

Ingebjørg Frislid
Andreas Letnes
Ada Holsether Savjord

Analyse av lønnsforskjeller mellom kvinner og menn i Norge og Storbritannia

Bacheloroppgave i Samfunnsøkonomi
Veileder: Eric Myran Wee
Mai 2023

Ingebjørg Frislid
Andreas Letnes
Ada Holsether Savjord

Analyse av lønnsforskjeller mellom kvinner og menn i Norge og Storbritannia

Bacheloroppgave i Samfunnsøkonomi
Veileder: Eric Myran Wee
Mai 2023

Norges teknisk-naturvitenskapelige universitet
Fakultet for økonomi
Institutt for samfunnsøkonomi



Kunnskap for en bedre verden

Sammendrag

I forskningsmiljøet har det de senere årene vært økende interesse for yrkesaktive kvinner både nasjonalt og internasjonalt. Ettersom kvinner utgjør en stadig større andel av arbeidsstyrken, har lik lønn mellom kvinner og menn blitt et viktig politisk mål. I denne studien har vi undersøkt lønnsgapet mellom kvinner og menn i to europeiske land, Storbritannia og Norge, ved å bruke data fra Programme for the International Assessment of Adult Competencies (PIAAC). Dette er en internasjonal undersøkelse utført av Organisasjonen for økonomisk samarbeid og utvikling (OECD).

Analysen fokuserer på hvordan både humankapitalvariabler og arbeidsrelaterte variabler påvirker lønnsnivået til kvinner og menn. Studien identifiserer et lønnsgap mellom kvinner og menn i både Storbritannia og Norge, samt lønnsforskjeller mellom de to landene. Videre observeres et lønnsgap mellom privat og offentlig sektor, der norske menn i privat sektor har høyest timelønn. I Storbritannia, derimot, er timelønnen høyest for menn ansatt i offentlig sektor. Marginalavkastningen av utdanning og kognitive ferdigheter er også lik for begge kjønn i begge land.

Abstract

In the research community, there has been increasing interest in recent years in working women, both nationally and internationally. As women constitute an increasingly large proportion of the workforce, equal pay between women and men has become an important political goal. In this study, we have examined the gender wage gap in two European countries, the United Kingdom and Norway, using data from the Programme for the International Assessment of Adult Competencies (PIAAC). This is an international survey conducted by the Organisation for Economic Co-operation and Development (OECD).

The analysis focuses on how both human capital variables and work-related variables affect the wage level of women and men. The study identifies a wage gap between women and men in both the UK and Norway, as well as wage differences between the two countries.

Furthermore, a wage gap between the private and public sector is observed, with Norwegian men in the private sector having the highest hourly wage. In the UK, on the other hand, the hourly wage is highest for men employed in the public sector. The marginal returns of education and cognitive skills are also the same for both sexes in both countries.

Innholdsfortegnelse

1. Innledning	5
1.1 Problemstilling.....	5
1.2 Struktur	6
2. Teori	7
2.1 Humankapitalteori	7
2.2 Preferanseteori	8
2.3 Segregeringsteori	8
2.4 Lønnsgapet i dag.....	9
3. Datamaterialet	10
3.1 Om datamaterialet.....	10
3.2 Deskriptive analyser	11
3.2.1 Deskriptiv statistikk for avhengig variabel	12
3.2.2 Deskriptiv statistikk for uavhengige variabler	13
3.3 Styrker og svakheter ved datasettet	15
4. Metode	17
4.1 MLR- antagelser	17
4.2 Gjennomgang av teori og metode.....	19
5. Regresjonsanalyse	22
5.1 Modell A - Lønnsgap mellom kvinner og menn i Norge og Storbritannia	23
5.2 Modell B - Forskjell i lønnsnivå i Storbritannia og Norge.....	25
5.3 Modell C - Avkastning på utdanning for kjønn i Storbritannia og Norge.....	26
5.4 Modell D - Avkastning på ferdigheter for kjønn i Storbritannia og Norge	27
5.5 Modell E – Sammenligning av lønnsgap i offentlig og privat sektor for kjønn i Storbritannia og Norge	28
5.6 Modell F - Forskjell i lønnsgap med og uten arbeidskontrakt i Storbritannia og Norge	30

6. Konklusjon	32
7. Litteraturliste	33
8. Vedlegg	36
8.1 Valg av funksjonsformer	36
8.2 Test for uavhengighet	37
8.3 Test for normalitet	37
8.4 Test for multikollinearitet	38
8.5 Test for heteroskedastisitet	39
8.6 Resultater fra regresjonsanalysene	40

1. Innledning

Lønnsforskjeller mellom kvinner og menn har lenge vært et aktuelt tema i samfunnsdebatten og i forskningsmiljøene. Norge har aktivt arbeidet for likestilling i arbeidsmarkedet siden begynnelsen av 1970-tallet. I 1978 vedtok Stortinget Likestillingsloven med formål om å fremme likestilling og bekjempe diskriminering av kvinner. Den norske regjeringens politiske plattform legger vekt på en arbeidslivspolitik som likestilling og likelønn, utviklingsmuligheter og tilgang til kompetanseutvikling er sentrale komponenter.

I global gender trap report 2022, en analyse som ble gjennomført av World Economic Forum (WEF, 2022), ser vi at Norge ligger på tredje plass over land som har høyest likestilling mellom kvinner og menn. Storbritannia på den andre siden ser vi ligger på 22. plass. I henhold til Likestillings- og diskrimineringsloven § 34 er arbeidsgivere pålagt å lønne kvinner og menn likt, uavhengig av kjønn. Loven fastslår at «kvinner og menn i samme virksomhet skal ha lik lønn for samme arbeid eller arbeid av lik verdi. Lønnen skal fastsettes på samme måte, uten hensyn til kjønn». Til tross for dette, indikerer lønnsgapet at det fortsatt eksisterer en kjønnsbasert lønnsdiskriminering i det norske arbeidsmarkedet.

Norge og Storbritannia (ekskludert Nord-Irland) er to interessante land å sammenligne når det gjelder likestilling og lønnsforskjeller, ettersom begge landene har sterke velferdsstater, men ulike arbeidsmarkedsforhold og kulturelle normer. I denne bacheloroppgaven vil vi se nærmere på flere problemstillinger knyttet til lønnsforskjeller i disse to landene.

1.1 Problemstilling

Først vil vi se på spørsmålet om hvorvidt det er lik lønn for likt arbeid i Norge og Storbritannia, ved å analysere datasett for de to landene. Dette er et viktig tema, da lik lønn for likt arbeid er en grunnleggende rettighet som bidrar til å redusere lønnsforskjeller mellom kjønn og andre grupper.

Deretter vil vi undersøke ulikhetene i lønnsnivået mellom Norge og Storbritannia. Selv om begge landene er utviklede økonomier med en høy grad av velstand, er det fortsatt forskjeller i lønnsnivået, vi vil se på mulige årsaker til dette. Vi vil også undersøke avkastningen på

utdanning og ferdigheter for kvinner og menn i Norge og Storbritannia. Dette er et viktig tema da det kan bidra til å forklare forskjeller i lønn mellom kjønnene.

Videre vil vi se nærmere på lønnsgapet mellom offentlig og privat sektor for kjønnene i Norge og Storbritannia. Til slutt ser vi på lønnsforskjeller mellom kvinner og menn, med tanke på arbeidskontrakt. Gjennom denne bacheloroppgaven vil vi undersøke ulike aspekter ved lønnsforskjeller mellom kvinner og menn i Norge og Storbritannia, og se på hva som kan forklare disse forskjellene. Gjennom dette arbeidet håper vi å bidra til økt kunnskap om dette viktige temaet.

1.2 Struktur

Vi har nå tatt for oss oppgavens problemstilling og en innledning til kvinners arbeidsdeltakelse i både Norge og i Storbritannia i det første kapittelet. Det andre kapittelet vil presentere det teoretiske rammeverket og tidligere studier som er gjennomført. Videre skal vi i kapittel tre gå gjennom datamaterialet som er blitt bruk i oppgaven, for deretter å presentere funnene vi har fått fra den deskriptive analysen. Det fjerde kapittelet vil omhandle Mincers-modell som vi har valgt å benytte oss av i denne oppgaven. Deretter settes det opp seks ulike hypoteser som vil fortelle oss om hvilke faktorer som påvirker lønnsforskjellen mellom kvinner og menn. I det femte kapittelet skal vi så presentere de empiriske resultatene fra regresjonsanalysen. Deretter skal vi diskutere rundt resultatene ved bruk av teori og empiri, før vi i det sjette kapittelet skal oppsummere hovedfunnene fra analysene og deretter danne en konklusjon.

2. Teori

Denne oppgaven vil utforske og diskutere ulike teorier og perspektiver som bidrar til å forstå lønnsforskjeller mellom kjønnene. Vi vil ta for oss humankapitalteori, preferanseteori, segregeringsteori og til slutt se på lønnsgapet i dag og sammenligne situasjonen i Norge og Storbritannia. Disse teoriene og konseptene vil hjelpe oss med å få en dypere innsikt i de komplekse faktorene som påvirker lønnsforskjellene mellom kvinner og menn og hvordan disse faktorene varierer mellom land og arbeidsmarked.

2.1 Humankapitalteori

Humankapitalteori gir et viktig perspektiv for å forstå lønnsforskjeller mellom kjønnene. Ifølge Becker (1962) kan investeringer i menneskelig kapital bidra til å forklare lønnsforskjeller mellom kvinner og menn. Becker skriver i "Investment in Human Capital: A Theoretical Analysis" at en faktor som påvirker inntekt er investeringer i humankapital som inkluderer skolegang, opplæring på jobben, og andre aktiviteter som forbedrer ferdigheter og kunnskap (Becker, G s.9). Humankapital kan også ha innvirkning på lønnsforskjellene mellom kjønn. Dette skildres i en nyere studie av lønnsforskjeller innad i Norge, denne artikkelen fant at kvinner generelt opplever lavere avkastning på arbeidserfaring (Stokke, 2021). Becker påpeker at høyere utdanning og ferdigheter generelt fører til høyere inntekt, uavhengig av land og tidsepoke (Becker, G s.9). Imidlertid kan variasjoner i tilgang til og kvalitet på utdanning og opplæring mellom kjønnene føre til ulikheter i investering i humankapital på tvers av land. Disse ulikhetene kan igjen bidra til å forklare lønnsforskjellene mellom kvinner og menn. Når en høyere andel av kvinner investerer mindre i humankapital enn menn fører det til at kvinner får lavere lønn enn menn, gjennom for eksempel karriereavbrudd for omsorgspermisjon eller reduserte stillinger.

2.2 Preferanseteori

Preferanseteori er en økonomisk teori som beskriver hvordan individer gjør valg og tar beslutninger basert på deres subjektive preferanser. Denne teorien ble først introdusert av blant annet økonomene som Adam Smith og Jeremy Bentham på slutten av 1700-tallet og har siden blitt en sentral del av moderne mikroøkonomi (Stigler, 1950). Preferanseteorien kan brukes til å analysere valg og beslutninger som tas på en arbeidsplass, for eksempel en persons nyttefunksjon basert på faktorer som lønn, arbeidstid, fleksibilitet og jobbtilfredshet. Ved å forstå hvordan individuelle preferanser påvirker arbeidsmarkedet, kan vi få innsikt i hvorfor kvinner og menn kan ende opp med ulike lønninger. Tidligere studier har funnet forskjeller i preferanser mellom kvinner og menn når det gjelder arbeids- og familieliv. For eksempel har det blitt vist at kvinner er mer sannsynlig til å prioritere familie foran jobb enn menn, noe som kan føre til at de tar mer fleksible og mindre krevende stillinger (Goldin, 2014). Dersom kvinner og menn har forskjellige prioriteringer og avveininger når det kommer til arbeid, familie og fritid, kan dette resultere i ulike valg av yrker og arbeidstidsfleksibilitet, som igjen kan bidra til lønnsforskjeller.

2.3 Segregeringsteori

I Paula Englands artikkel "The gender revolution: Uneven and stalled" (2010) blir segregeringsteori relevant når vi ser på sammenhengen mellom kjønnssegregering i arbeidsmarkedet og ulikhetene i endringene i kjønnsystemet. Kjønnssegregering i arbeidsmarkedet oppstår når kvinner og menn konsentrerer seg i ulike yrker og sektorer. Englands forklaringer om kulturell og institusjonell nedvurdering av egenskaper og aktiviteter knyttet til kvinner, kan bidra til å forstå hvorfor kjønnssegregering oppstår og vedvarer. Arbeidsmarkedet forblir segregert etter kjønn når menn har lite insentiv til å gå inn i tradisjonelt kvinnelige yrker og kvinner har sterke økonomiske insentiver til å bevege seg inn i tradisjonelt mannlige felt. Bevegelse inn i mannsdominerte yrker og sektorer fører ikke nødvendigvis til en generell likestilling i arbeidsmarkedet. Stokkes funn understreker også kjønnssegregerings problemet gjennom at kvinner og menn ofte jobber i samme firma, men med ulike stillinger og lønnsnivåer (Stokke, 2021).

2.4 Lønnsgapet i dag

For å få en grundigere forståelse av lønnsgapet mellom kvinner og menn i dag, er det viktig å se på både nasjonale og internasjonale trender. I 2022 publiserte World Economic Forum (WEF) sin årlige rapport om Global Gender Gap; The Global Gender Gap Report 2022. Det er en av flere studier som har blitt gjennomført for å analysere lønnsforskjeller mellom kvinner og menn. Denne rapporten måler kjønnsforskjeller i fire kategorier: økonomisk deltakelse og muligheter, utdanning, helse og politisk deltakelse. I den økonomiske kategorien analyseres lønnsgapet mellom kvinner og menn. WEF definerer lønnsgapet som forskjellen i gjennomsnittlig timelønn mellom kvinner og menn på tvers av hele økonomien til ett samfunn. Rapporten viser at lønnsgapet fortsatt eksisterer i mange land, inkludert EU-landene, og at variasjonen mellom land er betydelig. Lønnsgapet mellom kvinner og menn har en tendens til å øke med alder, utdanning og ansiennitet. Rapporten fremhever at de store lønnsforskjellene mellom landene må forstås i sammenheng med andre faktorer i arbeidsmarkedet. Et betydelig lønnsgap er ofte forbundet med et sterkt kjønnsdelt arbeidsmarked eller en høy andel av kvinner som jobber med redusert stillingsprosent. Den globale utviklingen viser at lønnsgapet mellom kvinner og menn gradvis avtar, men det er fortsatt mye arbeid som må gjøres for å oppnå full likelønn.

Lønnsforskjeller i Norge og Storbritannia er påvirket av en rekke faktorer inkludert utdanning, arbeidserfaring, yrke, sektor og kjønn (OECD, 2021). Selv om det er noen likheter mellom de to landene, er det også betydelige forskjeller i hvordan lønnsforskjeller oppstår og håndteres. Den norske velferdsstaten tilbyr en rekke støtteordninger som bidrar til å redusere økonomiske ulikheter i samfunnet. Selv om det norske arbeidsmarkedet fortsatt er kjønnsdelt og har lønnsforskjeller mellom kvinner og menn, er forskjellene generelt lavere enn i mange andre land, inkludert Storbritannia (OECD, 2021).

Lønnsgapet mellom kjønnene i Norge påvirket av sektorfordelingen og valg av yrker, der kvinner oftere jobber i offentlig sektor og i tradisjonelle kvinneyrker (Statistisk sentralbyrå, 2023). Kjønnbaserte lønnsforskjeller er også et betydelig problem i Storbritannia, selv om landet har innført tiltak for å redusere lønnsgapet, er forskjellene mellom kvinner og menn fortsatt høyere enn i Norge. Dette kan blant annet skyldes at kvinner i Storbritannia i større grad jobber i lavtlønnede stillinger og i deltidsstillinger. Det vil påvirke den samlede lønnsstatistikken (Office for National Statistics, 2022).

3. Datamaterialet

I denne delen skal vi presentere datasettet som vi har benyttet oss av i bacheloroppgaven. Deretter skal vi presentere funnene våre i de deskriptive analysene vi har foretatt, og til slutt skal vi diskutere styrkene og svakhetene ved datamaterialet.

3.1 Om datamaterialet

Dataene som vi har benyttet oss av i denne bacheloroppgaven er hentet fra den internasjonale undersøkelsen om lese- og tallforståelse (PIAAC), som ble gjennomført av OECD i perioden august 2011 til april 2012 i 21 land. Målet med denne undersøkelsen var å se på virkningen de tre kognitive ferdigheter - leseforståelse, tallforståelse og problemløsning i IKT-miljø – vil ha på utdanning hos personer i alderen 16 til 65 år. Undersøkelsen inkluderte også spørsmål om andre faktorer som kjønn, utdanningsnivå, alder, lønn, arbeidssted og etnisitet.

I denne oppgaven har vi valgt å benytte oss av data fra et representativt utvalg av britiske og norske arbeidstakere, hvor vi har valgt å ha timelønn som den avhengige variabelen. Det er kun de respondentene som har oppgitt sin timelønn i spørreskjemaet som er inkludert i våre analyser. Vi antar at utelatelsen av informasjon om timelønn er tilfeldig, og at vårt utvalg derfor vil være representativt for populasjonen. I denne oppgaven ønsker vi å sette søkelyset på respondenter som er i den såkalte «prime working age» (OECD, 2014), derfor har vi valgt å ekskludere individer som er under 25 år. Dette gjorde vi for å unngå eventuelle skjevheter i gjennomsnittsverdiene.

Vårt utvalg inkluderer både deltids- og heltidsansatte, samt ansatte i både privat og offentlig sektor. Vi har imidlertid valgt å ekskludere de som arbeider i non-profit organisasjoner, som for eksempel veldedighets- eller religiøse organisasjoner. Dette fordi arbeidere i non-profit organisasjoner kan ha en lønnsstruktur som er forskjellig fra det som er typisk i for-profit sektoren. I noen tilfeller kan lønnen være betydelig lavere, enten fordi arbeidstakerne frivillig aksepterer lavere lønn på grunn av deres engasjement for organisasjonens formål, eller fordi disse organisasjonene har færre midler tilgjengelig for lønn.

Dette kan potensielt påvirke analysen vår på to måter. For det første kan det redusere gjennomsnittlig timelønn i datasettet, ettersom lønnene i non-profitt sektoren kan trekke ned gjennomsnittet. For det andre kan det påvirke den laveste verdien i datasettet, ettersom timelønningen i non-profitt kan være blant de laveste.

3.2 Deskriptive analyser

Vi har valgt å gjennomføre deskriptive analyser hvor formålet vårt er å identifisere lønnsforskjeller mellom kvinner og menn i Storbritannia og Norge. I den følgende analysen er timelønn vår avhengige variabel. Vi har også valgt å inkludere følgende uavhengige variabler i analysen: utdanning, arbeidserfaring, leseferdigheter, tallferdigheter, sektor og arbeidskontrakt. Vi har valgt å benytte oss av timelønn som er oppgitt i amerikanske dollar, ettersom timelønningen da er justert for kjøpekraftsparitet (purchasing power parity) noe som gjør det lettere å sammenligne timelønningen på tvers av landegrensene.

I våre deskriptive analyser har vi benyttet oss av de binære uavhengige variablene kvinne, Storbritannia, offentlig sektor og arbeidskontrakt. Det innebærer for eksempel at dersom respondenten svarer at de er en kvinne, har hun verdien 1, og dersom respondenten er en mann har han verdien 0. Den uavhengige variabelen utdanning er et mål på antall år fullført utdanning, mens yrkeserfaring er et mål på antall yrkesaktive år.

Den uavhengige variabelen leseferdigheter kan defineres som respondentens evne til å forstå, vurdere og benytte seg av en skrevet tekst for å kunne delta aktivt i samfunnet, oppnå sine egne mål, og utvikle sine egne evner og kunnskaper. (Bjørkeng, 2013) Mens den uavhengige variabelen tallforståelse innebærer respondentens evne til å kunne tolke, benytte seg av og formidle matematisk informasjon (Bjørkeng, 2013). Leseferdigheter og tallforståelse vil måles på en skala fra 0 til 500. (Bjørkeng, 2013). Ettersom det i vårt datasett var svært få britiske respondenter som har gitt et svar på spørsmålet innenfor IKT, velger vi å se bort fra denne variabelen i våre videre analyser.

Den uavhengige variabelen alder representerer den faktiske alderen til respondentene, både i det britiske og norske datasettet. Det innebærer at dersom respondenten er for eksempel 34 år, vil det stå 34 i datasettet.

3.2.1 Deskriptiv statistikk for avhengig variabel

	Alle		Kvinner		Menn	
	Norge	Storbritannia	Norge	Storbritannia	Norge	Storbritannia
Gjennomsnittlig timelønn	25.05	17.46	23.63	16.75	26.45	18.48
Minimum	14.88	8.55	14.88	8.55	14.88	8.55
Maksimum	45.06	39.71	45.06	39.71	44.95	39.71
Standardavvik	6.46	7.12	5.65	6.96	6.89	7.22
Antall observasjoner	2 762	3 884	1 372	2 285	1 390	1 599

Tabell 1: Deskriptiv statistikk for lønn

Tabell 1 er en oversikt over timelønn for både kvinner og menn i Norge og Storbritannia. Etter vi har fjernet respondentene som ikke var aktuelle for å gjennomføre analysen ender vi opp med 6 646 respondenter, hvor 2 762 er norske og 3 884 er britiske. Dette er etter en korrigerings av ekstremverdier. Vi valgte å fjerne de øverste 5% og de nederste 5% på lønnsvariabelen ettersom at det ikke var et tydelig skille på ekstremverdier og vi ikke fant gode data på mulige minimums- og maksverdier. Denne metoden, også kjent som "trimming", er en vanlig statistisk teknikk for å redusere effekten av ekstremverdier og potensielle feilobservasjoner i dataene (Rousseeuw & Hubert, 2011).

Ut ifra oversikten i tabell 1 ser vi at den gjennomsnittlige timelønnen i Norge ligger på 25.01 US dollar. Den gjennomsnittlige timelønnen til norske kvinner ligger på 23.63, noe som tilsvarer 89.34% av menns timelønn. At norske kvinner tjener 89.34% av det norske menn tjener stemmer ganske godt med en statistikk som ble publisert av Statistisk sentralbyrå (SSB) i 2012, som viste at kvinners lønn utgjorde 86.5 % av menns lønn (SSB, 2013).

En artikkel fra gov.uk som ble lagt ut i 2022 hevder at den gjennomsnittlige timelønnen i Storbritannia for alle arbeidere er 13.57 pund (Gov.uk, 2022). Dersom vi gjør dette om til dollar ved å bruke valutakursen som var i 2012 får vi 20.36 dollar i timen. Dette svaret er ca.

14% større enn det som står i tabellen 1. Ettersom at PIAAC-undersøkelsen består av 3 884 tilfeldige respondenter anser vi ikke denne forskjellen som betydelig.

Vi kan også se at den gjennomsnittlige timelønnen i Storbritannia er betydelig lavere enn den gjennomsnittlige timelønnen i Norge. Den tilsvarer faktisk bare 69.7% av den gjennomsnittlige norske timelønnen. Dette stemmer ikke overens med tidligere funn. I en undersøkelse gjort av OECD ser vi at den gjennomsnittlige britiske lønnen er 86.87% av den norske gjennomsnittlige lønnen (OECD, 2022). Videre bemerker vi oss at lønnsgapet mellom kvinner og menn i Storbritannia er lavere enn det er i Norge. Kvinnene i Storbritannia tjener i snitt 90.64% av det en mann tjener, noe som er litt høyere enn i Norge hvor det var 89.34%. Videre forteller standardavviket oss at vi vil finne den største lønnsspredningen hos menn, både i Norge og i Storbritannia.

National minimum wage forteller oss hva som skal være minstelønn per time i Storbritannia. En artikkel lagt ut av gov.uk i 2012 viser til at minstelønnen i Storbritannia i oktober 2011 skulle ligge på 6.08 britiske pund i timen (gov.uk, 2011). Når vi gjør om dette til amerikanske dollar ut ifra valutakursen i 2011, ser vi at minstelønnen skulle ligget på 9.64 dollar i timen. Dette er et høyere tall enn det vi får som minstelønn i vår tabell. Dette kan forklares ved at det er arbeidstakere i Storbritannia som har en lønn som er mindre enn national minimum wage. Dersom national minimum wage hadde holdt ville minstelønnen i Storbritannia vært 35.22% mindre enn i Norge istedenfor 42.54% slik som den er i vår analyse.

3.2.2 Deskriptiv statistikk for uavhengige variabler

I den deskriptive statistikken for de uavhengige variablene har vi inkludert de variablene vi anser som relevante for vår analyse og som vi antar har en effekt på timelønnen, og dermed kan være med på å forklare lønnsgapet. Årsaken til at vi valgte akkurat disse uavhengige variablene er på grunn av at vi mener de kan være nyttige for å svare på problemstillingen vår. I teoridelen nevnte vi at lønnsforskjellene påvirkes av faktorer som utdanning, arbeidserfaring og sektor. Vi ønsker derfor å teste denne teorien for å finne ut om det stemmer.

I tabellen under ser vi gjennomsnittsdataen for hvert land, og for kvinner og menn separat.

	Alle		Kvinner		Menn	
	Norge	Storbritannia	Norge	Storbritannia	Norge	Storbritannia
Antall år utdanning	15.95 (3.69)	16.41 (5.39)	16.37 (3.59)	16.69 (5.34)	15.53 (3.73)	16.01 (5.42)
Arbeidserfaring i år	21.49 (11.04)	21.59 (10.97)	20.56 (10.31)	20.70 (10.51)	22.41 (11.64)	22.87 (11.48)
Leseferdigheter	288.04 (43.76)	280.32 (44.87)	287.78 (41.53)	279.56 (43.57)	288.27 (45.88)	281.41 (46.66)
Tallforståelse	291.59 (48.79)	272.85 (48.67)	285.24 (45.14)	268.16 (47.01)	297.87 (51.38)	279.55 (50.22)
Offentlig sektor	41,71%	38.29%	65.71%	71.62%	34.29%	28.38%
Har arbeidskontrakt	98.26%	95.05%	97.45%	94.99%	99.05%	95.14%

Tabell 2: Deskriptiv statistikk for utdanning, arbeidserfaring, leseferdigheter, tallforståelse og offentlig sektor i Norge og Storbritannia

I tabellen over har vi sammenlignet Norge og Storbritannia innenfor de fem uavhengige variablene vi har valgt å se nærmere på i vår analyse; utdanning, arbeidserfaring, leseferdigheter, tallforståelse, offentlig sektor og arbeidskontrakt. Ut fra tabellen ser vi at det er forskjeller i utdanningsnivå i Norge og Storbritannia. Vi ser at både norske kvinner i gjennomsnitt er høyere utdannet enn norske menn. Det samme ser vi for tallene i Storbritannia, bare at forskjellen ikke er like stor.

I PIAAC-rapporten, som fokuserer på sammenhengen mellom utdanning og ferdigheter, fremheves det at utdanning generelt har en positiv effekt på ferdighetene. Vi ser at utdanningsnivået er høyere i Storbritannia enn i Norge, dette kan vi forklare ved at den obligatoriske delen av skolen er et år lenger i Storbritannia enn i Norge. I Norge er grunnskolen tiåring, mens i Storbritannia er grunnskolen elleveårig (Gov.uk, 2016). Vi observerer videre at det er mennene i begge landene som har lengst arbeidserfaring. Dette kan

forklares gjennom preferanseteorien ved at kvinner ofte prioriterer familie i større grad enn menn.

I PIAAC-rapporten står det at gjennomsnittet på leseferdigheter og tallforståelse i OECD-landene ligger på 273 og 269 poeng. Ut ifra vår deskriptive analyse ser vi at både Norge og Storbritannia har et høyere gjennomsnitt enn dette i både leseferdigheter og tallforståelse. Videre ser vi at mennene har et høyere gjennomsnitt enn kvinnene i begge landene. Denne forskjellen anser vi ikke som signifikant når den korrigeres for arbeidserfaring, utdanning og sektor.

Videre ser vi at det er 41.71% av de norske, og 38.29% av de britiske respondentene som jobber i offentlig sektor. Blant de norske respondentene står kvinnene for 65.71% av de ansatte i offentlig sektor. I en artikkel lagt ut av SSB står det at 70% av ansatte i offentlig sektor er kvinner (Fredriksen, 2019), noe som stemmer ganske bra med tallet vi fikk. I Storbritannia får vi at 71.62% av ansatte i offentlig sektor kvinner. Det er altså en større andel av respondentene som har svart at de jobber i offentlig sektor i begge landene, som er kvinner. Dette kan indikere at landene er horisontalt segregert, noe som kan bidra til å forklare lønnsgapet mellom kvinner og menn.

I forhold til arbeidskontrakt ser vi at det er vanlig å ha arbeidskontrakt i både Norge og Storbritannia. I Norge ser vi at 98.26% har arbeidskontrakt, mens i Storbritannia er gjennomsnittet 95.05%. Videre ser vi også at det er flere menn enn kvinner som har arbeidskontrakt, men at forskjellen mellom kjønnene ikke er stor.

3.3 Styrker og svakheter ved datasettet

PIAAC-undersøkelsen (Programme for the International Assessment of Adult Competencies) tilbyr flere fordeler i en bacheloroppgave. Datasettet inneholder et mangfold av variabler som kan bidra til å forklare faktoren timelønn, inkludert utdanning, kjønn, arbeidserfaring, kognitive ferdigheter og opphav. Det store utvalget av populasjonen øker sannsynligheten for at resultatene er representative for hele befolkningen og styrker datasettets validitet opp mot problemstillingen vår. Selv om noen faktorer ikke kan forklares og blir fanget opp av støy, bidrar det omfattende datasettet til at en større del av timelønnen kan forklares av variablene.

Videre gir datasettet muligheten for internasjonale sammenligninger, noe som kan øke forståelsen av hvordan ulike faktorer påvirker lønn og kompetanse i forskjellige land og kulturer.

På den andre siden inneholder PIAAC-datasettet også svakheter. For det første tar det ikke hensyn til familieforhold og bosted, som kan påvirke lønnsnivået og sammenhengen mellom utdanning og inntekt. Dette innebærer at faktorer som foreldres inntekts- og utdanningsnivå, som har vist seg å påvirke utdanningens avkastning, ikke blir inkludert i analysen. For det andre kan datasettet ha begrensninger knyttet til utdanningsvariabelen, ettersom datasettet ikke tar høyde for ulik utdanning, derfor blir marginaeffekten av ett års ekstra utdanning lik for alle individer. I realiteten kan ulike utdanningsbakgrunner med samme varighet gi ulik avkastning, og avkastningen av ett års ekstra utdanning kan variere mellom for eksempel master- og bachelorstudenter. Det kan også være at datasettet ikke fanger opp alle relevante faktorer som påvirker lønnsnivået, som regionale lønnsforskjeller og variasjoner i lønnsnivået mellom ulike sektorer og yrker.

4. Metode

I dette kapitlet vil vi presentere regresjonsmodellen som benyttes i vår bacheloroppgave og de tilhørende hypotesene vi har utarbeidet. Målet er å gi en grundig forståelse av metodene og prinsippene som ligger til grunn for analysen vår. Ved å gjennomgå modellen og hypotesene vil vi sikre en solid basis for tolkning av resultatene og en effektiv måte å evaluere de faktorene som påvirker lønnsnivået.

4.1 MLR- antagelser

Multipel lineær regresjon (MLR) blir brukt i oppgaven for å analysere lønnsforskjeller mellom kvinner og menn i Storbritannia og Norge. MLR er en statistisk teknikk som lar oss modellere sammenhengen mellom en avhengig variabel og flere uavhengige variabler. I vårt tilfelle er den avhengige variabelen lønn, mens de uavhengige variablene inkluderer utdanning, erfaring og kognitive ferdigheter som leseforståelse og tallforståelse. MLR gir en fleksibel og effektiv måte å estimere effekten av flere faktorer på lønnsforskjeller samtidig. I tillegg lar MLR oss teste interaksjonseffekter mellom variabler, noe som kan være viktig for å forstå hvordan de uavhengige variablene påvirker den avhengige variabelen.

For å sikre at MLR-analysen gir pålitelige og gyldige resultater, er det viktig å vurdere om de fem MLR-antagelsene er oppfylt for datasettet og analysen vår (Wooldridge 83–93). I denne delen blir det gjennomgått hver av MLR-antagelsene og forklart hvordan de gjelder for vår studie.

Den første MLR-antagelsen handler om linearitet, hvilket innebærer at vi forventer et lineært forhold mellom de uavhengige variablene (som utdanning og erfaring) og den avhengige variabelen (lønn). Vi forventer således at lønnsforskjellene mellom kvinner og menn vil øke lineært med endringer i de uavhengige variablene. I vårt tilfelle, selv om linearitet ikke er fullt ut bekreftet i våre modeller, velger vi likevel å validere dem. Dette er på grunn av vår overordnede modellmålsetning, som er å fremstille forholdet mellom den avhengige variabelen og de uavhengige variablene. Vi støtter denne beslutningen ved å vurdere R-squared (se vedlegg kap. 8.6.), F-tester i modeller som er nødvendig for ekstra validering (se kap. 5.4 og 5.6), P-verdier (T-tester i samtlige modeller) og konfidensintervaller (se vedlegg

kap. 8.6), som alle kan leses direkte fra vår STATA-modell. Disse indikatorene viser at, til tross for mangel på perfekt linearitet, er våre modeller statistisk signifikante og gir meningsfylt innsikt i variablenes innvirkning på lønn.

Homoskedastisitet utgjør den andre MLR antagelsen, her antar vi at variansen i feilene er konstant over alle nivå av de uavhengige variablene. Dette betyr at spredningen av feilene er lik for alle observasjoner, uavhengig av deres verdi på de uavhengige variablene. Dersom feilene var heteroskedastiske, kunne dette føre til ineffektive estimater og svake analysen. For å teste antagelsen om homoskedastisitet brukte vi Breusch-Pagan testen, som sjekker for heteroskedastisitet i feilene. Siden testen i vedlegg 8.5 ikke avslører betydelig heteroskedastisitet, gir dette støtte til antagelsen om homoskedastisitet.

Uavhengige feil er den tredje MLR antagelsen, her antar vi at feilene i vår MLR-modell er uavhengige av hverandre, det vil si at det ikke er noen systematisk sammenheng mellom feilene. Dersom feilene var korrelert, kan dette føre til skjevheter i estimatene og redusere nøyaktigheten av analysen. For å teste antagelsen om uavhengige feil lager vi en stegvis test hvor vi sjekker residualene i modellen vår, dette måler graden av autokorrelasjon mellom feilene. Siden testen i vedlegg 8.3 indikerer at autokorrelasjonen er lav, styrker dette antagelsen om uavhengige feil.

Normalitet er den fjerde antagelsen, her antar vi at feilene i MLR-modellen vår er normalfordelte. Normalfordeling av feilene er viktig for å sikre at MLR-estimatene er upartiske og effektive, og at vi kan stole statistiske slutninger, som t-verdier og konfidensintervaller. For å sjekke antagelsen om normalitet gjennomførte vi normalitetstesten kvantil-kvantil (Q-Q) plott. Siden testen i vedlegg 8.3 ikke avslører ett betydelig avvik fra normalfordelingen, gir dette støtte til antagelsen om normalitet.

Ingen multikollinearitet er den femte antagelsen, den antar at de uavhengige variablene i vår MLR-modell ikke er for sterkt korrelerte med hverandre. Multikollinearitet kan føre til ustabile og upålitelige estimater, og gjøre det vanskelig å tolke effekten av individuelle uavhengige variabler på lønnsforskjellene. For å sjekke antagelsen om ingen multikollinearitet beregnet vi i vedlegg 8.4 Variance Inflation Factor (VIF) for hver uavhengige variabel i modellen. VIF-verdiene er lavere eller på den forhåndsbestemt terskel som er 10, dette støtte til antagelsen om ingen multikollinearitet. Unntaket er resultatet til D som er langt over den satte terskelen, dette kan tyde på at modellen har variabler som ikke er statistisk signifikante.

4.2 Gjennomgang av teori og metode

For å undersøke faktorene som påvirker lønnsnivået, er det nødvendig å etablere en lønningsformel. Denne formelen vil inneholde variablene vi anser som relevante for analysen vår, basert på problemstillingen.

Den forenklete lønningsformelen (1) viser at timebasert lønn er en funksjon av kjønn, og andre kontrollvariabler representeres av x . Variablene i x antas å representere menneskelig kapital og arbeidsrelaterte faktorer som utdanning, ferdigheter, arbeidserfaring og sektor. Vi forventer at disse elementene vil ha en positiv innvirkning på timelønnen ved å øke arbeidstakerens produktivitet.

$$(1) \text{timelønn} = f(\text{female}, x)$$

For å estimere en lønningsformel med dataene våre, er det nødvendig å spesifisere likning (1). En modell som ofte brukes av økonomer er Mincer's lønnsmodell, uttrykt ved likning (2). I denne modellen betraktes logaritmen til timelønnen ($\ln Y$) for individ i som en funksjon av utdanning ($X1$) og arbeidserfaring ($X2$) for det samme individet. Arbeidserfaring er uttrykt både lineært og kvadratisk ($X3$) for å ta hensyn til mulige ikke-lineære effekter av arbeidserfaring.

$$(2) \ln Y_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 (X_{i3})^2 + \varepsilon$$

Mincers lønnsmodell er relevant for analysen vår fordi den gir en praktisk og veletablert metode for å undersøke sammenhengen mellom lønn, utdanning og arbeidserfaring. Ved å bruke denne modellen kan vi estimere effekten av de uavhengige variablene (utdanning og arbeidserfaring) på den avhengige variabelen (timelønn) og dermed kvantifisere hvordan menneskelig kapital og arbeidsrelaterte faktorer påvirker lønnsnivået. Den logaritmisk-lineære funksjonsformen er valgt fordi den gir best tilpasning til vårt datasett, PIAAC. Vi vurderte dette ved å sammenligne modeller med ulike funksjonsformer og deres evne til å beskrive datasettet (se vedlegg kap. 8.1). Logaritmisk-lineære modeller er mye brukt i forskning, og tidligere studier har vist at denne funksjonsformen ofte gir bedre tilpasning (Heckman et al., 2006). Videre tillater logaritmisk-lineære formuleringer å sammenligne estimerte lønnsforskjeller mellom land, ettersom den estimerte lønnsforskjellen kan tolkes som en omtrentlig prosentvis lønnsforskjell.

Logaritmisk-lineære modeller er valgt for analysen av lønnsnivået på grunn av deres evne til å gi bedre tilpasning til datasettet og enkel tolkning av koeffisienter. Disse modellene forvandler ikke-lineære sammenhenger til lineære, noe som gjør det lettere å analysere og tolke resultatene. De tillater også sammenligning av estimerte lønnsforskjeller mellom ulike grupper og håndterer skjevhet i dataene. Dette gjør dem godt egnet for å studere forholdet mellom lønn, utdanning og arbeidserfaring.

Populasjonens regresjonslinje ($Y = \alpha + \beta X + \varepsilon$) er en ukjent linje som vi ønsker å finne det beste estimatet for. Ved hjelp av datasettet estimerer og kvantifiserer vi parameterne α og β , og estimerer en predikert regresjonslikning for utvalget vårt ($\hat{Y} = a + bX$). På grunn av utvalgsvariasjon vil ikke utvalgets regresjonslinje nødvendigvis sammenfalle med populasjonens regresjonslinje. I populasjonens regresjonslinje har den avhengige variabelen en lineær sammenheng med uavhengige variabler, og vi anser det som troverdig at forutsetningene for bruk av minste kvadraters metode (MKM) er oppfylt. Teorien bak MKM er å minimere summen av de kvadrerte avvikene mellom den faktiske verdien, Y , og den predikerte (observerte) verdien, \hat{Y} . Ved å minimere summen av de kvadrerte avvikene finner vi utvalgsparameterne a og b . Her er Y den faktiske verdien av Y for observasjon i , og \hat{Y} er den predikerte verdien av Y for observasjon i , basert på estimatene for a og b .

$$\min \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n (Y_i - a - bX_i)^2$$

Avviket mellom den faktiske timelønnen og predikerte timelønnen for observasjon i kalles for residualen. Det vil være den delen av lønnen som ikke blir forklart av vår modell.

Forklaringsfaktoren avhenger av hvilken funksjonsform vi bruker.

Minste kvadraters metode (MKM) er en standard teknikk innen statistikk og økonometri som brukes for å estimere koeffisientene i en regresjonsmodell. MKM fungerer ved å minimere summen av de kvadrerte avvikene mellom de faktiske verdiene (Y) og de predikerte verdiene (\hat{Y}) basert på regresjonsmodellen. Dette resulterer i en estimert regresjonslinje som gir den beste tilpasningen til datasettet.

Forutsetningene bak MKM inkluderer en antagelse om at det er en lineær sammenheng mellom avhengige og uavhengige variabler. Videre antas det at de uavhengige variablene ikke er stokastiske, det vil si at verdiene for X_{i1} og X_{i2} er konstante i hvert utvalg. Det antas også at det ikke er noen korrelasjon mellom de uavhengige variablene, noe som ville gjøre det

umulig å estimere koeffisientene. Feilleddet (støyleddet) antas å være stokastisk og ha et gjennomsnitt på null, mens kovariansen mellom feilleddet og de uavhengige variablene må være lik null. Det antas at det ikke er noen korrelasjon mellom feilleddene (støyleddene) for forskjellige observasjoner. Til slutt antas det at feilleddene (støyleddene) er normalfordelte med konstant varians, det vil si at det ikke er heteroskedasitet. Heteroskedasitet er et vanlig problem i tverrsnittsdata, så vi vil bruke Mincers lønnsprognosemodell i Stata og deretter teste for heteroskedasitet for å kunne justere for det senere.

Ved å bruke MKM estimeres koeffisientene i regresjonsmodellen, og det blir mulig å teste hypoteser om disse koeffisientene og deres signifikans. Dette gir grunnlaget for å analysere årsakssammenhenger og evaluere effektene av de uavhengige variablene på den avhengige variabelen.

5. Regresjonsanalyse

Formålet med denne studien er å undersøke kjønnsbaserte lønnsforskjeller i Storbritannia og Norge. Gjennom en beskrivende analyse identifiserte vi et lønnsgap mellom kvinner og menn i begge land, med det største gapet i Storbritannia. Videre vil vi utforske faktorene som påvirker timebasert lønn ved hjelp av regresjonsanalyser og vurdere om det er noen forskjeller i disse faktorene mellom de to landene. Deretter vil vi se på om det er et lønnsgap mellom ansatte med og uten arbeidskontrakt, undersøke årsakene til dette gapet, og evaluere hvordan det påvirker lønningene til kvinner og menn. Vi tar utgangspunkt i Mincer-modellen som beskrevet tidligere, og vil estimere seks varianter basert på Mincer-likningen. For å gjøre ligningene mer lesbare, velger vi å ekskludere repetitive variabler i modellene. Tar utgangspunkt i at alle variablene i modell A også inngår i modellene B-F.

$$A. \lnwage = \alpha + \delta_0 fem + \beta_1 edu + \beta_2 exp + \beta_3 exp^2 + \beta_4 num + \beta_5 lit + \varepsilon$$

$$B. \lnwage = (A) + \delta_1 GB + \rho_0(fem * gb) + \varepsilon$$

$$C. \lnwage = (A) + \delta_1 gb + \rho_0(fem * gb) + \rho_1(edu * gb) + \rho_2(edu * fem) + \rho_3(fem * edu * gb) + \varepsilon$$

$$D. \lnwage = (A) + \delta_1 gb + \rho_0(fem * gb) + \rho_4(fem * num) + \rho_5(fem * lit) + \rho_6(gb * num) + \rho_7(gb * lit) + \rho_8(fem * gb * num) + \rho_9(fem * gb * lit) + \varepsilon$$

$$E. \lnwage = (A) + \delta_1 gb + \delta_2 pub + \rho_0(fe * gb) + \gamma_0(pub * fem) + \gamma_1(pub * gb) + \gamma_2(pb * fem * gb) + \varepsilon$$

$$F. \lnwage = (A) + \delta_3 con + \delta_4 (gb * con) + \delta_5 (fem * con) + \delta_6 (gb * fem * con) + \varepsilon$$

T-tester har blitt utført for å undersøke om det observerte lønnsgapet mellom kvinner og menn er statistisk signifikant. Nullhypotesen antar at timebasert lønn for kvinner er den samme som for menn. Dersom t-verdien overgår den kritiske verdien, avvises nullhypotesen, og vi kan konkludere med at det finnes et kjønnsbasert lønnsgap.

Vi har estimert et lønnsgap mellom kvinner og menn i modellene A - F som varierer mellom 7.6% og 11.5 %, og er statistisk signifikante (se vedlegg kap. 8.6). Lønnsgapet er lavest i modell C, når vi tester for avkastning i utdanning, og størst i modell E der vi tester for om jobbsektor har betydning for lønnsgapet. I modell F tester vi for arbeidskontrakt, og ser at forskjellen i lønnsgapet er størst i Storbritannia, hvor de som jobber uten kontrakt har vesentlig lavere lønn sammenlignet med Norge.

5.1 Modell A - Lønnsgap mellom kvinner og menn i Norge og Storbritannia

I denne analysen antar vi at timebasert lønn for norske og britiske arbeidere er den samme. Vi vil benytte denne modellen for å undersøke om det finnes et lønnsgap mellom kjønnene, noe som vil være representert ved en prosentvis forskjell i gjennomsnittlig lønn. Videre vil vi bruke modellen for å fortolke den prosentvise endringen i timelønn som et resultat av ett ekstra års utdanning og arbeidserfaring. Kognitive ferdigheter representerer den prosentvise lønnsøkningen, som tilsvarer ett standardavviks økning i testresultater.

Vi begynner med å undersøke δ_0 for å fastslå om det finnes et lønnsgap mellom kjønnene. Vi oppretter en nullhypotese som sier at timelønnen til kvinner og menn er det samme. $H_0: \delta_0 = 0$. Alternativhypotesen er at kvinners lønnsnivå er lavere enn menns. $H_A: \delta_0 < 0$. Vi konstruerer en testobservator (TS) for å bestemme om vi kan beholde nullhypotesen eller ikke.

$$TS: \frac{b_0 - \beta}{s_b} \sim t(n - k) = \frac{-0.1098 - 0}{0.0085} \sim t^{0.05, 6635} = -12.92 < -1.645$$

Her ser vi at testobservatoren er lavere enn kritisk verdi ved 5% signifikansnivå, og vi kan forkaste nullhypotesen. Dette tyder på at kvinners timelønn er lavere enn menns i begge landene. Den estimerte timelønnen for kvinner er tilnærmet 11% lavere enn for menn, noe som samsvarer med funnene i kapittel 3. Videre vil vi undersøke koeffisienten β_1 , som representerer den prosentvise endringen i timelønn som et resultat av ett ekstra utdanningsår. Ifølge tabellen fører ett ekstra år med utdanning til en 2.2% økning i timelønn, og dette resultatet er statistisk signifikant ved 5% signifikansnivå. Dermed kan vi konkludere med at det finnes lønnsinsentiver knyttet til utdanning.

Videre ser vi at ett ekstra års arbeidserfaring, representert ved β_2 , gir en statistisk signifikant positiv effekt på timelønnen med 1.75%. Vi inkluderer arbeidserfaring i kvadrert form for å undersøke om arbeidserfaring har en lineær effekt. Vi etablerer en hypotesetest der $H_0: \beta_3 = 0$ og $H_A: \beta_3 < 0$ for å vurdere om arbeidserfaring har en lineær eller kvadratisk effekt på timelønnen. Med et 5% signifikansnivå aksepterer vi alternativhypotesen om at arbeidserfaring har en avtagende effekt på timelønnen.

Ferdighetene for tallforståelse indikerer at timelønnen øker med 0.3%, der økt leseforståelse gir en negativ effekt på timelønnen med 0.05%, dersom verdien på ferdighetene øker med et standardavvik. Alene kan vi ikke konkludere med at ferdighetene har noe signifikant effekt på timelønnen. Vi vil derfor ved hjelp av en testobservator avgjøre om ferdighetene samlet sett har en positiv og signifikant effekt på timelønnen. Testobservatoren inkluderer en modell uten restriksjoner (uten ferdighetene) og en modell med restriksjoner (med ferdighetene).

$$TS = \frac{\frac{(R_U^2 - R_R^2)}{h}}{\frac{(1 - R_U^2)}{(n - k)}} = \frac{\frac{(0.2741 - 0.1788)}{3}}{\frac{(1 - 0.2741)}{(6642 - 3)}} = 290.53 > 2.6059$$

Testobservatoren overstiger kritisk verdi i t-fordelingen ved 5% signifikansnivå, og vi konkluderer derfor med at de kognitive ferdighetene har en signifikant effekt på timelønnen.

Ved hjelp av modellen finner vi den predikerte timelønnen for kvinner og menn, gitt at verdiene for utdanning, arbeidserfaring og ferdigheter er lik null.

For menn:

$$\ln wage = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_3)^2 + \beta_4 X_3 + \beta_5 X_4$$

$$\ln wage = 1.88 + 0.0223X_1 + 0.0175X_2 + (-0.00024)X^2 + (-0.0006)X_3 + 0.0030X_4$$

For kvinner:

$$\ln wage = (\alpha + \delta_0) + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_2)^2 + \beta_4 X_3 + \beta_5 X_4$$

$$\ln wage = (1.88 - 0.11) + 0.0223X_1 + 0.0175X_2 + (-0.00024)X^2 + (-0.0006)X_3 + 0.0030X_4$$

Vi beregner den forventede timelønnen for kvinner ved å opphøye den naturlige logaritmen e i 1.77, og for menn i 1.88. Dette gir oss en forventet timelønn på 5.87 amerikanske dollar for kvinner og 6.55 for menn. I vår analyse, hvor timelønnen er justert for kjøpekraftsparitet og uttrykt i amerikanske dollar, mener vi det er mer hensiktsmessig å fokusere på koeffisientene i stedet for den forventede timelønnen.

5.2 Modell B - Forskjell i lønnsnivå i Storbritannia og Norge

I denne analysen utforsker vi om lønnsforskjellen kan forklares av at arbeidstakeren er fra Storbritannia, og om lønnsgapet mellom kvinner og menn er mindre i Norge. I kapittel 3 observerte vi en lønnsforskjell mellom Storbritannia og Norge, hvor lønnsgapet var høyere i Norge. Vi utvider modellen ved å inkludere en binær variabel som indikerer om individet er britisk, samt en interaksjonsvariabel som viser om individet er både britisk og kvinne. Først undersøker vi om det å være britisk har en signifikant effekt på lønnen ved å se på δ_1 , hvor vi tester om det er et lønnsgap mellom landene. Vi oppretter en nullhypotese der lønnsnivået i landene er likt, $\delta_1 = 0$, mot en alternativ hypotese der $\delta_1 < 0$. Nullhypotesen kan forkastes på 1% signifikansnivå, og vi kan konkludere med at det å være britisk har en negativ effekt på timelønnen (se vedlegg kap. 8.6). I Storbritannia er timelønnen 38.7 % lavere sammenlignet med Norge, noe som samsvarer med funnene i kapittel 3.

Koeffisienten ρ_0 indikerer at lønnsgapet mellom kvinner og menn er 0.9% større i Storbritannia sammenlignet med Norge. Siden vi bruker et 5% signifikansnivå, kan vi ikke konkludere med at lønnsgapet mellom kjønnene er statistisk ulikt mellom Storbritannia og Norge. Lønnsforskjellene mellom kvinner og menn er omtrent de samme som i modell (A), noe som antyder at det kjønnsbaserte lønnsgapet i liten grad kan forklares av om individet er britisk.

5.3 Modell C - Avkastning på utdanning for kjønn i Storbritannia og Norge

I denne modellen vil vi undersøke om lønnsgapet kan forklares av systematiske forskjeller i utdanningsnivå mellom kjønnene. I den deskriptive analysen observerte vi at den britiske befolkningen i gjennomsnitt har en litt høyere utdanning enn den norske, og at britiske kvinner i gjennomsnitt har høyere utdanning enn britiske menn. Vi vil utforske om avkastningen fra ett års ekstra utdanning varierer mellom kvinner og menn, og om dette kan bidra til å forklare noe av lønnsgapet mellom kjønnene.

Først vil vi undersøke om utdanningsnivået har en signifikant effekt på lønnen. Dette gjøres ved å teste $H_0: \beta_1 = 0$ mot $H_A: \beta_1 \neq 0$ ved hjelp av en t-test. Testobservatoren overstiger den kritiske verdien, og vi avviser dermed nullhypotesen om at utdanningsnivået ikke påvirker lønnen. Avkastningen på utdanning i modell (C) er omtrent 100×0.0236 , noe som betyr at ett ekstra års utdanning gir omtrent 2.36% høyere lønn.

Videre skal vi undersøke om den marginale effekten av utdanning varierer for kvinner og menn. Koeffisienten ρ_2 indikerer at avkastningen på ett ekstra års utdanning er 0.26% lavere for kvinner enn for menn. Siden denne koeffisienten ikke er statistisk signifikant ved 5% signifikansnivå, kan vi ikke med sikkerhet si at den har noe effekt på timelønnen.

Koeffisienten ρ_3 viser at avkastningen av ett ekstra års utdanning er 1.1% høyere for britiske kvinner enn for norske kvinner, og denne er statistisk signifikant.

I modell C ser vi at det estimerte lønnsgapet mellom kvinner og menn har redusert seg til 7.66%. Dette indikerer at selv om vi har tatt hensyn til forskjeller i utdanning mellom kjønnene, kan ikke lønnsgapet fullstendig forklares av disse systematiske forskjellene i formell utdanning. Det er fortsatt en betydelig lønnsforskjell mellom kjønnene som ikke kan forklares ut ifra utdanningsnivået alene. I den deskriptive analysen observerte vi at kvinner i

Norge i gjennomsnitt har høyere utdanning enn menn. Dette indikerer at selv om kvinner oppnår høyere utdanningsnivåer, oversettes ikke dette nødvendigvis til en tilsvarende økning i lønn, noe som bidrar til det observerte lønnsgapet.

5.4 Modell D - Avkastning på ferdigheter for kjønn i Storbritannia og Norge

I denne modellen vil vi utforske om lønnsgapet kan tilskrives systematiske forskjeller i ferdigheter. I den deskriptive analysen ser vi at Norge skårer høyere enn Storbritannia i både lese- og skriveferdigheter, hvor norske menn skårer høyere enn norske kvinner. Vi ønsker derfor å se om avkastningen på ferdigheter er mindre for kvinner sammenlignet med menn, og om avkastningen i Storbritannia er lavere enn i Norge. Vi benytter standardiserte testskårene for å lette sammenligningen av koeffisienter på tvers av land. Dette betyr at tolkningen av koeffisientene blir den prosentvise lønnsøkningen for hver økning på ett standardavvik i testverdi.

Det ser ut til at kognitive ferdigheter påvirker lønnsnivået, der tallforståelse fører til en økning i lønn med 0.13% og leseforståelse gir en ubetydelig reduksjon i lønn på 0.023%. Imidlertid viser analysen at koeffisienten for leseforståelse ikke er statistisk signifikant. Vi utfører en F-test for å undersøke nullhypotesen $H_0: \beta_4 = \beta_5 = 0$ i modell D mot en alternativ hypotese om at minst en av koeffisientene β_4 eller β_5 er ulik null.

$$F = [(SS_R - SS_{UR}) / (df_R - df_{UR})] / (SS_{UR} / df_{UR})$$

$$F = [(508.93587 - 513.623258) / (6627 - 6629)] / (513.623258 / 6629)$$

$$F = [(-4.687388) / (-2)] / (0.077481258)$$

$$F = (2.343694) / (0.077481258)$$

$$F = 30.2366$$

$$\text{Kritisk nivå} = 1.827$$

Testobservatoren overstiger den kritiske verdien i F-fordelingen. Derfor kan vi avvise nullhypotesen om at kognitive ferdigheter er lik null, noe som betyr at minst en av de kognitive ferdighetene har en signifikant effekt på timelønnen.

Da de estimerte parameterne p_4 , p_5 , p_6 , p_8 , og p_9 ikke er statistisk signifikante, kan vi ikke fastslå at marginalavkastningen på kognitive ferdigheter varierer etter kjønn eller mellom Norge og Storbritannia. Vi observerer at leseferdigheter gir en svært liten lønnsavkastning i Storbritannia.

Vi konkluderer med at avkastningen av ferdigheter er lik for både menn og kvinner, uavhengig av om individet er fra Norge eller Storbritannia. Videre observerer vi at lønnsforskjellen mellom kvinner og menn har blitt redusert til 3.34%. Imidlertid er denne koeffisienten ikke statistisk signifikant, noe som betyr at vi ikke kan være sikre på at det faktisk er en reell forskjell.

Vår analyse inkluderer mange interaksjonsledd i modellen. Disse kunne teoretisk sett ha tillatt effekten av individuelle variabler å variere, avhengig av verdien på andre variabler. Imidlertid viste det seg at de estimerte koeffisientene for disse interaksjonsleddene ikke var statistisk signifikante. Det betyr at vi ikke fant noen bevis for at effekten av kognitive ferdigheter på lønn varierer i henhold til kjønn eller land.

Sammenfattet kan lønnsgapet mellom kvinner og menn i liten grad forklares av systematiske forskjeller i ferdigheter, selv om ferdigheter i seg selv ser ut til å ha en statistisk signifikant effekt på lønnen. Konklusjonen er at lønnslikningen er lik modell (B) for alle individer, uavhengig av kjønn og nasjonalitet

5.5 Modell E – Sammenligning av lønnsgap i offentlig og privat sektor for kjønn i Storbritannia og Norge

I denne delen ønsker vi å undersøke om jobbsektoren påvirker lønnen og om lønnsforskjellene mellom kvinner og menn varierer mellom de to sektorene. Til slutt skal vi sammenligne lønnsgapet mellom privat og offentlig sektor i Storbritannia med Norge og se om det er noen forskjeller i størrelsen på lønnsgapet.

Vi utvider basismodellen ved å legge til en binær variabel for offentlig sektor, som får verdien 1 dersom individet jobber i offentlig sektor og 0 dersom individet jobber i privat sektor.

Deretter lager vi interaksjonsvariabler mellom kjønn, Storbritannia og britiske kvinner. Dette lar oss utforske lønnsgapet mellom kjønn og sektorer i Norge og Storbritannia, samt variasjonene i lønnsgapet mellom landene.

Lønnsforskjellen mellom offentlig og privat sektor i Norge er representert ved δ_2 , mens γ_1 viser om lønnsforskjellen mellom disse sektorene er den samme i Storbritannia. For å undersøke om lønnsgapet varierer mellom kjønn i offentlig og privat sektor i Norge, ser vi på γ_0 . Deretter analyserer vi γ_2 for å finne ut om lønnsforskjellen mellom kjønn i offentlig sektor i Storbritannia skiller seg fra situasjonen i Norge.

Estimert δ_2 er -0.997, noe som betyr at timelønnen i offentlig sektor er nesten 10% lavere enn i privat sektor. Nullhypotesen om at $\delta_2 = 0$ testes mot alternativhypotesen $\delta_2 < 0$ og vi forkaster nullhypotesen på 1% signifikansnivå. Vi konkluderer dermed med at individer i offentlig sektor tjener mindre sammenlignet med de som jobber i privat sektor.

Etter å ha identifisert en lønnsforskjell mellom offentlig og privat sektor, undersøker vi om lønnsgapet mellom kjønnene varierer i de to sektorene. Koeffisienten γ_0 indikerer at lønnsgapet mellom kjønnene er 0.067, det vil si at kvinner som arbeider i offentlig sektor tjener 6.7% mer enn menn. Vi setter opp en nullhypotese om at lønnsgapet er likt i begge sektorer, mot en alternativhypotese om at lønnen er høyere i offentlig sektor. Med et signifikansnivå på 5% konkluderer vi med at lønnsgapet mellom kvinner og menn i Norge, er 6.7% høyere i offentlig sektor enn privat sektor.

Når vi analyserer hvordan lønnsgapet mellom offentlig og privat sektor er i Storbritannia, studerer vi koeffisienten γ_1 . Den forteller oss at individer i offentlig sektor i Storbritannia tjener 16.2% mer enn ansatte i privat sektor, og koeffisienten er statistisk signifikant på et 5% nivå.

Til slutt undersøker vi om lønnsgapet mellom kjønnene er større eller mindre i offentlig sektor i Storbritannia sammenlignet med Norge. Koeffisienten γ_2 indikerer at lønnsgapet mellom kvinner og menn er 1.1% lavere i Storbritannia enn i Norge. Til 5% signifikansnivå kan vi ikke konkludere med at lønnsgapet mellom kvinner og menn i offentlig sektor er mindre i Storbritannia.

5.6 Modell F - Forskjell i lønnsgap med og uten arbeidskontrakt i Storbritannia og Norge

I denne modellen skal vi undersøke om lønnsgapet kan skyldes om individet jobber med eller uten arbeidskontrakt. Vi tar utgangspunkt i basismodellen vår og introduserer en ny binær variabel for arbeidskontrakt. Den får verdien 1 dersom individet jobber uten arbeidskontrakt og 0 dersom individet jobber med arbeidskontrakt. I tillegg lager vi interaksjonsvariabler mellom Storbritannia og arbeidskontrakt, kvinner og arbeidskontrakt og med britiske kvinner og arbeidskontrakt.

Koeffisienten δ_3 er på 0.1839 i vår modell, hvilket betyr at individer, uavhengig av kjønn og land, tjener i snitt 18.4% mer enn individer med kontrakt. Med 5% signifikansnivå kan vi konkludere at denne koeffisienten er statistisk signifikant.

Lønnsgap mellom arbeidstakere som jobber i Storbritannia uten arbeidskontrakt er representert ved koeffisienten δ_4 . Den sier at individer som jobber uten arbeidskontrakt i Storbritannia tjener 53.9% mindre enn norske arbeidstakere uten arbeidskontrakt. Med et 5% signifikansnivå kan vi konkludere at denne koeffisienten også er statistisk signifikant.

Når vi analyserer koeffisienten δ_5 ser vi at kvinner, uavhengig av land, som jobber uten arbeidskontrakt tjener 2.6% mer enn menn. Midlertidig er denne koeffisienten ikke statistisk signifikant, noe som tyder på at denne sammenhengen ikke gir noen effekt på timelønnen. Ved å analysere koeffisienten δ_6 ser vi at lønnsgapet mellom kjønnene i Storbritannia er 8.9% høyere enn i Norge. Men til 5% signifikansnivå kan vi heller ikke konkludere at dette er riktig.

Det ser ut til at arbeidskontrakt påvirker timelønnen, men koeffisientene for kvinner uten arbeidskontrakt og britiske kvinner uten arbeidskontrakt ikke er statistisk signifikante. Vi utfører en F-test for å undersøke nullhypotesen $H_0: \delta_5 = \delta_6 = 0$ i modell F mot en alternativ hypotese om at minst en av koeffisientene er ulik null.

$$F = [(717.589047 - 718.085467) / (6631 - 6633)] / (718.085467 / 6633)$$

$$F = [(-0.496420) / (-2)] / (0.108259531)$$

$$F = [0.248210] / (0.108259531)$$

$$F = 2.291559$$

Kritisk nivå: 3.002

Testobservatoren overstiger ikke den kritiske verdien i F-fordelingen, og vi kan derfor ikke avvise nullhypotesen om at alle koeffisientene for arbeidskontrakt er lik null. Dette betyr at vi ikke har tilstrekkelig bevis for å konkludere med at minst én av koeffisientene for arbeidskontrakt har en signifikant effekt på lønnen.

Vi kan heller ikke konkludere med at lønnen for arbeidstakere uten kontrakt er høyere for menn enn kvinner, siden parameterne δ_5 og δ_6 ikke er statistisk signifikante.

Videre, ved å analysere våre modellparametre, ser vi at lønnsforskjellen mellom kvinner og menn er 11.9%. Denne koeffisienten er statistisk signifikant, noe som indikerer at vi med en høy grad av sikkerhet kan si at det er en reell forskjell i lønn mellom kjønnene. Faktisk gir resultatene fra vår modell en indikasjon på at lønnsgapet mellom kvinner og menn i en viss grad kan forklares av systematiske forskjeller i arbeidskontrakter.

Imidlertid, selv om lønnsforskjellen mellom kjønnene er statistisk signifikant, er det viktig å merke seg at typen arbeidskontrakt i seg selv ikke ser ut til å ha en statistisk signifikant effekt på lønnsgapet mellom kvinner og menn i Storbritannia og Norge. Dette kan vi se ved å observere at de estimerte parameterne for koeffisientene δ_5 og δ_6 , som representerer lønnsforskjellen for arbeidstakere uten kontrakt mellom menn og kvinner, og mellom Norge og Storbritannia, ikke er statistisk signifikante. Derfor kan vi ikke trekke en konklusjon om at lønnen for arbeidstakere uten kontrakt er høyere for menn enn kvinner, eller at lønnsgapet er høyere i Norge sammenlignet med Storbritannia.

6. Konklusjon

Estimatene fra lønnsmodellene bekrefter et lønnsgap mellom kvinner og menn i både Storbritannia og Norge, med en forskjell som varierer mellom 7.6% og 11.5 % avhengig av hvilke faktorer vi kontrollerer for (se vedlegg kap. 8.6). Funnene tyder på at lønnsgapet i liten grad kan tilskrives systematiske forskjeller mellom kjønnene når det gjelder arbeidserfaring, utdanningsnivå eller evaluerte kognitive evner.

Resultatene indikerer at en del av lønnsgapet i det norske arbeidsmarkedet kan forklares av det faktum at kvinner er overrepresentert i offentlig sektor sammenlignet med menn. I den deskriptive analysen ble det observert at 65.71% av de sysselsatte i offentlig sektor i Norge er kvinner. Regresjonsanalysen estimerer en signifikant negativ effekt på timelønnen på 10% for de som arbeider i offentlig sektor. Regresjonsresultatene antyder at lønnsgapet mellom kjønnene er mindre i offentlig sektor enn i privat sektor for begge land. Når det gjelder arbeidstakere med eller uten arbeidskontrakt, merket vi at det var en økning i timelønn på 18.9% for de uten arbeidskontrakt. Imidlertid var koeffisientene for kvinner uten arbeidskontrakt og britiske kvinner uten arbeidskontrakt ikke statistisk signifikante. Det gjør at vi ikke kan trekke en klar konklusjon på at det er lønnsforskjeller basert på arbeidskontrakt. Videre kan vi ikke trekke en konklusjon om at lønnsgapet er lavere i Storbritannia sammenlignet med Norge. Ut fra de oppnådde resultatene synes det å være en mangel på systematisk variasjon i lønnsgapet basert på utdanningsnivå eller målte kognitive ferdigheter, men det er en betydelig variasjon mellom sektorene.

Denne bacheloroppgaven undersøker et begrenset utvalg av variabler, noe som reflekteres i forklaringsgraden til de ulike regresjonsmodellene. I den første modellen er forklaringsgraden relativt lav, med en verdi på 28.1%. Denne øker til 50.7 % i modell E, for deretter å reduseres til 29.8 % i modell F. For å forbedre forklaringsgraden og dermed øke forståelsen av lønnsgapet mellom kjønnene, kunne det være relevant å inkludere ytterligere variabler i analysen. Eksempler på slike variabler inkluderer bostedsforhold, stillingsnivå, husholdsansvar, sivilstatus, antall barn, bransje og yrkeskategori. Inkludering av disse variablene kan bidra til en mer omfattende og robust analyse av lønnsgapet mellom kvinner og menn.

7. Litteraturliste

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, 70(5, Part 2), 9-49. Hentet 23.3.23 fra:
https://www.jstor.org/stable/pdf/1829103.pdf?refreqid=excelsior%3A892502fe8b112570bf510a03006991df&ab_segments=&origin=&initiator=&acceptTC=1

Bjørkeng, B. (2013): Ferdigheter i voksenbefolkningen. Resultater fra den internasjonale undersøkelsen om lese- og tallforståelse (PIAAC). Rapport 42/2013. Statistisk Sentralbyrå.

Booth, A. L., & Van Ours, J. C. (2008). Job Satisfaction and Family Happiness: The Part-Time Work Puzzle. *The Economic Journal*, Volume 118, Issue 526, 1 February 2008, Pages F77–F99, Hentet 15.4.23 fra: <https://doi.org/10.1111/j.1468-0297.2007.02117.x>

Bye, K. S. «Lønn, 2012» SSB, 2013. Hentet 4.4.2023 fra: <https://www.ssb.no/arbeid-og-lonn/statistikker/lonnansatt/aar/2013-03-20>

England, P. (2010). The Gender Revolution: Uneven and Stalled. *Gender & Society*, 24(2), 149-166. Hentet 14.4.23 fra:
<https://journals.sagepub.com/doi/epdf/10.1177/0891243210361475>

Fredriksen, K (2019). *På stedet hvil: kvinner dominerer offentlig sektor, menn privat sektor.* (SSB) Hentet 12.5.2023 fra: <https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/pa-stedet-hvil-kvinner-dominerer-offentlig-sektor-og-menn-privat>

Goldin, C. (2014). A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104(4), 1091–1119. Hentet 30.3.23 <https://doi.org/10.1257/aer.104.4.1091>

Gov.uk (2022) *Average hourly pay*. Hentet 8.5.2023 fra: <https://www.ethnicity-facts-figures.service.gov.uk/work-pay-and-benefits/pay-and-income/average-hourly-pay/latest>

Gov.uk (2016) *Education system in the UK*. Hentet 8.5.2023 fra: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/219167/v01-2012ukes.pdf

Gov.uk (2011) *National Minimum Wage to rise from 1 October 2011*. Hentet 7.5.2023 fra: <https://www.gov.uk/government/news/national-minimum-wage-to-rise-from-1-october-2011>

Heckman, J. J., Lochner, L. J., & Todd, P. E. (2006). *Earnings Functions, Rates Of Return And Treatment Effects: The Mincer Equation And Beyond*. Hentet 30.4.23 fra: [https://doi.org/10.1016/S1574-0692\(06\)01007-5](https://doi.org/10.1016/S1574-0692(06)01007-5)

OECD (2022). *Average wages*. Hentet 11.5.2023 fra: <https://data.oecd.org/earnwage/average-wages.htm>

OECD (2015). *Employment rates by age group*. Hentet 8.4.23 fra: <https://www.oecd-ilibrary.org/docserver/factbook-2014-53-en.pdf?expires=1684001264&id=id&accname=guest&checksum=95CE17209561FDDF1DD1B97EFCDA01DE>

OECD (2021). *Employment: Gender wage gap*. Hentet 14.3.23 fra: https://stats.oecd.org/Index.aspx?DataSetCode=GENDER_EMP#

Office for National Statistics. (2022). *Gender pay gap in the UK: 2022*. Hentet 14.3.23 fra <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/bulletins/genderpaygapintheuk/2022>

Rousseeuw, P. J., & Hubert, M. (2011). Robust statistics for outlier detection. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1), 73–79. Hentet 30.4.23 fra: <https://doi.org/10.1002/widm.2>

Statistisk sentralbyrå. (2023). *Likestilling - Fakta om befolkningen*. Hentet 24.4.23 fra <https://www.ssb.no/arbeid-og-lonn/sysselsetting/statistikk/arbeidskraftundersokelsen>

Statistisk sentralbyrå (2013). *Lønn, 2012*. Hentet 27.03.2023 fra: <https://www.ssb.no/arbeid-og-lonn/statistikker/lonnansatt/aar/2013-03-20>

Stigler, G. J. (1950). *The Development of Utility Theory. I. Journal of Political Economy*, 58(4), 307–327. Hentet 18.3.23 fra: <http://www.jstor.org/stable/1828885>

Stokke, H. E. (2021). *The gender wage gap and the early-career effect: the role of actual experience and education level. LABOUR*, 35(2), 135–162. Hentet 3.5.23 fra: <https://doi.org/10.1111/labr.12191>

Wooldridge, Jeffrey M. *Introductory Econometrics : A Modern Approach*. 5th ed., Mason Cengage Learning, 2013, pp. 83–93.

World Economic Forum. (2022). *The Global Gender Gap Report 2022*. Hentet 14.3.23 fra: https://www3.weforum.org/docs/WEF_GGGR_2022.pdf

8. Vedlegg

8.1 Valg av funksjonsformer

I denne studien utføres en komparativ analyse av diverse funksjonsformer med det formål å identifisere den mest passende regresjonsmodellen som optimalt tilpasser seg det gitte datasettet. En regresjonsanalyse implementeres ved hjelp av statistikkprogrammet Stata, med fokus på variablene timelønn og antall år med utdanning.

For å finne funksjonsformen som beskriver relasjonen mellom de to variablene, sammenlignet vi ulike modeller basert på deres justerte R-kvadrat, AIC og BIC. Det er viktig å merke seg at direkte sammenligning av SSR mellom lineære og logaritmiske modeller kan være utfordrende, ettersom de har ulik verdi på den avhengige variabelen. Derfor tar vi høyde for andre kriterier som justert R-kvadrat, AIC (Akaike Information Criterion) og BIC (Bayesian Information Criterion) og Zarembka-metoden for å sammenligne modeller, da disse er mer allment akseptert og tar hensyn til både modelltilpasning og modellkompleksitet. Det er også viktig å merke seg at AIC og BIC ikke kan sammenlignes direkte, ettersom de er beregnet for ulike antall observasjoner.

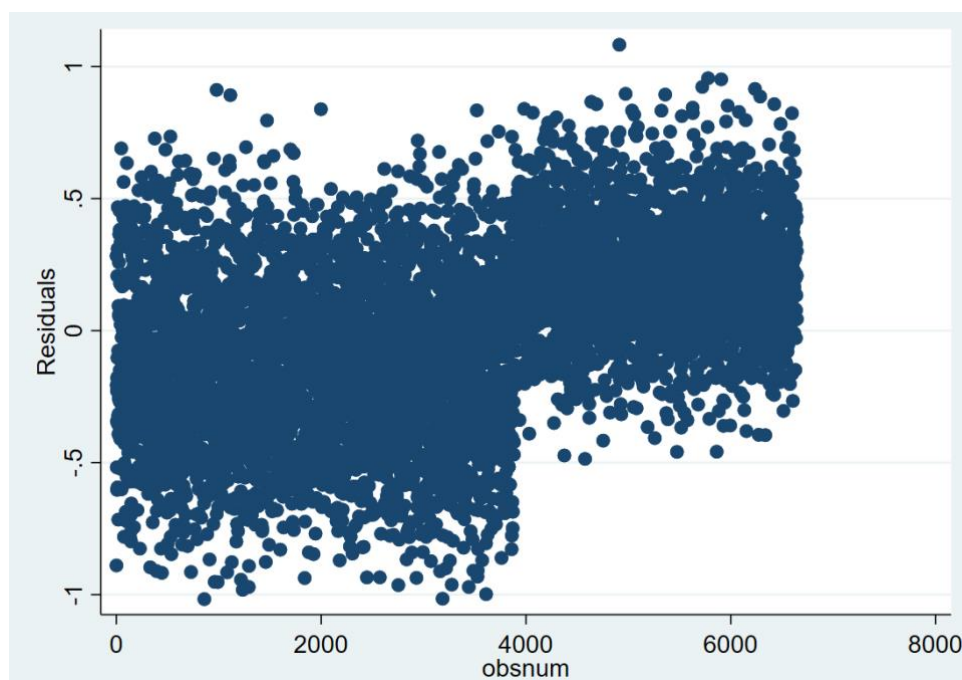
Basert på vår analyse valgte vi å bruke log-lin-modellen, ettersom den hadde høyere justert R-kvadrat og lavere AIC- og BIC-verdier. Dette indikerer at log-lin modellen gir en bedre tilpasning til vårt datasett og dermed en mer nøyaktig og pålitelig beskrivelse av relasjonen mellom variablene. Denne beslutningen er basert på en grundig vurdering av modellenes statistiske egenskaper, sammen med en teoretisk forankring og praktiske hensyn knyttet til den spesifikke konteksten for vår studie.

Funksjonsform	Formel	Justert R-kvadrat	AIC	BIC
Lineær	$Y = \alpha + \beta X$	0.2626	44117.84	44158.64
Lin-log	$Y = \alpha + \beta \ln(X)$	0.2519	23812.06	23843.07
Invers modell	$Y = \alpha + \beta \frac{1}{X}$	0.1785	24153.72	24184.73
Log-lin	$\ln Y = \alpha + \beta X$	0.2735	4303.306	4344.113

Log-log	$\ln Y = \alpha + \beta \ln(X)$	0.2754	2282.628	2313.64
---------	---------------------------------	--------	----------	---------

8.2 Test for uavhengighet

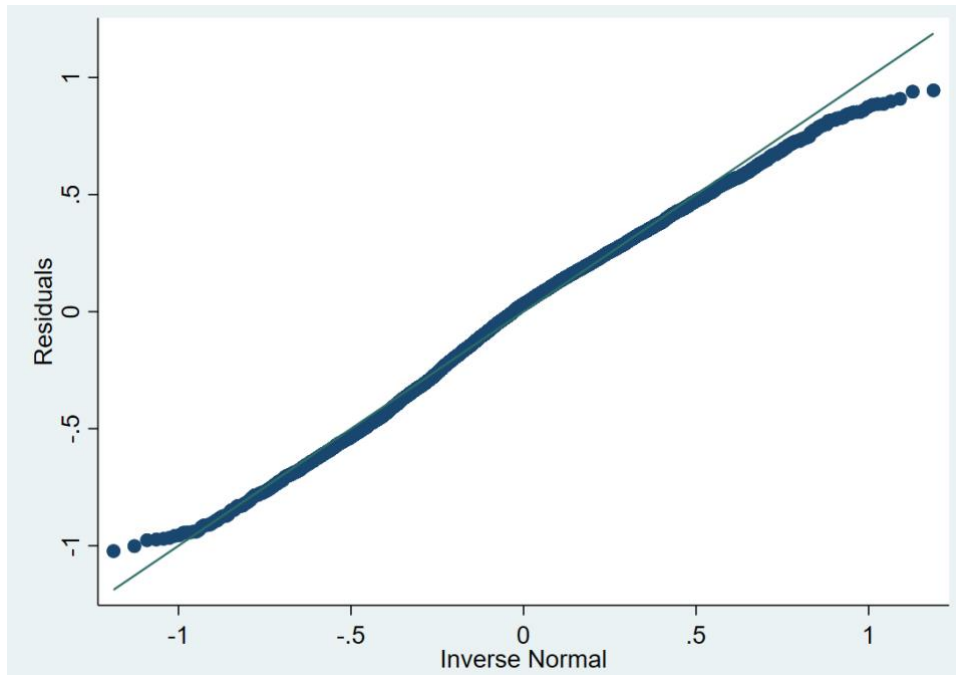
I henhold til MLR-3 antagelsen tester vi for uavhengighet for å finne indikasjoner på autokorrelasjoner. Siden datasettet vårt er ikke en tidsserie, lager vi en stegvis test hvor vi sjekker residualene i vår modell i Stata.



Her ser vi at residualene er spredt tilfeldig rundt null og ikke viser noen åpenbar trend eller mønstre. Dette tyder på at det ikke er noen autokorrelasjon.

8.3 Test for normalitet

Vi antar at feilene (residualene) er normalfordelte. Dette tester vi for i Stata (Q-Q-plott):



Analyse av residualene, visualisert i et Q-Q-plot, indikerer at disse i stor grad følger en normalfordeling. Dog, ved ytterpunktene av distribusjonen (ved -1 og 1) observeres avvik fra linjen, noe som tyder på "tyngre haler" enn forventet i en ideell normalfordeling. Dette er et tegn på en leptokurtisk fordeling, der vi ser flere ekstreme verdier enn hva som er typisk for en normalfordeling. For å tilpasse oss denne observasjonen, har vi anvendt en "trimming"-metode ved å fjerne de laveste og høyeste 5% av observasjonene. Med denne justeringen, kan vi anta at forutsetningen om normalfordeling er tilfredsstillt i vår analyse.

8.4 Test for multikollinearitet

Ingen multikollinearitet innebærer at de uavhengige variablene ikke er for sterkt korrelerte med hverandre. Dette sjekker vi for i Stata ved bruk av "vif" kommandoen og får følgende tabell:

Modell	A	B	C	D	E	F
Mean VIF	7.18	6.35	10.34	215.66	7.41	10.52

8.5 Test for heteroskedastisitet

Heteroskedastisitet oppstår når feilleddene i en modell ikke har konstant varians over observasjonene, noe som bryter forutsetningene for minste kvadraters metode (MKM). Når heteroskedastisitet er til stede, kan MKM-estimatorer være ineffektive, og dette resulterer i unøyaktige estimater av feilleddenes og parameterens varians. For å påvise heteroskedastisitet i modellen, anvender vi Breusch-Pagan-testen, som undersøker om antagelsen om konstant feilleddvariens er oppfylt.

Nullhypotesen for Breusch-Pagan-testen er homoskedastisitet, det vil si at variansen i feilleddene er konstant over observasjonene. Alternativhypotesen er heteroskedastisitet, der variansen i feilleddene er ulik over observasjonene.

Resultatene fra testen (utført i Stata):

```
Breusch-Pagan/Cook-Weisberg test for heteroskedasticity
Assumption: Normal error terms
Variables: All independent variables

H0: Constant variance

      chi2(5) = 254.58
Prob > chi2 = 0.0000
```

Testresultatet gir en p-verdi nær null. Dette fører til at vi avviser nullhypotesen om homoskedastisitet og konkluderer med tilstedeværelsen av heteroskedastisitet i modellen. Heteroskedastisitet innebærer at forklaringsvariablene påvirker variansen i feilleddene.

For å korrigere for heteroskedastisitet og oppnå mer pålitelige inferenser, vil vi benytte robuste standardavvik i de etterfølgende regresjonsanalysene. Robuste standardavvik er mindre følsomme for antagelsesbrudd som heteroskedastisitet og gir dermed mer nøyaktige konfidensintervaller og hypotesetester.

8.6 Resultater fra regresjonsanalysene

Tabell 4: Komplette modelleringsestimatresultater. Estimerte standardavvik i parentes

	(1)	(2)	(3)	(4)	(5)	(6)
	A	B	C	D	E	F
<i>VARIABLES</i>	<i>lnhour</i>	<i>lnhour</i>	<i>lnhour</i>	<i>lnhour</i>	<i>lnhour</i>	<i>lnhour</i>
<i>Kvinner</i>	-0.1098*** (0.0085)	-0.1047*** (0.0084)	-0.0766*** (0.0185)	-0.0334 (0.0525)	-0.1146*** (0.0114)	-0.1129*** (0.0086)
<i>Utdanning</i>	0.02230*** (0.0010)	0.0270*** (0.0009)	0.0236*** (0.0012)	0.0262*** (0.00088)	0.0263*** (0.00089)	0.0220*** (0.001)
<i>Erfaring</i>	0.0175*** (0.0016)	0.0196*** (0.0012)	0.0194*** (0.0012)	0.0193*** (0.0012)	0.0187*** (0.0012)	0.0173*** (0.00148)
<i>Erfaring kvadrert</i>	-0.0002*** (0.00003)	-0.0003*** (0.00002)	-0.0003*** (0.00002)	-0.0003*** (0.00002)	-0.0003*** (0.000025)	-0.0002*** (0.00003)
<i>Leseferdighet er</i>	-0.0005*** (0.0001)	0.0004*** (0.0001)	0.0005*** (0.00016)	-0.0002 (0.00027)	0.0005*** (0.00016)	-0.0006*** (0.00018)

<i>Tallferdighet er</i>	0.0030*** (0.0001)	0.0016*** (0.0001)	0.0014*** (0.00014)	0.0013*** (0.00027)	0.0014*** (0.00014)	0.0029*** (0.00016)
<i>Storbritannia</i>		-0.3870*** (0.0101)	-0.3849*** (0.0101)	-0.7694*** (0.05991)	-0.4311*** (0.01205)	
<i>Britisk kvinne</i>		0.0089 (0.0131)	-0.097*** (0.0213)	-0.071 (0.0809)	-0.0197 (0.0174)	
<i>Kvinne utdanning</i>			-0.0027 (0.0018)			
<i>Britisk kvinne utdanning</i>			0.0109*** (0.0019)			
<i>Kvinne tallf</i>				-0.0002 (0.00033)		
<i>Kvinne lesef</i>				-0.00007 (0.00036)		
<i>Britisk tallf</i>				0.0002 (0.0004)		
<i>Britisk lesef</i>				0.0011**		

	(0.00044)	
<i>Britisk kvinne tallf</i>	0.00012 (0.00053)	
<i>Britisk kvinne lesef</i>	0.00017 (0.00058)	
<i>Offentlig sektor</i>		-0.0997*** (0.0127)
<i>Kvinne offentlig sektor</i>		0.0677*** (0.0166)
<i>Storbritannia offentlig sektor</i>		0.1619*** (0.0214)
<i>Britisk kvinne off sektor</i>		-0.0113 (0.0273)
<i>Uten kontrakt</i>		0.1839** (0.0733)
<i>Storbritannia uten kontrakt</i>		-0.5385*** (0.0818)

<i>Kvinne uten kontrakt</i>						0.0258 (0.08804)
<i>Britisk kvinne uten kontrakt</i>						0.0812 (0.0999)
<i>Konstantledd</i>	1.883*** (0.0332)	2.208*** (0.0278)	2.233*** (0.0278)	2.432*** (0.0391)	2.254*** (0.0269)	1.911*** (0.0328)
<i>Observasjoner</i>	6642	6642	6642	6642	6642	6642
<i>R²</i>	0.281	0.495	0.498	0.502	0.507	0.298

