLIGT

Bilag 1 Blinder-Oaxaca-dekomponeringsanalyse

For at beregne, i hvor høj grad den observerede, gennemsnitlige lønforskel mellem mænd og kvinder hænger sammen med udvalgte forklarende variabler, anvender vi den traditionelle Blinder-Oaxaca-dekomponeringsmetode (se fx Oaxaca & Ransom, 1994). Ved hjælp af denne metode dekomponerer vi lønforskellen, som er defineret som den procentsats, kvinders løn skal stige med, for at de får samme gennemsnitsløn som mænd. Vi foretager analysen i to trin. Først udfører vi en regressionsanalyse for hhv. mænd og kvinder, og dernæst dekomponerer vi lønforskellen ved hjælp af lønregressionerne. Beregningsformlen for lønregressionerne er:

$$(A.1) \quad \ln(w_m) = X_m\beta_m + \varepsilon_m$$

$$(A.2) \quad \ln(w_k) = X_k\beta_k + \varepsilon_k$$

Den afhængige variabel er logaritmen til timelønnen for hhv. mænd ($\ln(w_m)$) og kvinder ($\ln(w_k)$).³¹ De forklarende variabler (fx uddannelse, erhvervs erfaring, sektor og arbejdsfunktion) er angivet ved vektorerne X_m for mænd og X_k for kvinder. De forklarende variabelers betydning for hhv. mænds og kvinders timeløn måles ved β_m og β_k . Disse størrelser er koefficienter til de forklarende variabler. Da de forklarende variabler ikke forklarer timelønnen helt præcist, indeholder regressionsligningerne også et fejlelement, ε_m og ε_k .

Lønforskellen mellem kvinder og mænd er defineret som:

$$(A.3) \quad F_{mk} = (\bar{w}_m - \bar{w}_k) / \bar{w}_k,$$

hvor \bar{w}_m og \bar{w}_k angiver det geometriske gennemsnit af hhv. mænds og kvinders løn. Lønforskellen (ganget med 100) angiver, hvor mange pct. mænd gennemsnitligt tjener mere end kvinder.

Resultaterne af regressionsanalysen gør det muligt at opdele lønforskellen i en forklaret og en uforklaret del. Med andre ord kan lønforskellen dekomponeres på følgende måde: Først omskrives lønforskellen ved hjælp af en logaritme-transformation til:

$$(A.4) \quad F_{mk} \cong \ln(F_{mk} + 1) = \ln(\bar{w}_m) - \ln(\bar{w}_k)$$

Logaritmen til hhv. mænds og kvinders gennemsnitsløn, $\ln(\bar{w}_m)$ og $\ln(\bar{w}_k)$, findes ud fra regressionsligningerne:

$$(A.5) \quad \ln(\bar{w}_m) = \bar{X}_m\hat{\beta}_m$$

$$(A.6) \quad \ln(\bar{w}_k) = \bar{X}_k\hat{\beta}_k,$$

hvor \bar{X}_m og \bar{X}_k angiver gennemsnittet af de forklarende variabler for hhv. mænd og kvinder, og $\hat{\beta}_m$ og $\hat{\beta}_k$ er de estimerede koefficienter. Lønforskellen i (A.4) kan nu dekomponeres på følgende måde:

$$(A.7) \quad \ln(F_{mk} + 1) = \bar{X}_m\hat{\beta}_m - \bar{X}_k\hat{\beta}_k = \bar{X}_m\hat{\beta}_m - \bar{X}_k\hat{\beta}_k + \bar{X}_k\hat{\beta}_m - \bar{X}_k\hat{\beta}_m$$

³¹ Det er standard i lønanalyser at anvende logaritmen til lønnen for at opfylde forudsætningen om en normalfordelt funktionsform. Den statistiske fordel herved er, at lønninger, som er meget høje i forhold til gennemsnittet, kommer til at vægte mindre i analysen.

$$= (\bar{X}_m - \bar{X}_k) \hat{\beta}_m + \bar{X}_k (\hat{\beta}_m - \hat{\beta}_k)$$

Det første led (efter lighedstegnet i anden linje) er lig med den 'forklarede' del, dvs. kønsforskellen mellem gennemsnittet af forklarende variabler ganget med koefficienterne fra mænds regressionsligning. Det andet led (i A.7) er lig med den uforklarede del, dvs. kønsforskelle i aflønning af samme karakteristika (inkl. et konstantled), der her er beregnet som forskellen mellem koefficienterne i mænds og kvinders regressionsligninger ganget med kvinders gennemsnitlige forklarende variabler.

Det er imidlertid også muligt at lave den 'omvendte' dekomponering, som er baseret på mænds karakteristika og kvinders koefficienter. Formlen for denne dekomponering er:

$$(A.8) \quad \ln(F_{mk} + 1) = (\bar{X}_k - \bar{X}_m) \hat{\beta}_k + \bar{X}_m (\hat{\beta}_k - \hat{\beta}_m)$$

I Larsen (2010) er vist, hvilke resultater der opnås, hvis den 'omvendte' dekomponering er anvendt.

FORTROLIGT

Bilag 2 Beskrivende statistik og resultater af regressionsanalyser

Bilagstabel 2.1 Beskrivende statistik (gennemsnitsværdier¹, X) og fase 3-resultater af regressionsanalyser, hvor logaritmen til timeløn er den afhængige variabel (koefficienter, β) særskilt for kvinder og mænd, 2016.

	X (mænd)	X (kvinder)	β (mænd)	β (kvinder)
<i>Uddannelseslængde, år (gns.)</i>	17,5	17,9	0,005	0,004
<i>Uddannelsesretning</i>				
Grundskole	13,9	8,6	0,093	0,048
Gymnasiale uddannelser	5,6	4,7	0,062	0,026
Forberedende uddannelser	0,1	0,0	0,016 ^{l.s.}	-0,005 ^{l.s.}
Undervisning og læring	3,0	6,1	-0,039	-0,002 ^{l.s.}
Humanistisk	1,8	3,9	-0,077	-0,022
Audio-visuel teknik og medieproduktion	0,9	0,5	-0,057	-0,076
Kunstnerisk	0,5	1,0	-0,085	-0,076
Samfundsvidenskab	3,0	3,6	-0,004	0,024
Erhvervsøkonomi, administration og jura	16,4	24,0	0,004	-0,034
Naturvidenskab	1,7	1,5	-0,037	0,040
Informations- og kommunikationsteknologi (IKT)	2,4	0,7	-0,008	0,020
Teknik, teknologi og industriel produktion	13,1	3,2	-0,050	-0,059
Mekanik, jern og metal	12,5	0,4	-0,079	-0,041
Bygge- og anlægsteknik	7,4	1,6	-0,070	-0,017
Jordbrug, skovbrug og fiskeri	2,1	1,1	-0,094	-0,076
Social og sundhed	6,4	33,2	0,003	0,014
Service	1,5	3,1	-0,091	-0,093
Forsvar, politi og sikkerhed	2,1	0,4	-0,114	-0,138
Transport	1,3	0,1	-0,053	-0,080
Studiefag uoplyst	3,1	1,8	-0,074	-0,100
Uoplyst mv.	1,2	0,7	0,753	0,647
I alt	100,0	100,0		
<i>Erhvervs erfaring</i>				
Erhvervs erfaring, år (gns.)	18,0	16,6	0,019	0,019
Erhvervs erfaring ² , år (gns.)	450,2	386,5	0,000	0,000
<i>Kvindeandele</i>				
Kvindeandel, sektor	45,0	58,9	-0,107	0,007
Kvindeandel, jobcelle	25,4	76,8	-0,093	-0,102
<i>Branche</i>				
Anden uoplyst	0,0	0,0	0,049 ^{l.s.}	-0,002 ^{l.s.}
Andre serviceydelser mv.	1,2	1,5	-0,059	0,001 ^{l.s.}
Bygge og anlæg	8,7	0,8	0,000 ^{l.s.}	-0,028

	X (mænd)	X (kvinder)	β (mænd)	β (kvinder)
Ejendomshandel og udlejning	1,1	0,8	-0,072	-0,029
Elektronikindustri	0,9	0,5	-0,005 ^{l.s.}	-0,067
Energiforsyning	0,8	0,4	0,151	0,139
Finansiering og forsikring	4,2	3,8	0,210	0,196
Forlag, tv og radio	1,3	0,9	0,053	0,028
Forskning og udvikling	0,8	0,6	0,042	0,021
Fremstilling af elektrisk udstyr	0,6	0,3	-0,016	-0,078
Føde-, drikke- og tobaksvarerindustri	2,7	1,2	-0,008 ^{l.s.}	-0,044
Handel	13,1	8,1	-0,065	-0,087
Hoteller og restauranter	1,5	1,7	-0,203	-0,162
It- og informationstjenester	2,8	0,9	0,150	0,117
Kemisk industri	0,9	0,5	0,044	0,068
Kultur og fritid	1,4	1,4	-0,117	-0,094
Landbrug, skovbrug og fiskeri	0,0	0,0	-0,080	0,011 ^{l.s.}
Maskinindustri	4,1	1,0	0,029	0,025
Medicinalindustri	1,3	1,4	0,154	0,181
Metalindustri	2,5	0,4	-0,045	-0,052
Møbel og anden industri mv.	1,6	0,7	-0,013 ^{l.s.}	0,050
Offentlig administration, forsvar og politi	7,4	7,2	-0,088	-0,022
Olieraffinaderier mv.	0,1	0,0	0,287	0,355
Plast-, glas- og betonindustri	1,7	0,5	-0,041	-0,045
Rejsebureauer, rengøring, anden operationel service	7,1	6,6	-0,113	-0,085
Reklame og øvrig erhvervsservice	0,7	0,7	0,003 ^{l.s.}	-0,027
Rådgivning mv.	4,0	2,6	0,081	0,066
Råstofindvinding	0,2	0,0	0,139	0,103
Sociale institutioner	6,1	26,4	-0,141	-0,096
Sundhedsvæsen	2,9	12,9	-0,029	-0,051
Tekstil- og læderindustri	0,2	0,1	-0,032	-0,063
Telekommunikation	0,9	0,4	-0,002 ^{l.s.}	-0,004
Transport	6,5	2,0	-0,066	-0,043
Transportmiddelindustri	0,4	0,1	-0,066	-0,055
Træ- og papirindustri, trykkerier	1,1	0,3	-0,073	-0,075
Undervisning	8,5	13,0	-0,060	-0,023
Uoplyst aktivitet	0,0	0,0	0,057 ^{l.s.}	-0,118
Vandforsyning og renovation	0,9	0,2	-0,061	-0,013
I alt	100,0	100,0		
<i>Jobstatus</i>				
Elever og lærlinge	1,2	2,0	-0,503	-0,469
Ledere og mellemedere	11,0	5,6	0,240	0,256
Medarbejdere med særligt ansvar	5,0	4,1	0,058	0,087
Øvrige medarbejdere	82,8	88,3	-0,082	-0,033
Andet ²	0,0	0,0	0,287	0,159
I alt	100,0	100,0		

	X (mænd)	X (kvinder)	β (mænd)	β (kvinder)
<i>Antal ansatte på arbejdspladsen</i>				
0 ansatte	3,5	1,6	-0,046	-0,014
1-9 ansatte	6,0	7,1	-0,084	-0,044
10-49 ansatte	31,6	27,1	-0,036	-0,026
50-249 ansatte	30,7	28,2	0,026	0,014
250-999 ansatte	13,8	10,2	0,072	0,046
1.000 ansatte eller derover	14,5	25,7	0,068	0,024
I alt	100,0	100,0		
<i>Antal aftalte timer pr. uge</i>				
Under 30 pr. uge	12,7	17,2	-0,036	0,005
30-36 timer pr. uge	12,2	25,5	0,010	0,012
37 timer pr. uge	60,9	54,2	0,071	0,039
38 timer eller derover pr. uge	14,2	3,2	-0,045	-0,056
I alt	100,0	100,0		
<i>Fravær³</i>				
Fraværstimer/betalte timer (gns.)	3,4	5,8	-0,221	-0,125
Sygefravær 2015, antal dage (gns.)	6,4	12,0	-0,015	-0,005
Sygefravær 2014, antal dage (gns.)	6,3	11,7	-0,015	-0,006
Sygefravær 2013, antal dage (gns.)	6,2	10,8	-0,013	-0,008
Sygefravær 2012, antal dage (gns.)	6,3	10,8	-0,013	-0,007
Sygefravær 2011, antal dage (gns.)	6,4	10,6	-0,018	-0,007
Barselsfravær 2015, antal dage (gns.)	1,9	21,8	-0,010	-0,006
Barselsfravær 2014, antal dage (gns.)	1,7	22,0	-0,006	0,005
Barselsfravær 2013, antal dage (gns.)	1,7	22,2	-0,006	0,001
Barselsfravær 2012, antal dage (gns.)	1,7	22,0	-0,006	-0,001
Barselsfravær 2011, antal dage (gns.)	1,7	23,3	0,000 ^{1.S.}	0,001
<i>Familieforhold</i>				
Antal hjemmeboende børn under 18 år (gns.)	0,9	1,0	0,020	0,011
Yngste barn, 0-2 år	11,8	10,8	-0,002	-0,002
Yngste barn, 3-6 år	9,9	11,3	0,006	0,008
Yngste barn, 7-17 år	18,8	24,1	0,009	0,004
Enlig	28,3	27,3	-0,011	0,000 ^{1.S.}
<i>Bopælsregion</i>				
Hovedstaden	33,9	34,9	0,042	0,046
Midtjylland	22,5	21,4	-0,012	-0,013
Nordjylland	9,7	9,5	-0,028	-0,028
Sjælland	13,9	14,2	0,014	0,009
Syddanmark	20,1	20,0	-0,015	-0,014
I alt	100,0	100,0		
Konstant			4,580	4,560

Anm.: ^{1.S.} angiver, at beta-koefficienten ikke er statistisk signifikant.

Note: ¹ Gennemsnitsværdier er angivet i procent, medmindre andet er angivet. ² Personer, der arbejder med produktion/ekspedition/administration" (en jobstatus-kategori, der ikke eksisterer mere, men fortsat anvendes i enkelte tilfælde). ³ Koefficienter for hhv. sygefravær og barselsfravær at ganget med 1.000.

Bilag 3 Fraværsvariabler

De fraværsvariabler, der anvendes som kontrolvariabler, er beskrevet i det følgende.

Fravær i ansættelsen i året

Vores opgørelse af fravær i ansættelsen er baseret på Lønstatistikken. Vi har anvendt variabelen TIMFRA, der angiver antallet af betalte fraværstimer i den aktuelle ansættelse i året. Betalte fraværstimer er fravær i forbindelse med egen sygdom, børns sygdom, barsel, ulykke og andet fravær, hvortil der ydes betaling, herunder feriefridage, børnefridage, omsorgsdage, senior-dage og lignende fravær (Danmarks Statistik, 2020e).

Vi har divideret antallet af betalte fraværstimer med TIMBET, som er det antal timer, som lønmodtageren reelt er blevet betalt for. Det vil sige det normale antal løntimer inkl. betalte overarbejdstimer (Danmarks Statistik, 2020d).

Danmarks Statistik vurderer helt overordnet, at fraværet i ansættelsen i året er undervurderet i både den private og de offentlige sektorer. For nærmere beskrivelsen herfra, se Danmarks Statistik (2020e).

Barsels- og sygefravær i de fem forudgående år

Vi har opgjort antallet af dage med henholdsvis barsels- og sygefravær særskilt for hvert år fem år tilbage i tid. Vores opgørelse heraf er baseret på Sygedagpengeregistret. Registeret er et udbetalingsregister, hvor alle hændelser, der har bevirket udbetaling af dagpenge eller refusion for udbetalt løn i forbindelse med barsel eller sygdom er registreret. Vi har anvendt variabelen SAGSART, der angiver hvilke regler den pågældende sag hører under, til at skelne mellem barsels- og sygefravær (Danmarks Statistik, 2020c). Vores inddeling er baseret på inddelingen i Styrelsen for Arbejdsmarked og Rekruttering (2015).

Vores opgørelse af sygefravær omfatter følgende sagsnumre:

- 11: Efter arbejdsgiverperioden (findes fra 1995)
- 12: Forsikrede arbejdsgivere (findes fra 1995)
- 13: Ingen arbejdsgiverperiode (findes fra 1995)
- 14: Kronisk syge (findes fra 1995)
- 15: Ophørt men ikke berettiget til løn. Ansat over 8 uger (findes fra 2009)
- 16: Forskudsvis udlagte dagpenge (findes fra 1995)
- 17: Arbejdsløse (findes fra 1995)
- 19: Ophørt men ikke berettiget til løn. Ansat over 13 uger (findes fra 2009)
- 21: Efter 2. uge, selvstændige (findes fra 1995)
- 22: Frivilligt sikrede selvstændige (findes fra 1995)
- 29: Kronisk syge, selvstændige, § 58a (findes fra 2012)
- 40: Flexjob, ansat efter 31.12.2012 (findes fra 2013)
- 41: Arbejdsskade, selvstændige (findes fra 1995)
- 42: Arbejdsskade, lønmodtagere (findes fra 1995)
- 43: Flexjob, ansat inden 31.12.2012 (findes fra 1998)
- 49: For sen anmeldelse, ingen dagpenge, medtages ej (findes fra 1995)

Vores opgørelse af barselsfravær omfatter følgende sagsnumre:

- 07: Orlov før adoption, lønmodtagere (findes fra 2005)
- 08: Orlov før adoption, selvstændige (findes fra 2005)
- 09: Ferieydelse, ej optjent feriepenge under barselsorlov (findes fra 2005)
- 10: Fødsel, lønmodtagere m/k (findes fra 1995)
- 18: Pasning syge børn, lønmodtagere (findes fra 1998)
- 20: Fødsel, selvstændige (findes fra 1995)
- 23: Pasning syge børn, selvstændige (findes fra 1998)
- 24: Fødsel, lønmodtagere, 32 uger fælles orlov (findes fra 2002)
- 25: Fødsel, selvstændige, 32 uger fælles orlov (findes fra 2002)
- 26: Adoption, lønmodtagere, 32 uger fælles orlov (findes fra 2002)
- 27: Adoption, selvstændige, 32 uger fælles orlov (findes fra 2002)
- 35: Retsb. udskudt orlov 8-13 uger inden 9 år, lønmodtagere (findes fra 2002)
- 36: Retsb. udskudt orlov 8-13 uger inden 9 år, selvstændige (findes fra 2002)
- 37: Aftaleb. udskudt orlov 32 uger inden 9 år, lønmodtagere (findes fra 2002)
- 38: Aftaleb. udskudt orlov 32 uger inden 9 år, selvstændige (findes fra 2002)
- 45: Bortadoption/dødfødsel mv., lønmodtagere (findes fra 2002)
- 46: Bortadoption/dødfødsel mv., selvstændige (findes fra 2002)
- 47: Graviditetsbetinget sygdom, lønmodtagere (findes fra 2002)
- 48: Graviditetsbetinget sygdom, selvstændige (findes fra 2002)
- 50: Adoption, lønmodtagere (findes fra 1995)
- 55: Forlængelse genoptaget arbejdstid, lønmodtagere (findes fra 2002)
- 56: Forlængelse genoptaget arbejdstid, selvstændige (findes fra 2002)
- 60: Adoption, selvstændige (findes fra 1995)
- 65: Graviditet, lønmodtagere (findes fra 1995)
- 70: Graviditet, selvstændige (findes fra 1995)
- 75: 2 uger efter fødsel, lønmodtagere (findes fra 1995)
- 77: 2 uger efter adoption, lønmodtagere (findes fra 2000)
- 80: 2 uger efter fødsel, selvstændige (findes fra 1995)
- 81: Fædreorlov uge 25-26, selvstændige (1998-2001)
- 82: 2 uger efter adoption, selvstændige (2000-2002)

Kun syge- og barselsfravær, hvortil der er knyttet en dagpengeret, indgår. Dette indebærer bl.a., at der er forskellige typer af barselsfravær, der ikke indgår i opgørelsen. For eksempel indgår offentlige ansatte kvinders 2-4 ugers ekstra graviditetsorlov ikke, fordi der ikke hverken dagpengeret eller dagpengefusion knyttet til denne del af orloven. Retten til 32 ugers forældreorlov (efter de 32 uger til deling med dagpenge) er også uden dagpengeret og indgår derfor heller ikke. Uden fra en antagelse om, at især kvinder benytter sig af denne ret til 32 ugers forældreorlov, må primært kvinders barselsfravær formodes at være underregistreret.

En anden fejlkilde er, at der ifølge Danmarks Statistik ikke er oplysninger om deltidssygdom og deltidbarsel i registret.

Bilag 4 Resultater for sektorer for 2016

Bilagstabel 4.1 Bruttolønforskelle og bidrag fra karakteristika for 'forklaret' og 'uforklaret' del særskilt for sektorer, 2016. Fase 3. Procent.

	Privat	Stat	Regional	Kommunal
Bruttolønforskelle	11,2	5,3	16,9	5,2
'Forklaret' del, i alt	8,7	1,4	10,9	3,1
Uddannelse	-2,4	-1,8	3,4	0,6
Længde	-1,5	-1,1	3,5	0,8
Retning	-0,9	-0,7	-0,1 ^{l.s.}	-0,2
Erhvervs erfaring	1,3	1,0	-0,6	-0,4
Kønsopdelt arbejdsmarked	8,4	2,9	9,5	2,6
Branche	0,1	0,3	-0,2	1,6
Jobcelle, kvindeandele	7,1	0,3	5,7	-0,3
Jobstatus	1,3	2,4	4,1	1,3
Antal ansatte	0,3	-1,0	0,0 ^{l.s.}	0,1
Arbejdstid	-0,1	0,4	0,4	1,0
Fravær	1,8	0,1	-1,6	-0,9
Fraværstimer	0,2	0,2	0,1	0,1
Sygefravær, 2011-2015	0,3	0,2	0,2	0,3
Barselsfravær, 2011-2015	1,2	-0,3	-1,9	-1,3
Børn	-0,2	-0,2	-0,3	-0,3
Enlig	0,0 ^{l.s.}	0,0	0,0 ^{l.s.}	0,0
Region	-0,4	-0,1	0,0	0,3
'Uforklaret' del, i alt	2,5	4,0	6,0	2,1
Uddannelse	0,6	-14,8	13,9	9,9
Længde	-1,3	-16,1	13,6	11,0
Retning	1,8	1,3	0,3 ^{l.s.}	-1,1
Erhvervs erfaring	-2,5	0,9	0,6 ^{l.s.}	-0,8
Kønsopdelt arbejdsmarked	-6,1	9,5	12,5	0,6
Branche	-0,3 ^{l.s.}	-0,8 ^{l.s.}	5,1	-4,2
Jobcelle, kvindeandele	-4,4	8,4	11,6	5,0
Jobstatus	-1,4 ^{l.s.}	1,9	-4,2	-0,3 ^{l.s.}
Antal ansatte	0,2	0,6	-0,3 ^{l.s.}	0,2
Arbejdstid	0,8	0,1 ^{l.s.}	1,3	-0,1 ^{l.s.}
Fravær	-2,0	-0,7	2,9	1,2
Fraværstimer	-0,5	-0,6	1,3	0,3
Sygefravær, 2011-2015	-0,4	-0,3	-0,3	-0,4
Barselsfravær, 2011-2015	-1,2	0,3 ^{l.s.}	1,9	1,3
Børn	0,8	1,0	-0,3 ^{l.s.}	-0,3 ^{l.s.}
Enlig	0,5	0,3	0,3	0,2
Region	-0,6	-0,2	-0,3	0,1
Konstant	10,9	7,1	-24,4	-8,8
Antal observationer	1.065.323	170.209	136.561	476.107

Anm.: ^{l.s.} angiver, at bidrag fra karakteristikum ikke er statistisk signifikant på 5-pct.-niveau.

Bilag 5 Resultater for hele arbejdsmarkedet for udvalgte år

Bilagstabel 5.1 Timeløn for alle og særskilt for mænd og kvinder, gennemsnit. Lønforskel, procent. Hele arbejdsmarkedet, 2010, 2013, 2016 og 2017.

	2010 ²	2013 ²	2016	2017
Alle	228,0	237,8	247,7	253,0
Mænd	248,5	258,4	266,5	272,0
Kvinder	210,0	220,2	230,6	235,5
Lønforskel ¹	16,9	16,0	14,5	14,4

Note: ¹ Lønforskellen er beregnet som forskellen mellem mænds og kvinders gennemsnitsløn divideret med gennemsnitslønnen for alle. ² De beregnede lønforskelte for især 2010, men også 2013, afviger fra de tilsvarende lønforskelte i Larsen & Houlberg (2013a) og Larsen & Larsen (2018). Det skyldes, at der anvendes et forskelligt observationsantal, se afsnit 3.1.

FORTROLIG

Bilagstabel 5.2 Bruttolønforskelle og bidrag fra karakteristika særligt for 'forklaret' og 'uforklaret' del, hele arbejdsmarkedet, 2010, 2013, 2016 og 2017. Fase 3. Procent.

	2010	2013	2016	2017
Bruttolønforskelle	13,9	12,7	11,5	11,5
<i>'Forklaret' del, i alt</i>	13,0	11,9	9,7	9,5
Uddannelse	-2,2	-3,1	-3,2	-3,3
Længde	-1,0	-2,0	-2,0	-2,2
Retning	-1,1	-1,0	-1,2	-1,1
Erhvervs erfaring	0,9	1,0	0,7	0,7
Kønsopdelt arbejdsmarked	12,0	12,3	11,8	12,1
Sektor, kvindeandele	1,1	2,7	1,5	1,7
Branche, dummyvariabler	2,3	2,3	3,2	3,4
Jobcelle, kvindeandele	6,5	4,9	4,9	4,9
Jobstatus	2,1	2,4	2,2	2,1
Antal ansatte på arbejdssted	-0,1	-0,4	-0,6	-0,6
Arbejdstid	0,4	0,4	0,0 ^{1.S.}	-0,1
Fravær	2,1	2,0	1,4	1,1
Fraværstimer	0,4	0,5	0,5	0,5
Sygefravær, t-1-t-5	0,3	0,3	0,4	0,3
Barselsfravær, t-1-t-5	1,4	1,1	0,5	0,3
Børn	-0,2	-0,3	-0,4	-0,4
Enlig	0,0	0,0	0,0	0,0
Region	0,0 ^{1.S.}	0,0 ^{1.S.}	-0,1	-0,1
<i>'Uforklaret' del, i alt</i>	0,9	0,9	1,8	2,0
Uddannelse	20,4	19,6	10,8	8,6
Længde	18,8	18,3	10,2	8,1
Retning	1,6	1,3	0,6	0,5
Erhvervs erfaring	-2,0	-0,2 ^{1.S.}	-0,1 ^{1.S.}	0,2
Kønsopdelt arbejdsmarked	-7,5	-13,5	-12,8	-11,4
Sektor, kvindeandele	-2,0	-8,3	-6,8	-5,9
Branche, dummyvariabler	-2,3	-1,2	-1,9	-1,6
Jobcelle, kvindeandele	0,8	1,8	0,6	0,4
Jobstatus	-4,0	-5,8	-4,7	-4,3
Antal ansatte på arbejdssted	-0,1	0,7	1,1	0,7
Arbejdstid	0,9	1,4	1,0	1,0
Fravær	-2,1	-2,2	-1,5	-1,4
Fraværstimer	-0,3	-0,7	-0,5	-0,6
Sygefravær, t-1-t-5	-0,4	-0,4	-0,5	-0,4
Barselsfravær, t-1-t-5	-1,5	-1,0	-0,5	-0,5
Børn	-2,1	0,6	0,8	0,3
Enlig	0,8	0,6	0,5	0,4
Region	-0,1	0,0 ^{1.S.}	-0,1	-0,2
Konstant	-7,1	-6,0	2,1	3,7
Antal observationer	1.722.222	1.691.640	1.848.200	1.864.385

Litteratur

- Albrecht, J., P.A. Edin, M. Sundström & S.B. Vroman (1999). Career interruptions and subsequent earnings: A reexamination using Swedish data, *Journal of Human Resources*, 294-311.
- Albæk, K., F. Casier & M. Larsen (2019). Er kvindefag stadig lavtlønsfag? En analyse af sammenhængen mellem løn og andelen af kvinder i forskellige arbejdsfunktioner. VIVE-rapport.
- Andersen, S. H. (2018). Paternity leave and the motherhood penalty: New causal evidence. *Journal of Marriage and Family*, 80(5): 1125-1143.
- Borchorst, A. (2010). Kollektive aktører og institutionelle spilleregler – barrierer og potentialer for ligeløn. I: M. Deding & H. Holt (red.) *Hvorfor har vi lønforskel mellem kvinder og mænd?* København, SFI Rapport 10:12.
- Card, D.A.R. Cardoso & P. Kline (2016). Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women, *The Quarterly Journal of Economics*, Volume 131, Issue 2, Pages 633–686, <https://doi.org/10.1093/qje/qjv038>
- Collischon, M. (2018). Can personality traits explain glass ceilings?, *SOEPpapers on Multi-disciplinary Panel Data Research*, No. 965, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin.
- Danmarks Statistik (2011). Lønstatistik – metode og nye begreber. TemaPubl 11:3. Tilgængelig på: <https://www.dst.dk/Site/Dst/Udgivelser/GetPubFile.aspx?id=17075&sid=loen>. Besøgt: 03-06-20.
- Danmarks Statistik (2017). Mænd får oftere betaling for overarbejde end kvinder. Nyt fra Danmarks Statistik nr. 234. Tilgængelig på: <https://www.dst.dk/Site/Dst/Udgivelser/nyt/GetPdf.aspx?cid=28414>. Besøgt: 14-08-20.
- Danmarks Statistik (2020a). Lønstatistik for den private, statslige og kommunale sektor. Tilgængelig på: <https://www.dst.dk/da/TilSalg/Forskningservice/Dokumentation/hoekvalitetsvariable/loenstatistik-for-den-private--statslige-og-kommunale-sektor>. Besøgt: 03-06-20.
- Danmarks Statistik (2020b). Statistikdokumentation for Lønstruktur 2018. Tilgængelig på: <https://www.dst.dk/da/Statistik/dokumentation/statistikdokumentation/loenstruktur>. Besøgt: 03-06-20.
- Danmarks Statistik (2020c). SAGSART. Tilgængelig på: <https://www.dst.dk/da/TilSalg/Forskningservice/Dokumentation/hoekvalitetsvariable/dagpenge-ved-sygdom-og-foedsel/sagsart>. Besøgt: 17-08-20.
- Danmarks Statistik (2020d). TIMBET. Tilgængelig på: <https://www.dst.dk/da/Statistik/dokumentation/Times/loenstatistik/timbet>. Besøgt: 17-08-20.
- Danmarks Statistik (2020e). TIMFRA. Tilgængelig på: <https://www.dst.dk/da/Statistik/dokumentation/Times/loenstatistik/timfra>. Besøgt: 17-08-20.

- Datta Gupta, N., R.L. Oaxaca & N. Smith (2006). Swimming Upstream, Floating Downstream: Comparing Women's Relative Wage Progression the United States and Denmark. *ILR Review*, 59 (2): 243-266.
- Deding, M. & K. Wong (2004). Mænds og kvinders løn. En analyse af løngabet 1997-2001. København: Socialforskningsinstituttet 04:10.
- Deding, M. & M. Larsen (2008). Lønforskelle mellem mænd og kvinder 1997-2006. København: SFI 08:28.
- Ferreira, D., E. Ginglinger, M. Laguna & Y. Skalli (2017): Closing the Gap: Gender Quotas and Corporate Board Composition, ECGI Working Paper Series in Finance, Working Paper N° 520/2017.
- Frederiksen, A., T. Halliday & A.K. Koch (2016): "Within- and Cross-Firm Mobility and Earnings Growth". *Industrial and Labor Relations Review*, (69)2, s. 320-353.
- Gallen, Y., R. V Lesner & R.M. Vejlin (2019): The labor market gender gap in Denmark: Sorting out the past 30 years, *Labour Economics*, 56, s. 58-67.
<https://doi.org/10.1016/j.labeco.2018.11.003>.
- Goldin, C. (2014): "A Grand Gender Convergence: Its Last Chapter". *American Economic Review*, (104)4, s. 1091-1119.
- Hirsch, B. (2005): "Why do Part-Time Workers Earn Less? The Role of Worker and Job Skills". *Industrial and Labor Relations Review*, (58)4, s 525-551.
- Holt, H. & M. Larsen (2011). Kønsopdelt lønstatistik og redegørelse om lige løn. Evaluering af loven. København: SFI 11:48.
- Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal*, (8) 4, s. 453-479.
- Joensen, J.S. & H.S. Nielsen (2016): Mathematics and Gender: Heterogeneity in Causes and Consequences, *The Economic Journal*, Volume 126, Issue 593, June 2016, Pages 1129–1163, <https://doi.org/10.1111/eoj.12191>.
- Kleven, H., C. Landais & J.E. Sogaard (2019): Children and Gender Inequality: Evidence from Denmark: *American Economic Journal: Applied Economics*. 11, 4, s. 181-209.
- Larsen, M. (2010). Lønforskelle mellem kvinder og mænd i 2007. Analyser for Lønkommissionen. København: SFI 10:10.
- Larsen, M., H. Holt & M.R. Larsen (2016). Et kønsopdelt arbejdsmarked. Udviklingstræk, konsekvenser og forklaringer. København: SFI 16:15.
- Larsen, M. & H.S.B. Houlberg (2013a). Lønforskelle mellem mænd og kvinder 2007-2011. København: SFI 13:24.
- Larsen, M. & H.S.B. Houlberg (2013b). Mere uddannelse, mere i løn? København: SFI 13:25.
- Larsen, M. & M.R. Larsen (2018). Forskelle mellem kvinders og mænds timeløn. Nye tal for perioden 2012-2016. VIVE-rapport.

- Lesner, R.V. (2019): Labor market sorting on personality traits and the gender wage gap, <https://doi.org/10.1080/13504851.2019.1646867>.
- Nielsen, H.S., M. Simonsen & M. Verner (2004): Does the Gap in Family-friendly Policies Drive the Family Gap?, *The Scandinavian Journal of Economics*, Vol. 106, 721-744.
- Risse, L., L. Farrell & T.R. L. Fry (2018): Personality and pay: do gender gaps in confidence explain gender gaps in wages? *Oxford Economic Papers*, Volume 70, Issue 4, Pages 919–949, <https://doi.org/10.1093/oep/gpy021>
- Rosholm, M. & N. Smith (1996). The Danish gender wage gap in the 1980s: A panel data study. *Oxford Economic Papers*, 48: 254-279.
- Schäfer, K.C. & J. Schwiebert (2018): The impact of personality traits on wage growth and the gender wage gap, *Bulletin of Economic Research*, Volume 70, Issue 1, 20-34, <https://doi.org/10.1111/boer.12115>.
- Simon H., E. Sanroma & R. Ramos (2017): "Full- and Part-Time Wage Differences in Spain: An Analysis along the Wage Distribution". *International Journal of Manpower*, (38)3, s. 449-469.
- Simonsen, M. & L. Skipper (2008): An Empirical Assessment of Effects of Parenthood on Wages. I: D. Millimet, J.A. Smith & E.J. Vytlačil (red.), *Advances in Econometrics*, 21: 359-380.
- Styrelsen for Arbejdsmarked og Rekruttering (2015): Sygedagpengereformen og det nye re- vurderingstidspunkt. Kvantitativ analyse. Tilgængelig på: <https://star.dk/media/1238/sygedagpengereformen-og-det-nye-revurderingstidspunkt.pdf>. Besøgt: 17-08-20.
- Tharp, D. T., M. Lurtz, K. S. Mielitz, M. Kitces, D.A. Ammerman (2019). Examining the gender pay gap among financial planning professionals: A Blinder-Oaxaca decomposition. *Financial Planning Review*, s. 1-20, <https://doi.org/10.1002/cfp2.1061>.
- Theodoropoulos, N., J. Forth & A. Bryson (2019): Are Women Doing It For Themselves? Gender Segregation and the Gender Wage Gap, IZA DP No. 12657.
- Wiswall, M. & B. Zafar (2018): Preference for the Workplace, Investment in Human Capital, and Gender, *The Quarterly Journal of Economics*, 133, s. 457–507, <https://doi.org/10.1093/qje/qjx035>.
- Østbakken, K.M., E. Barth & P. Schøne (2014): Hva betyr kjønnssegregering for lønn? I: Reisel, L. & M. Teigen (red.): *Kjønnsdeling og etniske skiller på arbeidsmarkedet*, s. 205-220. Oslo: Gyldendal Akademisk.

VIDEN
VELFÆRD

DET NATIONALE FORSKNINGS-
OG ANALYSECENTER FOR VELFÆRD