



# KJØNNSFORSKJELLER I LØNNSNIVÅ

og hvilken påvirkning arbeidssituasjon har på denne  
ulikheten

## Innhold

1. Innledning.....	1
1.1. Motiveringen bak valget av problemstilling .....	1
2. Teoretisk rammeverk og tidligere forskning .....	2
2.1. Innledning .....	2
2.2. Teoretisk rammeverk .....	2
2.3. Tidligere forskning og argumentasjon rundt problemstillingen .....	3
2.4. Oppsummering .....	4
3. Datamaterialet.....	5
3.1. Innledning .....	5
3.2. Definisjoner og kilder til datamaterialet .....	5
3.3. Deskriptiv statistikk for avhengig variabel .....	6
3.4. Figurer og videre analyser av avhengig variabel.....	7
3.5. Deskriptiv statistikk for kontrollvariabel .....	8
3.6. Fordeler og ulemper ved datasettet.....	9
3.7. Oppsummering .....	10
4. Økonometrisk modell.....	11
4.1. Innledning .....	11
4.2. Empirisk strategi.....	11
4.3. Valg av funksjonsform .....	14
4.4. Oppsummering.....	16
5. Empiriske resultater .....	17
5.1. Innledning .....	17
5.2. Empiriske hovedresultater.....	17
5.3. Analyse av tilleggsspørsmål.....	25
5.4. Oppsummering.....	26
6. Oppsummering og konklusjon.....	27
7. Referanser.....	28
8. Appendiks.....	29

## 1. Innledning

### 1.1. Motiveringen bak valget av problemstilling

I 2008 oppnevnte regjeringen Stoltenberg en kommisjon for å utrede en undersøkelse om forskjeller i lønn mellom kvinner og menn, og hva som kunne være de bakenforliggende årsakene til denne vesentlige kjønnsdiskrimineringen (NOU 2008: 6). Utredningen kom frem til at det forekommer substansielle forskjeller i lønnsnivå mellom kjønnene. Siden den gang har kjønnsforskjellene blitt noe redusert, men de er fremdeles tilstedeværende i det norske samfunn (Gunnes, 2019). Ytterligere er likestilling av kjønnene i aller høyeste grad en relevant problemstilling i dagens samfunn, og det er en samfunnsøkonomisk utfordring som debatteres hyppig i offentlige kanaler (Ibid).

Valget av bacheloroppgavens problemstilling motiveres derfor av ønsket om å avdekke en ytterligere forklaring bak lønnsforskjeller basert på kjønn. Bacheloroppgavens problemstillinger som skal analyseres og besvares er som følger:

- Er det kjønnsforskjeller i lønn og lønnsavkastning av utdanning og arbeidserfaring?
- Er det en relasjon mellom kjønnsforskjeller i lønn og partners arbeidssituasjon i Norge?

I tillegg vil oppgaven besvare dette tilleggsspørsmålet ved problemstillingen:

- Opplever kvinner ulike kjønnsforskjeller i lønn mellom offentlig og privat sektor?

Oppgaven vil ta for seg en analyse av disse elementene gjennom en empirisk statistikkundersøkelse som ligger til grunn for oppgavens problemstilling. Først vil det bli foretatt en gjennomgang av teoretisk rammeverk og tidligere forskning. Deretter blir det en gjennomgang av datamaterialet oppgaven skal undersøke. Det vil så foretas en gjennomgang av den økonometriske modellen analysen skal benytte seg av.

Avslutningsvis skal de empiriske resultatene analyseres og diskuteres, frem til at oppgaven besvares i en oppsummering og konklusjon.

Oppgaven konkluderes i at det forekommer ingen kjønnsforskjeller i lønn av utdanning og arbeidserfaring, og heller ingen kjønnsforskjeller i lønnsavkastning av utdanning. Det forekom en svært liten, men statistisk signifikant, forskjell i avkastning av arbeidserfaring, hvor kvinner har lavere avkastning enn menn. Det forekom ingen betydelig signifikant sammenheng mellom lønn og partners arbeidssituasjon for fulltidsansatte, arbeidsledige og partnere i andre arbeidssituasjoner. Derimot resulterte analysen i at det forekom en

statistisk signifikant sammenheng mellom lønn og deltidsansatte partnere, hvor menns lønn ble påvirket positivt av å ha en deltidsarbeidende partner, og hvor kvinners lønnsnivå ble påvirket negativt av å ha en deltidsarbeidende partner. I tillegg forekommer det en større grad av kjønnsforskjeller i lønn i privat sektor enn offentlig sektor. Det viste seg at lønnsnivåene er også høyere i privat sektor. At arbeidende kvinner da er konsentrert rundt deltidsarbeid i offentlig sektor konkluderes derfor til en stor bidragsyter til at det forekommer lønnsforskjeller mellom kvinner og menn, i tråd med tidligere forskning.

## 2. Teoretisk rammeverk og tidligere forskning

### 2.1. Innledning

Dette kapitlet vil gjennomgå hvilke teorier oppgaven tar for seg for å besvare problemstillingen, samt litteraturer og tidligere forskning om lignende problemstillinger som drøfter årsaken til kjønnsforskjeller i lønn.

### 2.2. Teoretisk rammeverk

Årsakene til at det observeres en kjønnsforskjell i lønnsnivå kan være flerfoldig. Ifølge en rapport fra Statistisk sentralbyrå er arbeidstid, særlig deltidsarbeid blant kvinner, et vesentlig bidrag til at kvinner opplever lavere lønnsnivå enn menn (Gunnes, 2019). En annen kjent teori av den amerikanske økonomen Gary Becker er teorien om humankapital. Denne teorien tilsier at en økning i utdanningsnivå skal ytterligere øke et individs lønnsnivå (Becker, 1964). Noe som er interessant med denne teorien er at kjønn i utgangspunktet ikke skal ha noen påvirkning på hvor mye inntekten vil øke for økt utdanning. Dette betyr at ifølge Beckers teori skal en kvinne ha like god avkastning på utdanning som menn. Denne påstanden er skal videre analyseres senere i oppgaven.

En annen teori om årsaker til kjønnsforskjeller i lønn er den av Begg et al. om lønnsdiskriminering. Kjønnsdiskriminering forekommer hvor tilsvarende produktive arbeidere utbetales ulik som har årsak i ulik personlig karakteristikk, som alder, rase eller kjønn (Begg, 2014; s.243). Her skiller Begg et al. mellom smaksbasert diskriminering, som er lønnsdiskriminering basert på preferanser, og statistisk diskriminering. Sistnevnte er lønnsdiskriminering basert på statistikk over hvilke arbeidsgrupper som er mer produktive enn andre. Eksempelvis kan en arbeidsgiver som mener at kvinner er

gjennomsnittlig dyktigere regnskapsførere enn menn, gi mer lønn til kvinnelige regnskapsførere enn mannlige.

Beggs videre teorier om kjønnsforskjeller mellom kvinner og menn går ut på at lønnsforskjeller kan oppstå fordi at særlig unge kvinner har en tendens til å avbryte karrieren for å få barn (Begg, 2014; s.244). Dette avbruddet hos kvinnelige ansatte gjør at investeringer bedriften gjør i form av kompetanse og opplæring av ansatte har bedre avkastning hos menn enn kvinner. Årsaken til lønnsgapet, mener Begg, er derfor ikke grunnet i systematisk lønnsdiskriminering, men heller at kvinner er mer utsatt for permisjon, som fører til at bedrifter investerer flere ressurser i mannlige ansatte. Resultatet gir de mannlige ansatte relativt høyere lønn enn kvinnene i samme bedrift (Ibid).

### 2.3. Tidligere forskning og argumentasjon rundt problemstillingen

Av tidligere forskning på temaet trekkes den sjette norske offentlige utredningen fra 2008 frem, med tittelen «Kjønn og lønn – Fakta, analyser og virkemidler for likelønn». Utredningen ble etterspurt for å analysere de bakenforliggende årsakene til at det norske samfunn opplevde lønnsforskjeller mellom kvinner og menn. Av de mest relevante funnene for denne oppgaven trekkes det nå frem de konklusjonene som kan svare på vår problemstilling.

Utredningen kom frem til at lønnsforskjeller oppstår ikke på grunn av direkte diskriminering av kvinner, men i opphav av at kvinner og menn har ulike stillinger (NOU 2008:6, s.54). Ytterligere viste det seg at lønnsnivået er gjennomsnittlig lavere i kvinnedominerte yrker enn i mannsdominerte yrkesområder. Flere kvinner enn menn tar i dag høyere utdanning, men det er fremdeles en fremtredende lønnsforskjell mellom kjønnene med likt utdanningsnivå, særlig av høyere utdanning (NOU 2008:6, s.58).

Omtrent 50% av lønnsforskjeller mellom kjønnene kan forklares med at kvinner jobber deltid. Forskning viste at heltid- og deltidstimplønnen er lite forskjellig, alt annet likt, og kunne ikke hevde at de som jobber deltid blir utsatt for timelønnsdiskriminering.

Kjønnsforskjeller kommer snarere frem i at deltidsansatte har færre arbeidstimer enn fulltidsansatte, og at dette har bidratt til en netto lønnsforskjell mellom kvinner og menn. Det viste seg at deltidsarbeid konsentrerte seg i kvinnedominerte yrkesgrupper, og at deltidsarbeid dermed kan ha bidratt til å videreutvikle et gjennomsnittlig lavere lønnsnivå hos kvinner enn hva som forekommer blant menn (NOU 2008:6, s.72).

Deretter viste det seg at langt flere kvinner jobber i offentlig sektor, særlig omsorgsyrkene. Det kom også frem at lavutdannede i offentlig sektor tjente bedre enn i privat sektor, mens høyutdannede i offentlig sektor hadde et lavere lønnsnivå enn høyutdannede i privat sektor. Ytterligere er det en høy konsentrasjon av kvinner i de mest lavtlønnede private yrkesområdene (NOU 2008:6, s.64). Utredningen konkluderer i at lønnsforskjeller mellom kjønnene oppstår i en ulik konsentrasjon av kjønnene i stillinger, bransjer og sektorer med forskjellig lønnsnivå (NOU 2008:6, s.75). Lønnsforskjeller mellom kjønnene er dog mindre i Norge enn i de aller fleste andre land. Trolig pga. en mer egalitær lønnsstruktur i samfunnet vårt enn hvordan lønnsstrukturen forekommer i andre samfunn (Ibid).

I rapporten fra SSB vises det til at kjønnsforskjeller i lønn er blitt gradvis redusert siden utredningen i 2008. Kvinners gjennomsnittlige månedslønn utgjorde 86,7% av menns gjennomsnittlige månedslønn i 2017, mens tilsvarende i 2008 var 85,0%. Dog viste rapporten som nevnt at en særlig påvirkning til lønnsforskjellene er en høy konsentrasjon av kvinner i deltidsstillinger. Nesten 1 av 2 kvinner (48,4% av alle kvinner) jobbet i deltidsstillinger i 2018, hvor bare 26,4% av alle menn gjorde det samme.

Kjønnsforskjeller i lønnsnivå mellom kvinner og menn var nærmest ikke-eksisterende i deltidsstillinger, mens kvinner som jobbet i heltidsstillinger i gjennomsnitt fikk kun 88,5% av hva en tilsvarende mannlig heltidsarbeider gjennomsnittlig fikk utbetalt i månedslønn.

#### 2.4. Oppsummering

Dette kapittelet har drøftet det teoretiske rammeverket oppgaven skal innebefatte. Dette innebærer teorier om hvorfor kjønnsforskjeller i lønn kan oppstå. Disse teoriene inkluderer kvinners arbeidssituasjon, humankapitalinvestering, og Begg et al. definisjoner på lønnsdiskriminering. Videre er det foretatt en presentasjon av tidligere forskning som er relevant for oppgavens helhetlige mål, som videre analyse skal bygges på, og som resultater og funn vil sammenlignes og drøftes mot.

### 3. Datamaterialet

#### 3.1. Innledning

I dette kapitlet vil det presenteres hvilket datamateriale som benyttes for å besvare oppgaven. Innunder dette inngår definisjoner og kilder av datamaterialet, samt deskriptiv statistikk og analyser av ulike variabler som oppgaven tar høyde for.

#### 3.2. Definisjoner og kilder til datamaterialet

Variablene fra datasettet som skal benyttes i oppgaven er følgende:

earnhr, female, yrsqual, j\_q02c, og d\_q03

Disse variablene vil nå ytterligere forklares i den rekkefølgen som er angitt ovenfor.

Variabelen earnhr defineres som timelønn i lokal valuta, og vil bli omtalt som timelønn videre i oppgaven. Variabelen female er en kjønnsvariabel i form av en såkalt dummyvariabel, hvor variabelen er lik 1 dersom respondenten til statistikken er en kvinne, og er lik 0 dersom respondenten er en mann. Denne variabelen er en sentral forklaringsvariabel i oppgaven. Videre defineres variabelen yrsqual som antall år med utdanning.

Variabelen j\_q02c defineres som arbeidssituasjonen til respondentens samboer eller partner. Denne variabelen defineres ytterligere inn i disse dummyvariablene som skal analyseres videre i oppgaven: fulltid hvis respondentens partner jobber i en fulltidsstilling, deltid hvis partner jobber deltid, arbeidsledig dersom partner ikke driver med inntektsgivende arbeid, og til slutt en dummyvariabel for annen arbeidssituasjon, dersom partneren er i en annen uspesifisert jobbsituasjon. Dette innebærer om observasjonens partner er elev, student, lærling, pensjonist, funksjonsnedsett/uføretrygget, i militærtjeneste eller hjemmeværende. Disse fire variablene for arbeidssituasjonen til observasjonens partner utnevnes som de mest interessante variablene å undersøke og ta stilling til i forhold til oppgavens problemstilling. Grunnen til at denne variabelen benyttes kontra variabelen for fulltidsarbeid (som kun viser om en observasjon jobber fulltid eller ikke) er fordi disse variablene kan brukes til å vise ytterligere nyanser ved arbeidssituasjonen hos kjønnene enn hva den enkle fulltidsvariabelen gjør. I tillegg er det interessant å undersøke hvordan, og særlig hvorfor, en påvirkes av partners arbeidssituasjon.

Videre benyttes også variabelen d\_q03, som definerer hvilken økonomisk sektor respondenten jobber i. Denne variabelen deles også videre inn i disse to dummyvariablene som skal benyttes i regresjonsanalysen: offentlig sektor dersom respondenten jobber i en offentlig stilling, og privat sektor, dersom respondenten jobber i det private.

### 3.3. Deskriptiv statistikk for avhengig variabel

	Alle	Kvinner	Menn
Gjennomsnittlig timelønn	214.2932	197.8306	230.2995
Minimum	50	50	50
Maksimum	572.6316	572.6316	571.4286
Standardavvik	78.78021	66.08572	86.4741
Antall observasjoner	3558	1754	1804

Tabell 1

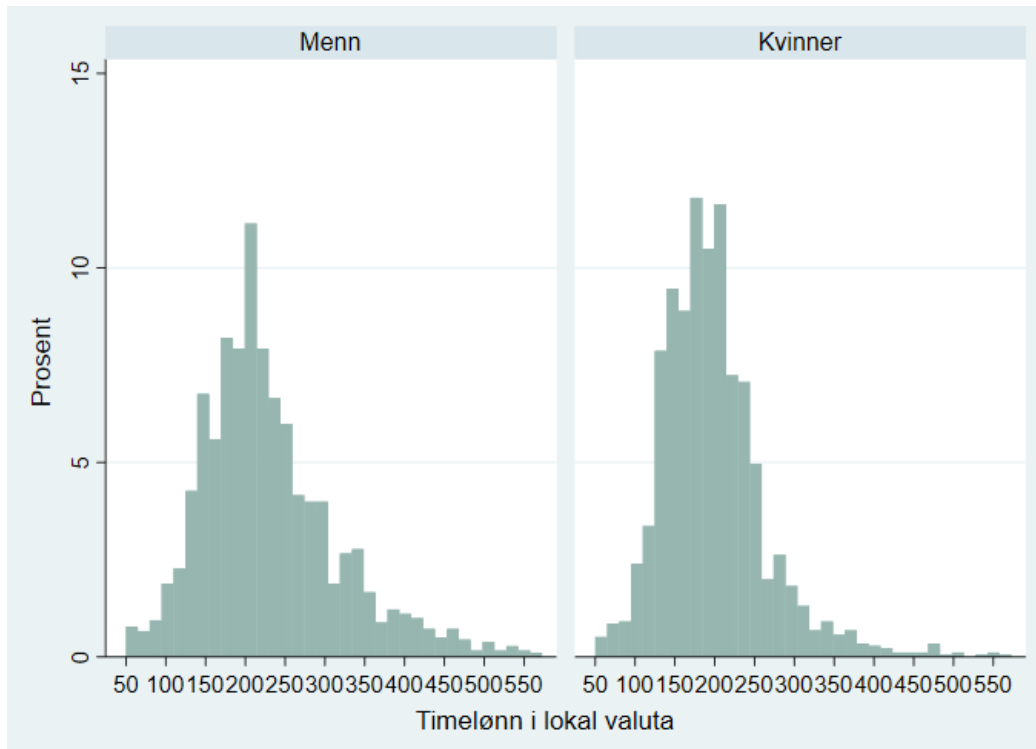
Tabellen ovenfor er hentet fra datasettet og viser til en fundamental forskjell i gjennomsnittlig lønnsnivå mellom kvinner og menn isolert sett. Gjennomsnittlig timelønn blant kvinner er kun 86% av hva som er det gjennomsnittlige lønnsnivået for menn. Derimot ser vi at både minimum- og maksimumsverdiene for lønn er like blant kjønnene, noe som vil tilsa at det må være en større konsentrasjon av lavtlønnede kvinner enn menn. Dog er standardavviket høyere blant menn enn hos kvinner, som betyr at lønnsnivået er mer spredt hos menn enn hos kvinner. Dette kan bekrefte hvorfor det er en konsentrasjon av lavtlønnede kvinner i utvalget.

Hvorfor gjennomsnittlig inntektsnivå ligger så mye høyere for menn enn kvinner er hva oppgaven skal nærmere undersøker. Uten videre innsikt kan en anta at årsakene, utenom systematisk lavere lønn for kvinner uansett arbeid, ligger i ulike egenskaper ved gjennomsnittlig arbeid hos hvert kjønn. Årsaken til forskjellen må altså befinne seg i at menn jobber i en mer høytlønnnet sektor, en mer høytlønnnet arbeidssituasjon, eller et mer høytlønnnet yrke enn hva som gjelder det tilsvarende for kvinner. Det blir fort klart at kjønnsforskjeller i lønn ikke lar seg forklare av lønnsvariabelen alene, og at vi må introdusere flere variabler for å få en nyansert modell for finne svar på problemstillingen.



### 3.4. Figurer og videre analyser av avhengig variabel

Konsentrasjonen av kvinner med lavere lønn kommer ytterligere frem i histogrammet nedenfor som viser den prosentvise lønnsfordelingen hos hvert kjønn.



Figur 1

Her ser vi at det er en særlig høyere prosentvis konsentrasjon i timelønn fra 100-250 kr hos kvinner. Denne konsentreringen er langt større relativt til det tilsvarende området hos menn. Videre illustreres den høyere spredningen i lønnsnivå hos menn ved at histogrammene er mer fordelt utover timelønnsnivåene enn hva det gjør hos kvinner. Det er også svært små utslag i lønnsnivå fra 400 kr i timen og oppover hos kvinner sammenlignet med menn. Dette indikerer at det er flere menn med høyere lønnsnivå enn andre menn, enn hva det er kvinner med høyere lønnsnivå enn andre kvinner. Dette innebærer også at det er langt flere menn med høyere lønnsnivå enn kvinner, som også kommer til uttrykk i kjønnsforskjellene i gjennomsnittlig timelønn på forrige side. Årsaken til hvorfor dette fenomenet forekommer er hva oppgaven skal analysere videre for å finne en forklaring på.

### 3.5. Deskriptiv statistikk for kontrollvariabel

	Alle			Kvinner			Menn		
	Gj.snitt	Std	Obs	Gj.snitt	Std	Obs	Gj.snitt	Std	Obs
Kvinne	0.4822543	0.4997337	5128	1	0	2473	0	0	2655
Antall år utdanning	14.14118	2.596513	4951	14.1569	2.594078	2390	14.12651	2.599204	2561
Antall år arbeidserfaring	18.60284	12.73899	4716	17.12857	11.90103	2279	19.98153	13.33123	2437
Offentlig sektor	0.3578655	0.4794329	3954	0.488051	0.49999	1883	0.2394978	0.4268801	2071
Privat sektor	0.6264542	0.48308063	3954	0.4949549	0.5001074	1883	0.7460164	0.4353935	2071
Partner jobber fulltid	0.6640203	0.4724068	3149	0.7823219	0.4128035	1516	0.5541947	0.4972065	1633
Partner jobber deltid	0.1365513	0.3434276	3149	0.0530897	0.2262693	1516	0.2131047	0.4096265	1633
Arbeidsledig partner	0.0200064	0.140044	3149	0.0151715	0.1222751	1516	0.0244948	0.1546268	1633
Partner i annen arbeidssituasjon	0.179422	0.3837663	3149	0.1484169	0.3556301	1516	0.2082058	0.4061492	1633

Tabell 2

Tabellen ovenfor viser den deskriptive statistikken for kontrollvariablene som analysen baseres på. Som vi ser i tabellen er 48% av observasjonene er kvinner, og dermed er 52% av observasjonene menn. Dette kommer til uttrykk på antall observasjoner av hvert kjønn. Antall år utdanning varierer mellom 7 til 21 hos begge kjønnene, og gjennomsnittet mellom begge kjønn ligger på cirka 14 år. Antall års arbeidserfaring varierer fra 0 til 47 år hos begge kjønn. Her kommer det frem at kvinner ligger noe under gjennomsnittet for begge kjønnene, mens menn ligger noe over.

Når det kommer til økonomisk arbeidssektor observeres det et tydelig skille mellom kjønnene. Mellom begge kjønn jobber i gjennomsnitt 36% i offentlig sektor.

Gjennomsnittlig jobber 48% av alle kvinner i offentlig sektor, mens bare 24% av alle menn også befinner seg i denne sektoren. I privat sektor jobber 63% av arbeidsstyrken. Bortimot 75% av alle menn jobber i privat sektor, mens kun omtrent 50% av alle kvinner gjør det samme. En tolkning av dette er at det er mulig at flere yrker i offentlig sektor, som f.eks. sykepleiere og lærere, er særlig kvinnedominerte.

Observasjonene hos partners arbeidssituasjon viser til at 78% av kvinners partnere jobber i en fulltidsstilling, mens bare 55% av menns partnere gjør det samme. Dette gjenspeiles i

observasjonene om at bare 5% av kvinners partnere jobber deltid, mens over 21% av menns partnere har deltidsstillinger. Begge kjønn har en forholdsvis lav og lik prosentvis andel partnere som er arbeidsledig, med et kjønnsveid gjennomsnitt på 2%. Vi ser også at flere menn enn kvinner har partnere i en annen arbeidssituasjon, med at omtrent 15% kvinners partnere er i en slik situasjon, mens nesten 21% av menns partnere er det samme.

En tolkning av arbeidssituasjonen kan ha sammenheng med kjønnsfordelingen i økonomiske sektorer, hvor f.eks. sykepleiere, en særlig kvinne-dominert offentlig stilling, er i høy grad en deltidsstilling.

### 3.6. Fordeler og ulemper ved datasettet

En av fordelene til datasettet er at det tilbyr tilstrekkelige varierte variabler slik at vi kan skille særegne individer fra hverandre, og isolere de i egne grupper som videre kan analyseres og sammenlignes mot hverandre. Vi kan derfor isolere datamaterialet om hvorvidt det er snakk om kvinner eller menn, om individet har 4 års arbeidserfaringen eller 40, hvor lang gjennomsnittlig utdanning denne gruppen har, hvilken økonomisk sektor de tilhører, og så videre. I tillegg tilbyr datasettet de relevante variablene oppgaven ønsker å undersøke gjennom dens analyse, slik som timelønn, erfaring, utdanning, og mer.

Svakhetene til datasettet kommer frem i blant annet hvor generell og uspesifisert statistikken kan være. I forhold til arbeidserfaring sier datamaterialet ingenting om hva slags type erfaring dette innebærer. F.eks. kan arbeidserfaring fra to forskjellige sektorer ha liten betydning for lønn i hver sektor, dersom arbeidet er fullstendig urelatert til hverandre. Her blir vi derfor nødt til å anta at arbeidserfaringen har en uavhengig positiv effekt på lønnsnivå. Alder hadde også vært nyttig å spesifisert, ettersom arbeidsmarkedet i dag og for 30 år siden trolig skiller seg fra hverandre på flere måter. Antall år utdanning spesifiserer heller ikke omfanget av utdanningen gjennomført. Eksempelvis kan to ulike individer, en med fullført mastergrad og en med fem årstudier, ha like mange års utdanning å vise til. Videre kan hvilken type utdanning gjennomført ha et utslag på lønn. Hvis en antar at en lege har høyere timelønn enn en litteraturviter, kan en også anta at fem års legeutdanning har høyere lønnsavkastning enn fem år med litteraturstudier. En annen antagelse analysen må foreta blir derfor at all utdanning blir fullført til normert tid, at en formell utdanning blir foretatt, og at all utdanning har lik effekt på timelønn, uavhengig av studieretning eller karakterresultater på nevnt utdanning.

I forhold til variabelen for partners arbeidssituasjon hadde det også vært nyttig med mer spesifisert informasjon. Det er f.eks. ingen betinget nødvendighet at individets partner er av motsatt kjønn, som vil ha mye å si for hvilke kjønnsdiskrimineringer partneren står ovenfor angående timelønn. Eventuelt hadde det også vært interessant å undersøkt om homofile par stod ovenfor lønnsdiskriminering, og eventuelle årsaker til dette. Derimot er seksuell legning ikke oppgitt i datasettet, som gjør en slik analyse umulig.

Videre hadde det vært ideelt om flere observasjoner rundt partners arbeidssituasjon hadde blitt foretatt i datasettet for å gi en mer representativ analyse. Eksempelvis har omtrent 12% av 1516 undersøkte kvinner, som tilsvarer omtrent 182 kvinner, en arbeidsledig partner. Å analysere denne statistikken og redegjøre økonometriske modeller på dette grunnlaget blir i svært liten grad tilfredsstillende representativt. Å utføre statistisk signifikante konklusjoner om det virkelige samfunnsbildet basert på et så lite utvalg kan avvike fra den faktiske sannheten. Dog vil modellen likevel tilby en representasjon av utvalget som kan bære frem viktige funn om hvorfor det oppstår kjønnsforskjeller.

På bakgrunn av dette gjøres det en antagelse om at observasjonene foretas av heterofile forhold, for å bringe innsikt til hvorvidt de ulike kjønnene befinner seg i ulike arbeidssituasjoner. Spesifikt fulltidsarbeid, deltidsarbeid, eller arbeidsledighet, presenteres ikke i noen annen variabel i datasettet. Ulikhetene i arbeidssituasjon kan forklare hvorfor det oppstår lønnsforskjeller mellom kjønnene, og kan derfor vise seg å være en særdeles nyttig variabel til tross for at utvalget kunne vært noe mer omfattende i størrelse.

### 3.7. Oppsummering

I dette kapitlet er det blitt foretatt en gjennomgang og en definering av datasettets variabler, etterfulgt av deskriptiv statistikk for avhengig variabel, samt en figur for denne. Deretter ble det utført en tilsvarende deskriptiv statistikk for kontrollvariablene, en kort tolkning av enkelte observasjoner av relevant statistikk, samt fordelene og ulempene datasettet bringer med seg til den videre analysen av oppgaven.

#### 4. Økonometrisk modell

##### 4.1. Innledning

I dette kapittelet vil det foretas en presentasjon og drøfting av den økonometriske modellen oppgaven skal analysere. Dette vil innebære den empiriske strategien besvarelsen benytter, og en definisjon av modellens funksjonsform.

##### 4.2. Empirisk strategi

Analysens regresjonsmodell vil være minste kvadraters metode (OLS – ordinary least square). OLS er en lineær regresjonsmetode en benytter for å forklare forholdet mellom en avhengig variabel og uavhengige kontrollvariabel(er) ved å estimere de ukjente parameterne som kreves for å sette variablene i en lineær modell (Thomas, s.266). OLS estimerer parameterne ved å minimisere residualkvadratsummen, altså summen av kvadratdifferansene mellom den observerte avhengige variabelen. For å forklare nærmere ser vi for oss følgende regresjonsligning:

$$(1) Y_i = \alpha + \beta X_i + \varepsilon_i$$

Her er  $Y_i$  den uavhengige variabelen,  $\alpha$  er modellens konstantledd,  $\beta$  er en ukjent parameter for observasjonen  $X_i$ .  $\varepsilon_i$  er et stokastisk restledd som fanger opp påvirkningen på  $Y_i$  av uforklarte faktorer i regresjonen.  $i = 1, 2, \dots, n$  for antall observasjoner.

Ligningen ovenfor kan være regresjonen for en fullstendig populasjon, men siden dette ikke tillater seg å effektivt måle, benytter vi OLS for å estimere en tilsvarende lineær regresjonsmodell for utvalget. Avviket fra den faktiske uavhengige variabelen  $Y_i$  og den observerte uavhengige variabelen i utvalget  $\hat{Y}_i$  defineres som residualen  $e_i$ . Altså:

$$(2) e_i = Y_i - \hat{Y}_i$$

$e_i$  kan også ansees som det empiriske motstykket til  $\varepsilon_i$ . Hvis vi videre sier at  $a$  er en estimator for  $\alpha$ , og at  $b$  er en estimator for  $\beta$ , kan vi skrive om likning (1) på følgende vis:

$$(3) \hat{Y}_i = a + bX_i + e_i$$

For å finne en regresjonsmodell ved minste kvadraters metode ønsker vi å estimere  $a$  og  $b$  på en slik måte at residualkvadratsummen, altså  $e_i^2$ , blir minst mulig (Ibid).

Residualkvadratsummen kan også uttrykkes ved SSR. En finner minimal SSR for å finne en estimert  $\hat{Y}_i$  som er likest mulig den faktiske  $Y_i$ , slik at resultatet blir så korrekt som

mulig. Grunnen til at den kvadreres er for å ta høyde for at  $\hat{Y}_i$  kan være både mindre og større enn  $Y_i$ . Etter omregning på likning (3) kommer vi frem til følgende formler for å finne de optimale estimatorene for  $\alpha$  og  $\beta$  er:

$$(4) a = \bar{Y} - b\bar{X}$$

$$(5) b = \frac{\sum(Y_i - \bar{Y})(X_i - \bar{X})}{\sum(X_i - \bar{X})^2}$$

Hvor  $\bar{Y}$  og  $\bar{X}$  er gjennomsnittsverdier for  $Y_i$  og  $X_i$ . En gjenstår da med en empirisk lineær regresjonsmodell som er likest mulig det faktiske statistiske forholdet en uavhengig variabel  $Y_i$  står ovenfor en populasjon. Men dette krever at følgende forutsetninger om det stokastiske restleddet  $\varepsilon_i$  er oppfylt: (Thomas, s.359 og s.394)

1. Hver verdi  $X_i$  kan ikke være stokastisk.
2. Forventningsverdien av restleddet  $E(\varepsilon_i) = 0$ . Dette antyder at alle gjennomsnittlige forstyrrelser for hver  $X_i$  er lik, slik at tilsvarende  $Y_i$  antas å være likt plassert under og over regresjonslinjen for hver instans av  $i$ .
3. Variansen av restleddet  $Var(\varepsilon_i) = \sigma^2$ , altså den har en konstant varians. Slik blir spredningen for hver  $Y_i$  lik for hver instans av  $i$ . Restleddet kan med dette anses som homoskedastisk, altså lik residualspredding for alle  $X_i$ .
4. Kovariansen mellom to vilkårlige restledd  $i$  og  $j$   $Cov(\varepsilon_i, \varepsilon_j) = 0$ . Dette forekommer av at det ikke kan være to observasjoner,  $X_i$  og  $X_j$ , som samvarierer siden de er fra et tilfeldig utvalg. Derfor innebærer denne forutsetningen også at  $Cov(X_i, X_j) = 0$ .
5. Hver  $\varepsilon_i$  må være normalfordelt. Hvis likning (1) er konstant og  $\varepsilon_i$  er normalfordelt, er også  $Y_i$  normalfordelt over gjennomsnittet av konstanten  $E(Y_i) = \alpha + \beta X_i$ . Denne forutsetningen er en nødvendighet for hypotesetesting, som bør være av normalfordelt variabler.

Dersom forutsetning 1 ikke er oppfylt, gjelder forutsetning 2-5 betinget på realiserte verdier av  $X_i$ . Hvis forutsetning 5 ikke gjelder vil ikke estimatorene vi får for parameterne gjennom OLS-metoden være forventningsrette og effesiente. En kan da ikke anse estimatorene som BLUE (Best Lineær Unbiased Estimator)(Ibid). Modellen ansees da å inneholde heteroskedastisitet, altså variansen for residualen er ulik over alle verdier av  $X_i$  (Thomas, s. 479).

For å forsikre oss om at modellen oppfyller disse forutsetningene vil alle regresjonene som foretas i oppgaven kjøres gjennom en robusthetstest. Testen vi benytter oss av kalles Breusch-Pagan testen, og den undersøker om residualverdien avviker fra et lineært forhold av de eksogene variablene. Breusch-Pagan testen er gitt ved:

$$V(\varepsilon_i) = f(\alpha_1 + \alpha_2 W_2 + \alpha_3 W_3 + \dots + \alpha_m W_m)$$

Hvor  $W$  er variabler som kan påvirke variansen av restleddet (Thomas, s.480). Hvis det forekommer heteroskedastisitet i modellen vil modellen som sagt kjøres med robuste standardfeil, som vil si at regresjonen tilpasser seg slik at den blir så lite heteroskedastisk som mulig.

For å anslå hvor mye av modellen som forklares av variablene regresjonen består av, benyttes determinasjonskoeffisienten  $R^2$ . Determinasjonskoeffisienten defineres av den forklarte variasjonen i den avhengige variabelen som forklares av modellen ( $SSE = \sum(\hat{Y}_i - \bar{Y})^2$ ) fordelt på totalt variasjon i modellen ( $SST = \sum(Y_i - \bar{Y})^2$ ). Dette avviket i variasjon kan ytterligere illustreres ved avviket mellom

$$(6) R^2 = \frac{\text{forklart variasjon i modellen}}{\text{total variasjon i modellen}} = \frac{SSE}{SST} = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2}$$

Videre kan den totale variasjonen i modellen ( $SST$ ) uttrykkes gjennom residualkvadratsummen ( $SSR$ ). Det var denne variasjonen vi ønsket å minimere ved hjelp av likning (4) og (5) for å finne en regresjonsmodell som forklarte mest mulig av variasjonen. Dette fordi det er denne variasjonen som gjenstår uforklart i modellen, som vi konsekvent ønsker minst av. Med denne kan totalvariasjonen uttrykkes ved:

$$(7) SST = SSE + SSR$$

Bruker vi dette til å flytte om på likning (7) kan vi også uttrykke determinasjonskoeffisienten ved hjelp av residualkvadratsummen:

$$(8) R^2 = \frac{SST - SSR}{SST} = 1 - \frac{SSR}{SST}$$

Innen statistikk benyttes hypotesetesting for å kontrollere signifikansen rundt enkelte parametere, for eksempel om tilstedeværelsen av en parameter påvirker regresjonen av en modell og er av statistisk betydning (Thomas, s.126). Dette kan formuleres i hypoteser, hvor nullhypotesen  $H_0$  antyder at variabelen  $\beta_1$  ikke er av betydning ( $\beta_1 = 0$ ), som skal

testes mot alternativhypotesen  $H_1$  at variabelen  $\beta_1$  har en betydning ( $\beta_1 \neq 0$ ). I alternativhypotesen kan  $\beta_1$  både være mindre og større enn null. Vi sier da at vi har et tosidig alternativ, og hypotesen må testes mot både positive og negative verdier av  $\beta_1$  (Ibid). Signifikansnivået er graden av sannsynlighet for å feilaktig forkaste nullhypotesene (Thomas, s.129). Dette begrepet benyttes i oppgaven for å gi en indikator på korrektheten til et gitt estimat, og det er vanlig å benytte 5%-signifikansnivå. Med et 5%-signifikansnivå har vi et  $(100-5) = 95\%$  konfidensintervall. Dette tilsier at det er 95% sannsynlighet for at faktisk regresjonsverdi er innenfor estimatet (Thomas, s.114).

Videre benyttes signifikansnivået til å estimere kritisk verdi, eller p-verdi, ved en t-test. Den kritiske verdien er en verdi gitt av signifikansnivået  $\alpha$ , antall observasjoner  $n$ , og antall koeffisienter  $i$ , og uttrykkes som  $t_{\frac{\alpha}{2}, n-i}$  (Thomas, s.171). Basert på estimert verdi for variabelen  $\beta_1$ , som i henhold til likning (3) er  $b_1$ , og estimert standardavvik for den samme variabelen,  $S_{b_1}$ , får vi testobservatoren (TS), som uttrykkes  $\frac{b_1}{S_{b_1}} \sim t_{n-i}$ . Benevnningen  $\sim t_{n-i}$  antyder at TS t-fordeles med  $n$  antall frihetsgrader fratrukket antall koeffisienter  $i$ . Dersom absoluttverdien av TS overstiger kritisk verdi ved t-fordelingen kan vi nemlig forkaste nullhypotesen og godta alternativhypotesen til angitt signifikansnivå.

#### 4.3. Valg av funksjonsform

Vi skal nå foreta oss et valg av funksjonsformen til regresjonsmodellen, hvor vi ønsker å finne en modell med høyest mulig determinasjonskoeffisient. La oss begynne med en funksjonsform med de relevante variablene som ligner på likning (1):

$$(9) \text{timelønn}_i = \alpha + \beta_1 \text{female}_i + \beta_2 X_i + \varepsilon_i$$

Vi har nå en likning hvor avhengig variabel er  $\text{timelønn}_i$ ,  $\text{female}_i$  er kontrollvariabelen som representerer kjønn, og  $X_i$  er andre kontrollvariabler vi ønsker å undersøke senere, som utdanning, arbeidserfaring, etc. I tillegg består modellen av et stokastisk restledd, angitt som  $\varepsilon_i$ . Ved hjelp av OLS-estimatoren kan vi måle determinasjonskoeffisienten ved denne funksjonsformen, en såkalt lineær-lineær form. Ytterligere kan vi teste om funksjonsformen forklarer mer av sammenhengen i variablene dersom vi benytter oss av en transformasjonsteknikk og gjør begge sidene av likning (9) logaritmiske, i stedet for lineære. Dette gjør vi ved å ta logaritmen av begge sider, og så teste forklaringseffekten av denne versjonen. Videre kan vi også teste om funksjonsformen forklarer enda mer av



variabelforholdet dersom den er dellogaritmisk, altså at bare den ene siden av likningen er logaritmisk, mens den andre er lineær.

Et viktig poeng i sammenligningen av ulike funksjonsformer er at det kun er relevant å sammenligne forklaringseffekten hos funksjonsformer som har lik avhengig variabel. En logaritmisk funksjon av timelønnen lar seg ikke sammenligne med en lineær funksjon av samme variable når det kommer til forklaringskraft. Etter omfattende testing hvor dette poenget er tatt høyde for kom det frem til at en logaritmisk-lineær funksjonsform var den best passende funksjonsformen til å besvare oppgaven etter hvert som likningen utvides. Ytterligere begrunnelse for denne funksjonsformen er at det er en form som ofte benyttes i tidligere studier, deriblant den norske offentlige utredningen fra 2008, som oppgaven ønsker å gjøre seg sammenlignbar med. Funksjonsformen til modellene i oppgaven kommer dermed til å se ut som dette, og utvides på dette fundamentet:

$$(10) \ln \text{timelønn}_i = \alpha + \beta_1 \text{female}_i + \beta_2 X_i + \varepsilon_i$$

Modellens egenskap ved funksjonsformen tilsier at  $100 * \beta_1$  gir en prosentvis lønnsforskjell mellom kjønnene ved endringer i variabelen  $X_i$ . Likningen viser en logaritmisk lønnsforskjell og korrelerer en prosentvis differanse i gjennomsnittslønn mellom kjønnene.

Gjennom analysen vil denne likningen stadig utvides med ytterligere variabler for andre kontrollvariabler vi ønsker å analysere opp mot den avhengige variabelen. Disse vil nå bli presentert, hvor verdien for  $X_i$  i likningen ovenfor stadig utvides med relevante variabler.

$$(11) \ln \text{timelønn}_i \\ = \alpha + \beta_1 \text{female}_i + \beta_2 \text{arbeidserfaring}_i + \beta_3 \text{arbeidserfaring}_i^2 + \varepsilon_i$$

$$(12) \ln \text{timelønn}_i \\ = \alpha + \beta_1 \text{female}_i + \beta_2 \text{arbeidserfaring}_i + \beta_3 \text{arbeidserfaring}_i^2 \\ + \beta_4 \text{utdanning}_i + \varepsilon_i$$

$$(13) \ln \text{timelønn}_i \\ = \alpha + \beta_1 \text{female}_i + \beta_2 \text{arbeidserfaring}_i + \beta_3 \text{arbeidserfaring}_i^2 \\ + \beta_4 \text{utdanning}_i + \beta_5 \text{arbeidsstiasjon}_i + \varepsilon_i$$

$$\begin{aligned}
(14) \ln \text{timelønn}_i &= \alpha + \beta_1 \text{female}_i + \beta_2 \text{arbeidserfaring}_i + \beta_3 \text{arbeidserfaring}^2_i \\
&+ \beta_4 \text{utdanning}_i + \beta_6 \text{arbeidssektor}_i + \varepsilon_i
\end{aligned}$$

Likning (11) tar høyde for effekten av arbeidserfaring kan ha på lønnsnivået mellom kvinner og menn. Merk at både nivået på arbeidserfaring, samt den kvadrerte verdien, foretas i analysen. Dette er for å medregne en potensiell ikke-lineær effekt av erfaring. I likning (12) utvides modellen for å også vurdere effekten av utdanningsnivå. Likning (13) utvider modellen nok en gang ved å medberegne effekten av partners arbeidssituasjon har på lønnsnivået, hvor i likning (14) tar den også observert økonomisk sektor i kalkuleringen.

#### 4.4. Oppsummering

I kapittel 4 er den empiriske strategien for oppgavens analyse presentert. Modellen benytter minste kvadraters metode for å estimere en modell for utvalget, og det er blitt vist til fundamentet denne estimeringen baserer seg på. Således ble oppgavens funksjonsform frembrakt og argumentert for. Funksjonsformen i den videre analysen vil være en logaritmisk-lineær funksjonsform. Ytterligere tilleggsvariabler ble lagt til i modellen og drøftet mot hva oppgavens analyse skal bevise.

## 5. Empiriske resultater

### 5.1. Innledning

I dette kapittelet skal det presenteres oppgavens empiriske funn basert på datasettet og modellens funksjonsform. Deretter skal funnene analyseres og drøftes. Så vil det foretas analyser av problemstillingens tilhørende tilleggsspørsmål, og en tilsvarende analyse og drøfting av disse skal frembringes.

### 5.2. Empiriske hovedresultater

Vi starter med å undersøke om det er kjønnsforskjeller i lønn og lønnsavkastning som følger av arbeidserfaring og utdanning. Undersøker derfor regresjonen i likningene (10), (11) og (12) som gir følgende estimater for modellenes parametere:

Modell:	10	11	12
VARIABLER	ln timelønn	ln timelønn	ln timelønn
female	-0.134654 (0.0118)	-0.132511 (0.0105)	-0.137214 (0.00950)
Arb.erfaring		0.0410321 (0.00153)	0.0298402 (0.00144)
Arb.erfaring <sup>2</sup>		-0.0007531 (3.57e-05)	-0.000494 (3.35e-05)
Utdanning			0.0580524 (0.00203)
Konstantledd	5.370181 (0.00830)	4.975908 (0.0146)	4.211876 (0.0298)
Observasjoner	3,558	3,556	3,555
R <sup>2</sup>	0.035	0.244	0.385

Tabell 3

Av kjønnsvariabelen female ser vi at timelønnsnivået er varierer mellom 13,3-13,7% mindre hos kvinner enn menn i de ulike modellene. Utfører vi en T-test av kjønnsvariabelen  $\beta_1$ , tester vi nullhypotesen  $H_0: \beta_1 = 0$  mot  $H_1: \beta_1 \neq 0$ . Denne testen gir at  $\beta_1 \neq 0$  er signifikant for alle modellene, og at nullhypotesen om at kjønn ikke påvirker lønnsnivå kan avkreftes. Dette tilsier at kjønnsforskjellen er statistisk signifikant i hver modell. Med andre ord så er det en signifikant forskjell i lønnsnivå utelukkende basert på kjønn i seg selv, men også når en tar både utdanning og arbeidserfaring i betraktning.

Av modell (11) ser vi at et års ytterligere arbeidserfaring har bidrar til en effektiv økning av lønnsnivå med ca. 4,1%, hvor denne effekten er signifikant. Effekten av den kvadrerte variabelen for arbeidserfaring i denne modellen viser til at en økende arbeidserfaring også virker negativt på timelønnsnivået med omtrent 0,0008% for hvert år ekstra med arbeidserfaring. Årsaken til disse motstridene effektene kommer av at det må være et punkt hvor arbeidserfaring begynner å virke negativt på lønnsnivå. Grunner til dette kan være at man kommer til en alder hvor en er mindre produktiv enn man ville vært ved mindre arbeidserfaring. Lønnseffekten av arbeidserfaring vil være positiv frem til en har akkumulert 27,2 år med erfaring, hvor den etter dette nivået vil påvirke lønnsnivået negativt.

Videre i analysen tester vi om forskjeller i arbeidserfaring kan forklare kjønnsforskjeller i timelønn. Vi analyserer nå estimatorene for modellen i likning (11). Til dette benytter vi en såkalt F-test for å undersøke om parameterne  $\beta_2$  og  $\beta_3$  påvirker lønnsnivå. En F-test er lik en t-test, bare hvor en tester effekten av flere variabler samtidig (Thomas, s. 208). Nullhypotesen blir dermed  $H_0: \beta_2 = \beta_3 = 0$ , som skal testes mot den alternative hypotesen om at minst  $\beta_2$  eller  $\beta_3$  påvirker lønnsnivå, altså  $H_1: \beta_2, \beta_3 \neq 0$ . Som følger av testen ser vi at kritisk verdi ved testen overskrides av testobservator, og nullhypotesen kan derfor avkreftes ved signifikant nivå. Dette innebærer at arbeidserfaring har en signifikant betydning for lønnsnivå. På grunnlag av dette kan vi også avvise at det er en lineær sammenheng mellom arbeidserfaring og lønn. Ytterligere ser vi også at arbeidserfaringen ikke har en særlig stor betydning for graden av lønnsforskjeller mellom kjønnene, ettersom kvinner fortsatt tjener  $-0,133 \cdot 100\% = 13,3\%$  mindre i timen enn menn for gitt nivå av arbeidserfaring. Selv om kjønnsforskjellen er redusert fra modell (10) med ca. 0,2% fordi vi nå i tillegg analyserer arbeidserfaring, kan vi fortsatt ikke si at forskjeller i arbeidserfaring gir en særlig større forklaring på hvorfor kvinner opplever en systematisk lavere timelønn enn menn. Med dette ser vi at vi må grave videre, og vi fortsetter analysen ved å utvide regresjonsmodellen.

I modell (12) analyseres også effekten økt utdanning har på lønn. Ifølge modellen ser vi at avkastningen ett års utdanning har på lønn er omtrent 0,058, som vil si at timelønnen øker med 5,8% for hvert ytterligere år med utdanning. Vi utfører også her en t-test for å undersøke om utdanning har en betydning for lønnsnivå. Dette gjør vi ved å undersøke nullhypotesen  $H_0: \beta_4 = 0$  mot  $H_1: \beta_4 \neq 0$ , og finner også ut her at nullhypotesen kan

avkrefte ettersom testobservator overskrider kritisk verdi. I denne modellen øker estimatet lønnsforskjeller basert på kjønn til 13,7%, en økning med nesten 0,5% fra forrige modell. Kjønnforskjeller i lønn kan med andre ord heller ikke forklares av forskjeller i utdanning. Likevel øker determineringskoeffisienten til modell (12) med over 14% fra modell (11). Dette betyr også at mer av variasjonene relatert til lønnsnivået mellom kjønnene forklares når utdanning også tar høyde for. Dette er en rettferdiggjort sammenligning ettersom begge modeller benytter den samme avhengig variabelen.

Nå skal oppgaven undersøke om det forekommer forskjeller i lønnsavkastning mellom kvinner og menn av utdanning, altså om kvinner har ulik avkastning av et ekstra år med utdanning enn hva menn har. Dette gjør vi ved å lage en interaksjonsvariabel mellom kjønns- og utdanningsvariabelen. Modellen vil da bli en utvidet versjon av likning (12), og de empiriske funnene av analysene presenteres i tabellen under.

Modell	(12')		(11')	
Variabler	ln lønn		ln lønn	
	Estimert koef.	Standardavvik	Estimert koef.	Standardavvik
female	-0,135584	(0.00967)	-0,137579	(0.00949)
Arb.erfaring	0,0298497	(0.001438)	0,031539	(0.00154)
Arb.erfaring2	-0,0004939	(0.0000355)	-0,0005092	(0.0000338)
Utdanning	0,0597329	(0.0027483)	0,0580144	(0.002028)
Utdanning * female	-0,0035499	(0.0039126)		
Arb.erfaring * female			-0,0024247	(0.0007931)
Konstantledd	4.187286	0,0402813	4,18736	(0.0308256)
Observasjoner	3555		3555	
R2	0,3856		0,3871	

Tabell 4

(12')  $\ln \text{timelønn}_i$

$$= \alpha + \beta_1 \text{female}_i + \beta_2 \text{arbeidserfaring}_i + \beta_3 \text{arbeidserfaring}_i^2 + \beta_4 \text{utdanning}_i + \gamma_1 ((\text{utdanning}_i - \overline{\text{utdanning}}) * \text{female}_i) + \varepsilon_i$$

Vi ser av likningen ovenfor at interaksjonsleddets estimeres som variabelen for utdanningsavvik fra gjennomsnittlig utdanningsnivå i utvalget multiplisert med kjønnsvariabelen. Hadde ikke gjennomsnittet blitt trukket fra ville lønnsforskjellene også blitt påvirket av

observasjoner uten utdanning. Når gjennomsnittet trekkes fra kan nemlig  $\beta_1$  tolkes som lønnsforskjell mellom kjønn som har et gjennomsnittlig antall år utdanning. Her kan også  $\gamma_1$  tolkes som effekten avviket fra gjennomsnittlig utdanningsnivå har på lønn. Estimater for  $\gamma_1$  er dog -0,004%. Dette kan tolkes som at kvinner tjener 0,004% mindre på ett års utdanning enn hva menn hadde gjort. En t-test konkluderer derimot at dette nivået er statistisk ikke-signifikant. Vi kan derfor ikke påstå at det forekommer forskjeller i lønnsavkastning mellom kvinner og menn, og at kjønnsforskjeller i lønn ikke kommer til syne på grunn av dette i oppgavens datasett. Dette er i tråd med Gary Beckers teori om humankapital, hvor økt utdannelse gir avkastning på lønnsnivå uavhengig av kjønnsnivå (Becker, 1964), men også imot tidligere forskning som hevdet at det forekom kjønnsforskjeller i avkastning på høyere utdanning (NOU 2008:6, s.58).

Kan det derimot forekomme forskjeller i lønnsavkastning mellom kvinner og menn av arbeidserfaring? Dette undersøker vi ved å lage en tilsvarende interaksjonsvariabel som ovenfor, men denne gang mellom kjønns- og arbeidserfaringsvariabelen. Dette gir oss en utvidet utgave av likning (11):

$$\begin{aligned}
 (11') \ln \text{timelønn}_i &= \alpha + \beta_1 \text{female}_i + \beta_2 \text{arbeidserfaring}_i + \beta_3 \text{arbeidserfaring}_i^2 \\
 &+ \gamma_2 \left( (\text{arbeidserfaring}_i - \overline{\text{arbeidserfaring}}) * \text{female}_i \right) + \varepsilon_i
 \end{aligned}$$

Argumentasjonen bak interaksjonsvariabelen er lik med den ovenfor. Her kan  $\beta_1$  tolkes som kjønnsforskjeller i lønn mellom observasjoner med lik gjennomsnittlig arbeidserfaring, mens  $\gamma_2$  tolkes som effekten avviket fra dette gjennomsnittet har på lønnsnivå. Av regresjonen får vi at estimatoren for  $\gamma_2$  er -0,0024, som vil si at kvinner har 0,24% lavere avkastning på arbeidserfaring enn menn. En t-test indikerer at dette resultatet er signifikant helt ned til 0,2% signifikansnivå, og at nullhypotesen om at en kvinnes arbeidserfaring ikke påvirker lønnsforskjeller kan avkreftes. Dette betyr at kvinner faktisk opplever en kjønnsdiskriminering i avkastningen av arbeidserfaring, dog viser resultatet fra dette datasettet at denne forskjellen er svært liten. Likevel er den signifikant betydelig, og vi kan konkludere i at avkastning på arbeidserfaring på lønnsnivå avhenger av kjønn.

En årsak til dette utslaget kan være av at vi ser fra delkapittel 3.5 at menn har noe høyere arbeidserfaring enn kvinner, som kan forklare hvorfor det fremdeles kommer frem en grad

av kjønnsforskjell i denne variabelen. Årsakene til at menn overhode har mer arbeidserfaring enn kvinner kan muligens komme av at kvinner utdanner seg lenger enn menn, som vil si at menn har mulighet for å komme ut i arbeid tidligere enn kvinner. I følge Begg kan denne forskjellen også forekomme av at kvinner har avstått fra arbeidserfaring for biologiske årsaker, med andre ord for å etablere familie og føde barn. Begg teoriserer at arbeidsgiver investerer flere ressurser i menn enn kvinner på grunn av at kvinner er mer utsatt for f.eks. foreldrepermisjon. Selv om resultatet ikke er særlig stort, kan den forekomme av det Beggs teori innebærer (Begg, 2014; s.244). Av tidligere forskning utpekes ikke kjønnsforskjeller i arbeidserfaringsavkastning som en hovedårsak til at lønnsforskjeller mellom kvinner og menn oppstår (NOU 2008:6, s.75).

Oppgaven vil nå begi seg ut på problemstillingen om det forekommer en relasjon mellom kjønnsforskjeller i timelønn og partners arbeidssituasjon i Norge. I den deskriptive statistikken observerte vi flere skillelinjer mellom kvinner og menn gjennom denne variabelen. Deriblant at mesteparten (78,2%) av alle kvinners partnere, antatte menn, jobber i fulltidsstillinger, og at over 1 av 5 (21,3%) av menns partnere, antatte kvinner, jobber i deltidsstillinger. Variabelen har en indirekte natur som gjør det vanskelig å benytte variabelens isolerte estimatører til å formulere en fornuftig regresjonsanalyse som kan benyttes til å tolke sammenhenger på kjønnsforskjeller i timelønn. Likevel viste det seg at modellen for likning (13) ga interessante resultater på hva av partnerens arbeidssituasjon som påvirker lønnsnivået på tvers av kjønnene. Det vil derfor foretas en kort gjennomgang av modellen, som det er foretatt enkelte endringer på:

$$\begin{aligned}
 (13') \ln \text{timelønn}_i &= \alpha + \beta_1 \text{female}_i + \beta_2 \text{arbeidserfaring}_i + \beta_3 \text{arbeidserfaring}_i^2 \\
 &+ \beta_4 \text{utdanning}_i + \beta_5 \text{arbeidssituasjon}_i \\
 &+ \gamma_3 (\text{arbeidssituasjon}_i * \text{female}_i) + \varepsilon_i
 \end{aligned}$$

I appendiks vil det være en tabell for de empiriske resultatene av modell (13) som vi vil benytte til å drøfte analysen. Til å begynne med ser vi at kvinner opplever 16,8-17% lavere timelønn enn menn, som er en økning fra modellene (10)-(12), når vi tar partners arbeidssituasjon med i estimeringen av kjønnsforskjeller. Det vil si at denne modellen forklarer mindre av kjønnsforskjeller i lønn enn de tidligere modellene. Dette gjenspeiles med at alle utgavene av modell (13') har en forklaringsvariabel på omtrent 0,298, som er en reduksjon fra modell (12) med sin determinasjonskoeffisient på 0,385. Modellen er

formulert slik at  $\beta_5$  viser hvordan menns lønnsnivå påvirkes av sin partners arbeidssituasjon. Det vil si at menn med eksempelvis en partner som jobber fulltid kan forvente  $\beta_5 * 100\%$  høyere lønn. Tilsvarende kan kvinner med partner som jobber fulltid forvente  $(\beta_5 + \gamma_3) * 100\%$  høyere lønn. Interaksjonsvariabelen  $\gamma_3$  benyttes for kombinere variabelen for en gitt arbeidssituasjon med kjønnsvariabelen.

Variabelen for arbeidssituasjon vil som sagt stå mellom fulltidsarbeid, deltidsarbeid, arbeidsledighet og andre arbeidssituasjoner, som spesifisert i 3.2. Det foretas en egen regresjon for estimatene av hver av de fire arbeidssituasjonene. Der hvor partner jobber fulltid, kommer det av utvalget at en mann med partner som jobber fulltid kan forvente  $(0,0017 * 100\%) = 1,7\%$  høyere lønn som følger av dette. En kvinne med partner som jobber fulltid kan forvente  $(0,0017 + 0,028) * 100\% \approx 3,0\%$  høyere lønn. Det som er interessant i denne påstanden er at det påvirker kvinner mer positivt å ha en partner i fulltidsarbeid enn hva det gjør for menn. Knytter vi dette opp mot deskriptiv statistikk for variabelen, ser vi at flesteparten (over 78%) av menn jobber fulltid. En mulig forklaring på at dette teller positivt for kvinner kan være at det forekommer en høy konsentrasjon av fulltidsansatte menn, som vi også vet hovedsakelig er ansatt i privat sektor (nesten 75% av alle menn). Vi har observert at det er her de mest høytlønnede mennene befinner seg (NOU 2008:6, s.65). En skulle i utgangspunktet tro at partners arbeidssituasjon ikke skulle ha en påvirkning på ens eget lønnsnivå, og det blir klart at denne hypotesen må testes.

Utfører vi en F-test om denne estimeringen kommer vi frem til at den er i høy grad statistisk ikke-signifikant. Heller ikke t-tester på enkeltvariablene  $\beta_5$  og  $\gamma_3$  er statistisk betydelig. Altså nullhypotesen om at estimatene ikke har betydning kan ikke avkreftes, verken for menn eller kvinner. Vi kan dermed avkrefte hypotesen om at ens partner jobber fulltid påvirker lønnsnivået hos noen av kjønnene. Av dette kan vi også deduktivt konkludere at om partner jobber fulltid ikke vil påvirke kjønnsforskjeller i lønn. Vi går videre med å analysere neste utgave av modellen, hvor vi drøfter virkningen en deltidsansatt partner har på lønnsnivået.

Når deltidsansatte partnere vurderes, så ser vi at kjønnsvariabelen reduseres fra forrige modellutgave med omtrent 0,002, og at forklaringskoeffisienten  $R^2$  øker med omtrent 0,001. Dette gir en liten antydning på at kjønnsforskjeller forklares i en større grad når deltidsansatte partnere tas i betraktning i stedet for fulltidsansatte partnere. Av variabelen



$\beta_5$  ser vi at menns lønnsnivå påvirkes  $(0,028 * 100\%) = 2,8\%$  positivt av at partneren jobber deltid. Mer interessant er at vi ser at estimatet for interaksjonsvariabelen  $\gamma_3$  er  $-0,086$ . At kvinners partner jobber deltid påvirker lønnsnivået til kvinner med  $(0,028 - 0,086) * 100\% \approx -5,8\%$ . Dette vil si at dersom en kvinne har en partner som jobber deltid, påvirker dette lønnen hennes negativt med omtrent 5,8%. Kvinner med deltidsarbeidene partnere opplever også  $(\beta_1 - 0,058) * 100\% = -22,9\%$  lavere lønn enn menn i tilsvarende situasjon. Utfører vi en t-test på hvert av estimatene får vi at estimatet for  $\beta_5$  kun er statistisk signifikant ned til 13,1%-signifikansnivå, som er svært lite. Derimot er interaksjonsvariabelen  $\gamma_3$  statistisk signifikant helt ned til 3,4%-signifikansnivå, som er betydelig. Vi kan derfor konkludere i at om kvinner har en partner som jobber deltid, vil dette negativt påvirke kvinnens lønnsnivå, og øke lønnsforskjeller mellom kvinner og menn. Vi kan foreløpig ikke si det samme om betydningen av en deltidsansatt partner for menn. Med en samkjørt F-test av begge variablene viser at vi kan avvise nullhypotesen om at partners arbeidssituasjon ikke har påvirkning på lønnsnivået hos begge kjønn ned til 8,9%-signifikansnivå. Dette betyr at hypotesen ikke vil gjelde dersom vi har et krav om 5%-signifikansnivå, men sammenhengen er derimot betydelig hvis vi benytter oss av 10%-signifikansnivå i stedet.

Særlig interessant med deltidsvariabelen for partners arbeidssituasjon er at det forekommer en positiv sammenheng med lønn for menn, men en negativ sammenheng med lønn for kvinner. Hvordan kan dette ha seg? Vi nevnte ovenfor at de fleste menn jobber fulltid, og at de fleste menn befinner seg i privat sektor. I denne sektoren er det også slik at deltidsstillinger er de dårligst betalte, og at fulltidssektorer er de best betalte stillingene (NOU 2008:6, s. 72). Deduktivt kan vi drøfte at dersom du er en kvinne med en partner som er en mann, og partneren din jobber deltid, jobber han mest sannsynlig i privat sektor. At partneren da ikke er den del av majoriteten av menn som jobber fulltid kan antyde at partneren har gjort livsvalg som har forårsaket at han havnet i en lavtlønnet deltidsstilling, sammenlignet med fulltidsstillingene i privat sektor. En kan anta at partnere er partnere nemlig fordi de tiltrekkes av mennesker med flere likhetstrekk enn ulikhetstrekk med seg selv. En kan dermed tenke seg at dersom en kvinne, hvis partner har gjort livsvalg som har ført til at han jobber deltid kontra det normative fulltidsarbeidet, har også en kvinne gjort tilsvarende livsvalg som fører henne til en lavere timelønn enn hun ellers ville hatt. Dersom kvinnen hadde foretatt livsvalg som hadde ført henne mot en partner som er innenfor det normative flertallet av menn med fulltidsstillinger, kan en

tenke seg at påvirkningen av partners arbeidssituasjon hadde vært positiv, slik den er i regresjonen av fulltidsansatt partner.

Tilsvarende forekom det en positiv påvirkning på menns lønnsnivå dersom partneren var deltidsansatt. Som vi ser av den deskriptive statistikken, og av tidligere forskning, jobber flere kvinner i deltidsstillinger. Dersom det er «vanlig» at menns partner jobber deltid, kan det hende at mannen også er mer innenfor det «vanlige» sjiktet med tanke på lønnsituasjon, med eksempelvis en fulltidsstilling i privat sektor. Derfor kan vi også tenke oss hvorfor menns lønnsnivå avhenger positivt av at mannen har en partner som jobber deltid. At det forekommer en tiltrekningseffekt hvor menn utenfor den høytlønnede normalen påfører økt sannsynlighet for kvinnen til å tjene mindre er ingen vanntett påstand, men likevel en mulig forklaring på at denne statistisk signifikante sammenhengen.

Videre skal vi også undersøke signifikansen av partners arbeidssituasjon på lønn dersom partneren er arbeidsledig. Når arbeidsledighet analyseres er estimert kjønnsvariabel for lønn, samt forklaringskoeffisienten  $R^2$ , omtrent den samme som i forrige modell. Vi ser at  $\beta_5$  estimeres til -0,102, som vil si at menns lønnsnivå påvirkes negativt med  $\beta_5 * 100\% = 10,2\%$  dersom partneren er arbeidsledig. Gjennomfører vi en t-test for å teste om denne variabelen er ubetydelig for lønnsnivå, kan denne hypotesen avkrefte helt ned til 5,6%-signifikansnivå. Vi kan derfor ikke avkrefte hypotesen om at menns lønnsnivå ikke avhenger av at partner er arbeidsledig ved 5%-signifikansnivå. Interaksjonsvariabelen estimeres til 0,01, som vil si at kvinners lønnsnivå påvirkes  $\beta_5 + \gamma_3 * 100\% = -9,0\%$  dersom partner er arbeidsledig. Dette estimatet kan heller ikke bekrefte at kvinners lønnsnivå påvirkes av at partner er arbeidsledig. At det forekommer en negativ sammenheng for begge kjønn er derimot forståelig, for en kan tenke seg at dersom partner er i en situasjon hvor han eller hun er utsatt for arbeidsledighet, kan det tenkes at en selv også befinner seg i en slik situasjon hvis vi bygger drøftingen på likhetstiltrekningen diskutert tidligere.

Vi ser at det forekommer et tilsvarende resultat for regresjonsmodellen når vi analyserer effekten på lønn av at partner befinner seg i en annen arbeidssituasjon enn de foregående arbeidssituasjonene. Hvilke arbeidssituasjoner denne dummyvariabelen innebærer er spesifisert i kapittel 3.2. Effekten for menns lønnsnivå av at partner befinner seg i en

annen arbeidssituasjon er  $\beta_5 * 100\% = -2,2\%$ , mens for kvinner er den  $\beta_5 + \gamma_3 * 100\% = -0,4\%$ . Heller ikke her er noen av estimatene statistisk signifikante, i tillegg til at de ikke er betydelig store effekter i seg selv. Interessant nok kan en tenke seg at dersom partner er i en annen arbeidssituasjon, som f.eks. kan være at partneren er student, er det for så vidt også trolig at en selv er student, lærling eller veldig nylig ferdigutdannet. At det da forekommer en negativ effekt av partners arbeidssituasjon på eget lønnsnivå høres i seg selv trolig ut. Grunnet t-testen kan vi likevel ikke bekrefte at en partner i en annen arbeidssituasjon enn de foregående påvirker eget lønnsnivå, og vi konkluderer at denne variabelen ikke har en betydelig effekt.

### 5.3. Analyse av tilleggsspørsmål

Vi skal nå undersøke om det forekommer forskjeller i timelønn mellom kjønnene avhengig av økonomisk sektor. Vi begynner med å undersøke likning (14) som ser om kjønnsforskjeller i lønn kommer tydeligere frem i offentlig eller privat sektor.

VARIABLER	ln timelønn
female	-0.139218 (0.0123)
Arb.erfaring	0.0296452 (0.00143)
Arb.erfaring <sup>2</sup>	-0.0004853 (3.34e-05)
Utdanning	0.0607598 (0.00209)
Offentlig sektor	-0.070492 (0.0154)
Offentlig sektor * female	0.037676 (0.0202)
Konstantledd	4.190032 (0.0302)
Observasjoner	3,554
R <sup>2</sup>	0.391

Tabell 5

$$(14') \ln \text{timelønn}_i = \alpha + \beta_1 \text{female}_i + \beta_2 \text{arbeidserfaring}_i + \beta_3 \text{arbeidserfaring}_i^2 + \beta_4 \text{utdanning}_i + \beta_6 \text{arbeidssektor}_i + \gamma_4 (\text{female}_i * \text{arbeidssektor}_i) + \varepsilon_i$$

La oss vurdere modellen med hensyn på offentlig sektor. Ved å utføre en regresjonsanalyse på denne modellen, vil modellen estimere lønnsnivået for menn i privat sektor dersom dummyvariablene for kvinner er lik null.  $\beta_1$  i modellen blir dermed kjønnsforskjellen i timelønn i privat sektor,  $\beta_6$  vil bli sektorforskjellen i lønn for menn, og  $\gamma_4$  blir lønnsavviket mellom kjønnene i hver sektor. Denne variabelen estimeres i regresjonen til å være 0,037, som antyder at lønnsforskjellen til kvinner i offentlig sektor er 3,7% lavere enn i privat sektor, siden  $\gamma_4$  er positivt. Derimot er ikke denne sammenhengen statistisk signifikant ved 5%-signifikansnivå, ettersom testobservator ikke overstiger kritisk verdi ved denne t-fordelingen. Hypotesen om at kvinner i offentlig sektor ikke opplever lavere lønn avkrefte helt ned til 6,1%-signifikansnivå, men sammenhengen blir ikke like tilfredsstillende ettersom den ikke overholder kravet om 5%-signifikansnivået. Ser vi vekk i fra dette inntil videre og går videre med modellen, observerer vi estimatet for  $\beta_1$  er -0,139, som tilsier en lønnsforskjell mellom kvinner og menn i privat sektor er omtrent  $100 * (-0,139) = -13,9\%$ . Kjønnsforskjellen i offentlig sektor blir summen av estimatene  $\beta_1$  og  $\gamma_4$ , og tilsier at kvinner tjener omtrent  $-13,9\% + 3,7\% = 10,2\%$  mindre enn menn i offentlig sektor. Av forklaringsvariabelen hos denne modellen på 0,391 ser vi at den forklarer mer av de observerte variasjonene enn noen av de andre modellen så langt i oppgaven.

Hvis vi vurderer estimatet for  $\gamma_4$  som statistisk signifikant kan vi konkludere i at det forekommer en lavere kjønnsforskjell i timelønn i offentlig sektor enn i privat sektor. Dette støttes opp av tidligere forskning, som viste til at det var større ulikhet i lønnsnivå mellom kjønnene i privat enn offentlig sektor, grunnet høy konsentrasjon av kvinner i lavtlønnede stillinger i privat sektor (NOU 2008: 6, s.65).

#### 5.4. Oppsummering

I dette kapitlet er empiriske resultater presentert, analysert og drøftet ut i fra oppgavens problemstillinger. Hittil kan vi konkludere i kjønnsforskjeller i avkastning i arbeidserfaring, at kvinners lønn påvirkes av at partner jobber deltid, og at det er større kjønnsforskjeller i lønn i privat sektor enn i offentlig. Likevel forekommer det en systematisk kjønnsforskjell som ikke lar seg forklare av oppgavens variabler.

## 6. Oppsummering og konklusjon

Oppgavens problemstilling var motivert av å avsløre ytterligere årsaker til at det forekommer lønnsforskjeller mellom kjønnene enn hva som er presentert i tidligere forskning. Oppgaven gjorde dette ved å undersøke kjønnsforskjeller i lønn og lønnsavkastning av utdanning og arbeidserfaring, og om det er en sammenheng mellom kjønnenes lønnsnivå og arbeidssituasjonen til eventuelle partnere. I tillegg undersøkte oppgaven om det forekom lønnsforskjeller mellom kvinner i offentlig og privat sektor sammenlignet med menn. Oppgaven etablerte et teoretisk rammeverk som presenterte teorier om hvordan og hvorfor kjønnsforskjeller i lønn forekom, som ble vurdert opp mot oppgavens funn. Tidligere forskning viste at særlig forskjeller i arbeidssituasjon og -sektor var en ledende årsak til at det observeres kjønnsforskjeller i lønnsnivå (NOU 2008:6, s.75). Videre ble det nevnt at lønnsforskjeller forekommer i avtagende grad over de siste årene (Gunnes, 2019).

Med dette ble det foretatt en presentasjon av oppgavens datamaterialet som oppgaven baserer sine funn og analyser på. Dette innebar en systematisk forklaring av de ulike variablene som ble benyttet, en presentering av deskriptiv statistikk for avhengig og uavhengige variabler, og en gjennomgang av datasettets fordeler og ulemper. En fordel ved datasettet er at det inkluderte spesifikke variabler som var ønskelig å undersøke, mens svakheten var at disse variablene i seg selv var i overkant uspesifiserte, hvor det var flere aspekter som også hadde vært interessant å undersøkt. Flere forutsetninger om datasettets variabler ble foretatt for å motvirke effekten av ulempene som ble drøftet. Deretter ble det gjennomgått hvilken empiriske strategien oppgaven benyttet, og det ble foretatt valg av modellenes funksjonsform, begge kritisk for resultatene oppgaven gjengir.

De empiriske resultatene konkluderte i følgende svar på problemstillingen: alle estimerte modeller for lønnsnivå viste at det forekommer en tydelig forskjell i lønn basert på kjønn, og at denne forskjellen var statistisk signifikant. Lønnsforskjellen varierte mellom 13-17% lavere lønn for kvinner, hvor modellenes variabler forklarte cirka 30-39% av kjønnsforskjellene. Kjønnsforskjellen kan i liten til ingen grad forklares av systematiske forskjeller mellom kvinner og menn i utdanningsnivå og arbeidserfaring. Disse variablene har en effekt på lønnsnivå, men ikke lønnsforskjeller mellom kjønnene. Avkastningen på utdanningen er heller ikke forskjellig mellom kjønnene. Det forekom en liten forskjell i avkastning på arbeidserfaring, hvor kvinner hadde dårligere avkastning enn menn, men denne forskjellen var svært liten, og kan ikke anses som en avgjørende årsak til kjønnsforskjeller.

Videre kunne oppgaven også konkludere at lønnsnivå hos kvinner påvirkes særlig av at partner jobber deltidsstillinger, og at denne påvirkningen er statistisk signifikant. Årsaken til denne sammenhengen kan være summen av en assosiasjons- og tiltrekningseffekt basert på datasettets utvalg, hvor deltidsvariabelen indirekte antyder at kvinner med deltidsarbeidende partnere var utsatt for lavere lønn enn andre. Selv om sammenhengen i seg selv er statistisk signifikant, må det likevel påpekes at modellen forklarte mindre av variasjonen enn tidligere lønnsmodeller i oppgaven. Videre kan det konkluderes at det forekommer mindre kjønnsforskjeller i lønn i offentlig sektor enn i privat sektor, med omtrent 3,7%. Til tross for analysens funn gjenstår likevel en del av kjønnsforskjellene uforklart av modellene, og utfordringen av systematiske kjønnsforskjeller i lønn vedvarer som et av flere samfunnsøkonomiske problemer landet fremdeles står ovenfor.

## 7. Referanser

Begg, D, Vernasca, G., Fischer, S. og Dornbusch, R. (2014) *Economics 11th Edition*. Berkshire, Storbritannia: McGraw-Hill Education

NOU 2008: 6 (2008) *Kjønn og lønn – Fakta, analyser og virkemidler for likelønn*. Oslo: Barne- og likestillingsdepartementet

Gunnes, N. (2019) *Lønnsforskjellene mellom kvinner og menn minker*. Tilgjengelig fra: <https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/lonnsforskjellene-mellom-kvinner-og-menn-fortsetter> (Hentet 22.03.2019)

Thomas, R. L. (2005) *Using statistics in economics*. Berkshire, Storbritannia: McGraw-Hill Education

Becker, G. S. (1964) *Human Capital*. Tilgjengelig fra: <https://www.econlib.org/library/Enc/HumanCapital.html> (Hentet 26.04.2019)

## 8. Appendiks

Utvidet tabell av modell (13):

Modell	13, fulltid	13, deltid	13, arbeidsledig	13, annet
VARIABLER	ln timelønn	ln timelønn	ln timelønn	ln timelønn
female	-0.173*** (0.0118)	-0.171*** (0.0118)	-0.169*** (0.0113)	-0.168*** (0.0114)
Arb.erfaring	0.0252*** (0.00192)	0.0254*** (0.00189)	0.0253*** (0.00189)	0.0251*** (0.00193)
Arb.erfaring <sup>2</sup>	-0.000418*** (4.19e-05)	-0.000425*** (4.12e-05)	-0.000424*** (4.12e-05)	-0.000417*** (4.23e-05)
Utdanning	0.0545*** (0.00249)	0.0548*** (0.00248)	0.0546*** (0.00248)	0.0546*** (0.00249)
Fulltid	0.00173 (0.0160)			
Fulltid * female	0.0280 (0.0264)			
Deltid		0.0284 (0.0188)		
Deltid * female		-0.0864** (0.0408)		
Arbeidsledig			-0.102* (0.0533)	
Arbeidsledig * female			0.0115 (0.0932)	
Annen arbeidssituasjon				-0.0220 (0.0211)
Annen arbeidssituasjon * female				0.0176 (0.0329)
Konstantledd	4.345*** (0.0419)	4.335*** (0.0416)	4.348*** (0.0415)	4.350*** (0.0423)
Observasjoner	2,427	2,427	2,427	2,427
R <sup>2</sup>	0.298	0.299	0.299	0.298

Standardavvik i parenteser