

Leonhard Åserud Spidsø
Preben Kristoffersen Paulsen
Stian Brandtzæg Havdal
Oliver Løbak Sæther

Empirisk analyse av lønnsforskjeller mellom kvinner og menn i Japan og i Norge

Bacheloroppgave i Samfunnsøkonomi
Veileder: Sepehr Cyrusian
Mai 2024

Leonhard Åserud Spidsø
Preben Kristoffersen Paulsen
Stian Brandtzæg Havdal
Oliver Løbak Sæther

Empirisk analyse av lønnsforskjeller mellom kvinner og menn i Japan og i Norge

Bacheloroppgave i Samfunnsøkonomi
Veileder: Sepehr Cyrusian
Mai 2024

Norges teknisk-naturvitenskapelige universitet
Fakultet for økonomi
Institutt for samfunnsøkonomi



Kunnskap for en bedre verden

Sammendrag:

Vi har i denne bacheloroppgaven tatt for oss og undersøkt lønnsgapet mellom kvinner og menn i Japan og i Norge med utgangspunkt i OECDs rapport om ulikheter i arbeidslivet. Vi har anskaffet og tatt i bruk data fra OECDs PIAAC-undersøkelse for å analysere påvirkningen av ulike variabler som utdanning, erfaring og relevante kunnskaper. Resultatene viser et lønnsgap i begge land, men fremhever seg som betydelig større i Japan. Videre i oppgaven konkluderes det i at de uavhengige variablene ikke kan forklare lønnsgapet, og andre innspill som preferanseteori og kulturelle forskjeller antydes som mulige årsaker til den større forskjellen i Japan sammenlignet med Norge.

Innholdsfortegnelse:

Sammendrag:	1
1. Innledning	3
1.1 Problemstilling	5
1.2 Struktur.....	5
2. Teori og Tidligere Litteratur	6
2.1 Humankapitalteori.....	6
2.2 Preferanseteori	7
2.3 Tidligere forskning- Lønnsgapet I dag.....	7
3. Datamateriale	8
3.1 Beskrivelse av datamaterialet	9
3.2 Deskriptive analyser.....	9
3.2.1 Deskriptiv statistikk for uavhengige variabler	11
3.3 Databehandling	13
4. Teoretisk rammeverk og modeller	14
4.1 Minste kvadraters metode.....	15
5. Regresjonsanalyse	17
5.1 Regresjonsanalyse – kvinner i Japan	18
5.2 Regresjonsanalyse- kvinner i Norge	19
5.3 Resultater	21
6. Signifikanstesting av uavhengige variabler.	21
6.1 Er kjønn en signifikant faktor for avkastning på timelønn?	22
6.2 Observasjon av signifikansnivå for resterende uavhengige variabler	24
7. Avslutning	26
7.1 Konklusjon	26
7.2 Kritikk	27
7.3 Videre forskning	27
8. Litteraturliste	28

1. Innledning

I tiårene etter andre verdenskrig har Norge og Japan, to land av ulike geografiske og kulturelle grunnlag, sett en drastisk økonomisk vekst innen landets økonomi. Selv for en progresjon innen personlig økonomi og arbeidsliv, publiserte OECD en rapport i 2015 (Organisation for Economic Co-operation and Development, 2015) som resulterte i å bevisstgjøre et stort økonomisk skille mellom menn og kvinner i Japan, i motsetning til Norge hvor lønnsgapet mellom menn og kvinner er betydelig mindre (2015). Vi skal i denne bacheloroppgaven i samfunnsøkonomi ta for oss disse lønnsforskjellene mellom kvinner og menn i Norge og Japan. Lønnsnivå er et dagsaktuelt tema, spesielt med tanke på at det er ønskelig å oppnå likestilling i samfunnet mellom kvinner og menn

Vi har i denne oppgaven valgt å ta for oss Norge og Japan på bakgrunn av deres fellestrekk når det kommer til økonomisk vekst, livskvalitet og trygghet, men samtidig deres motsetninger når det kommer til kultur og forventinger. Japan bygde opp sin økonomi igjen etter andre verdenskrig og har til i dag utviklet en svært diversifisert industri- og tjenesteøkonomi. Dette medfører at Japan er en av verdens største produsenter av motorvogner, stål og høyteknologiske industrivarer (Masai, Latz, & Hijino, 2024). Norge på den andre siden, tjente store summer på oljeproduksjon. Dette spilte en stor rolle i å drive frem en velstående norsk økonomi og gi nordmenn en av verdens høyeste inntekter per innbygger (Encyclopedia Britannica, 2024).

GDP & GDP per capita: Norge & Japan

	Land	
	Norge	Japan
GDP (USD) i milliarder	490,29	5011,87
GDP (USD) per capita	90559,48	39933,01

Tabell 1: GPD Norge og Japan

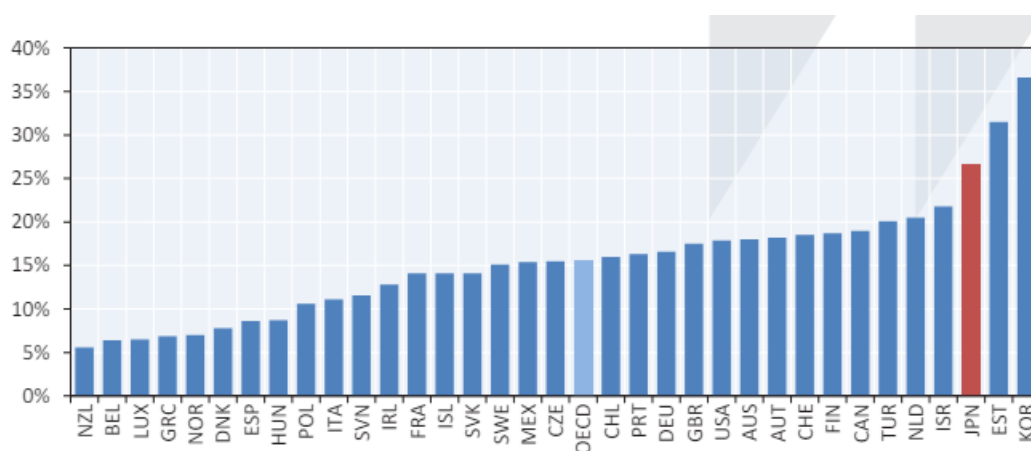
Ser vi på BNP ut ifra sektor, kan vi se at begge landene har sentralisert seg rundt å levere varer og tjenester. Ser vi på *Figur 1* ser vi at tjenester i Japan står for hele 70% av BNP, mens 58% av BNP i Norge omgår tjenester.



Figur 1: Japan GDP etter sektor (Statista 2023).

Figur 2: Norge GDP etter sektor (Statista 2023).

Selv om landene tilsynelatende har en relativt lik økonomisk vekst, er det store kulturelle forskjeller som gjenspeiles i lønnsforskjeller mellom kvinner og menn. The Global Gender Gap Index fra 2023, som er en indeks basert på økonomiske muligheter, utdanning, helse og politisk ledelse, står Norge oppført på 2.plass, mens Japan er nede på 125.plass (Zahidi, S, 2023). Noe som kan indikere at det er store forskjeller for hvilke muligheter kvinner i Japan får sammenlignet med kvinner i Norge. Siden begge landene er medlem av “*Organisation for Economic Co-operation and Development*” (OECD), forventes det kanskje et mindre skille. I rapporten “*Inequality*” i “*Japan Policy Brief*” av OECD blir det oppgitt at Japan rangerer som et av de dårligste landene i OECD nettverket når det kommer til likestilling i inntekter.



Graf 1: Gendergap in the OECD (Organisation for Economic Co-operation and Development, 2015).

1.1 Problemstilling

Vi ønsker i denne oppgaven å undersøke ved bruk av innsamlet kvantitativ data fra OECD's PIAAC undersøkelse, om det eksisterer et lønnsgap mellom kvinner og menn i Norge og Japan, og størrelsen på dette. Vi vil så se på forskjellene mellom Norge og Japan, og stille spørsmålet om disse forskjellene kan forklares på bakgrunn av uavhengige variabler som hovedsakelig handler om utdanning, kunnskaper og erfaring. Vi vil med dette se på om et større lønnsgap i Japan enn Norge, som OECD rapporten antyder, kan forklares i en lavere positiv avkastning på lønn for kvinner i Japan enn kvinner i Norge. Vi kan med dette frembringe følgende problemstilling og delproblemstilling:

“Eksisterer det en lønnsforskjell mellom kvinner og menn i Norge og i Japan?”

Delproblemstilling: “Hvordan påvirker utdanningsnivå, kognitive ferdigheter, arbeidserfaring og sektor størrelsen på lønnsgapet”

Vi konkluderer i denne oppgaven med at; ja, det eksisterer et lønnsgap mellom kvinner og menn i begge land, men skillet er betydelig større i Japan. At skillet er større i Japan enn i Norge kan ikke forklares i at det er spesielt store ulikheter på avkastningen til de ulike variablene i regresjonsanalysen. Resultatene er signifikante, og vi kan heller anta at lønnsgapet kan forklares i større grad ved bruk av oppgavens teorier.

1.2 Struktur

Oppgaven vil videre presentere to relevante teorier i “*Humankapitalteori*” og “*Preferanseteori*”, samt tidligere forskning på tema. Videre redegjøres det for oppgavens datagrunnlag, og relevant data presenteres og diskuteres. Oppgaven presenterer så relevant teoretisk rammeverk, der vi hovedsakelig fokuserer på en logaritmisk lineær regresjonsanalyse som videre blir brukt i analysen. Videre blir den lineære modellen utført i

Stata og resultatene blir diskutert i to individuelle regresjonsanalyser. Etter analysen testes resultatene for å forsikre oss om at de er statistisk signifikante.

Oppgaven avsluttes med en avslutning bestående av redegjøring av konklusjon, kritikk og videre forskning.

2. Teori og Tidligere Litteratur

I dette segmentet av oppgaven skal vi se på teorier og tidligere litteratur/empiri som kan hjelpe sette lys på lønnsforskjeller mellom kvinner og menn. Vi har i hovedsak tatt utgangspunkt i å presentere to teoretiske rammeverk: humankapitalteori og preferanseteori. Videre vil vi benytte oss av tidligere forskning knyttet til Japan og Norge. Ved å kombinere empiri med det teoretiske rammeverket og gjennomføre en regresjonsanalyse, kan vi undersøke lønnsgapet i dag og sammenlikne den nåværende situasjonen i de to landene.

2.1 Humankapitalteori

Humankapitalteori handler om investeringen i individer, og hvordan det kan påvirke både avkastning på både et individuelt nivå og i en større sammenheng på et bedriftsnivå. I denne oppgaven vil den relevante avkastningen være i form av lønn (*wage*). Investering i kapitalen hos mennesker kan gjøres gjennom utdanning, praksistrening («*on the job training*»), helseordninger og flere tiltak som bedrer individers velferd ifølge Becker (Becker, 1962, s. 9). Ved å se på individuell investering kan vi få et bredere perspektiv om det er forskjell mellom menn og kvinner. I PIACC undersøkelsen ser vi at både japanske og norske kvinner har høy investering i utdanning (PIACC). Ifølge Becker skal høyere utdanning føre til en større avkastning på lønn (Becker, 1962 s. 9). Om vi tar for oss et større nivå kan vi se om bedrifter i de respektive landene investerer mer i et kjønn enn det andre. Dersom det befinner seg en tydelig skjevhet mellom investeringer i kvinner og menn kan dette være en faktor som er med å forklare forskjellene i lønn.

2.2 Preferanseteori

Innenfor økonomi kan preferanseteori beskrives som å være en studie som undersøker årsakene til individers valg ut ifra preferanser. Teorien bygger på den utilitaristiske teorien utviklet av Jeremy Bentham hvor valg baserer seg på konsekvenser (Carson & Kosberg, 2018, s. 23). For å knytte dette opp mot det aktuelle temaet kan vi se på Catherine Hakims preferanseteori, som setter lys på kvinners valg mellom karriere-satsing eller familieliv (Hakim, 2000). Dette kan være med på å forklare lønnsforskjellene i land hvor den kulturelle samfunnsnormen for kvinner peker mer mot arbeid i hjemmet og ansvar for familien. Preferanser trenger nødvendigvis ikke spesifikt å omhandle kun de to ovennevnte retningene, men kan også omfatte en bredere spekter av faktorer og valgkriterier som påvirker individets karrierevei og arbeidsdeltakelse. Dette kan bidra til å forklare forskjellene mellom kjønn i landene.

2.3 Tidligere forskning- Lønnsgapet I dag

For å se på dagens lønnsforskjeller i Norge og Japan har vi valgt å benytte oss av World Economic Forum's Global Gender Gap Report. Totalt var det 146 land som ble inkludert i undersøkelsen. Rapporten tar utgangspunkt i fire områder: *økonomiske deltakelse og muligheter, utdanning, helse og politisk myndiggjøring*. I rapporten finner vi en indeksrangering av landene, både totalt og for hvert område. Indeksen beskriver landenes kjønnsparitet med en verdi fra 0 til 1, hvor 1 er fullstendig paritet mellom kjønnene (Zahidi, 2023). I rapporten kan vi se at Norge har fått en verdi på 0,879 i 2023 som tilsvarer en 2.plass blant deltakerne. Japan derimot har en verdi på 0,647 som tilsvarer en plassering på 125.plass.

Det er vel dokumentert at det er klare kjønnsforskjeller mellom kvinner og menn i Japan i dag. Vi kan spesielt se dette ved lønnsgapet. I dag er lønnsgapet mellom menn og kvinner blant heltidsansatte i Japan betydelige høyere enn i Norge. Dette måles ved forskjellen i median lønn mellom menn og kvinner sammenliknet med medianlønnen til menn (OECD, 2023). En annen faktor som kan gi et innblikk på forskjellene er arbeidstid og stilling i arbeidsmarkedet. I Japan har det lenge vært en kulturell rolle som kvinner oppfyller ved ekteskap. Ved stiftelsen av en familie, er det vanligvis kvinnen som tar hovedansvaret for

omsorgen for barnet hjemme. I en rapport fra OECD viser det at kvinner oftest kommer tilbake til arbeidslivet ved å ta en deltidsstilling, eller ikke kommer tilbake i det hele tatt (OECD, 2023). Tiden hjemme vil kan være med på å svekke kvinners «*on the job learning*», hvor et mulig utfall er en stagnering av humankapitalveksten.

Videre kan vi se at yrkesvalg og sektor kan spille en rolle i dagens lønns-gap. Som nevnt tidligere er Norge et av de landene som stiller best ved kjønnsparitet i WEF's undersøkelse. Likevel har kvinner mindre gjennomsnittslønn enn menn (Fløtre, 2024). Dette kan til en viss grad forklares av inndeling på arbeidsmarkedet og yrker hvor det er «overrepresentasjon» fra kjønnene. I den private sektoren er det gjennomsnittlig høyere lønn enn i den offentlige, og i en undersøkelse fra SSB nevnes det at den offentlige sektoren er «kvinnedominert» (Grini & Fløtre, s. 18.).

I analysen har vi valgt å justere etter kjøpekraftsparitet (PPP). «Kjøpekraftspariteter er prisnivåindikatorer som uttrykker prisnivået i et gitt land på et gitt tidspunkt, relativt til prisnivået i ett eller flere land» (Svennebye, 2013). Etersom det kan forekomme ulikheter ved prisnivåer i de utvalgte landene vil det være logisk å se på kjøpekraftsparitet. Ved å justere for kjøpekraftsparitet vil det gi et mer nøyaktig bilde ved sammenligningen av økonomiske faktorer mellom landene.

3. Datamateriale

I dette kapittelet skal vi redegjøre for og presentere datamengden som er funnet og benyttet i oppgaven. Videre skal vi presentere det som direkte er funnet i den deskriptive analysen, samt redegjøre for styrker og svakheter.

3.1 Beskrivelse av datamaterialet

Datamaterialet brukt i bacheloroppgaven inneholder data om Norge og Japan, og er hentet ut fra PIAAC-undersøkelsen, gjennomført i samarbeid med OECD. PIAAC-undersøkelsen er en internasjonal undersøkelse som administreres hvert tiende år. Undersøkelsen går ut på å spørre 5 000 individer fra hvert land, i alderen 16 til 65 år. PIAAC samler videre inn data basert på befolkningen sitt samlede ferdighetsnivå med å undersøke de tre områdene; problemløsning innen IKT, tallforståelse og leseferdigheter (OECD, 2016). OECD bruker disse ferdighetene som variabler fordi det beskriver hvor god en person er til å ta til seg læring samt potensialet til å utvikle andre ferdigheter fremover. Videre ble deltakerne i undersøkelsen spurt om andre faktorer som alder, antall år skolegang, kjønn, familie, sektor, arbeidserfaring m.m. I oppgaven har vi kun lastet ned nødvendig data for Norge og Japan, der vi videre har undersøkt resultatene ved bruk av programmet STATA.

Videre i analysen har vi valgt timelønn som vår avhengig variabler, og kun valgt individene i PIAAC-en som har oppgitt timelønn. Videre har vi inkludert for både privat og offentlig sektor, samt ubetalte verv. For å få et realistisk bilde av lønn har vi videre fjernet 2% med lavest oppgitt timelønn, samt 2% av individene med høyest oppgitt timelønn. Dette er for å fjerne ekstreme verdier som ikke er generaliserbart overfor befolkningen. Videre har vi valgt og separerer de som jobber i offentlig sektor, sammenlignet med de som jobber i en non-profit-organisasjon.

3.2 Deskriptive analyser

De deskriptive analysene har som mål å se på forskjellen i timelønn mellom kvinner og menn i Norge og Japan, og gi dypere forståelse for resultatet i regresjonsanalysen. I analysen for den avhengige variabelen (timelønn), har vi i “*Tabell 1*” sett på timelønn i Norge, samt justert timelønn etter USD, ved bruk av kjøpekraftsparitet teori og OECD sin PPP-kalkulator. Årsaken til at vi benytter kjøpekraftsparitetsteori, er for å ta hensyn til variasjon i to ulike valutakursen, samtidig som man tar hensyn til at prisnivået i ulike land varierer (Holden, 2016). For Japan blir samme justering gjort og presentert i “*Tabell 2*”.

For å analysere den uavhengige variabelen kvinne, også kalt dummy variabel, har vi gitt kvinne tallverdien 1, mens mann har fått verdi 0. En annen binær uavhengige variabel vi har i oppgaven er offentlig sektor. De ikke-binære uavhengige variablene er: Alder, antall år skolegang, tallforståelse, leseforståelse og erfaring.

Timelønn - Norge:

	Norske kroner - NOK			PPP - USD		
	Totalt	Menn	Kvinner	Totalt	Menn	Kvinner
Antall observasjoner	3628	1853	1775	3628	1853	1775
Gjennomsnittlig timelønn	215,31	232,3	197,73	24,25	26,17	22,27
Minimum	39,55	39,64	39,55	4,45	4,46	4,45
Maksimum	657,89	657,89	607,29	74,11	74,11	68,41
Standardavvik	83,16	92,39	68,06	9,37	10,41	7,67

Tabell 1: Deskriptiv statistikk for lønn i Norge. Presentert i NOK og USD, justert etter PPP.

Timelønn - Japan:

	Japanske Yen - JP ¥			PPP - USD		
	Totalt	Menn	Kvinner	Totalt	Menn	Kvinner
Antall observasjoner	3323	1735	1588	3323	1735	1588
Gjennomsnittlig timelønn	1647,05	2008,1	1253,83	15,42	18,8	11,74
Minimum	367,006	384,62	367,006	3,44	3,6	3,44
Maksimum	12019,23	12019,23	10509,72	112,51	112,51	98,38
Standardavvik	1212	1405,45	788,45	11,35	13,16	7,38

Tabell 2: Deskriptiv statistikk for lønn i Japan. Presenter i Yen og USD, justert etter PPP.

3.2.1 Deskriptiv statistikk for uavhengige variabler

Vi ønsker i denne oppgaven å få et innblikk i hva det er som forårsaker et lønnsgap mellom kvinner og menn i Norge og Japan, og om disse faktorene har mer betydning i et land fremfor det andre. For å analysere faktorer som påvirker timelønn tar vi for oss flere sentrale variabler med relevant forklaringskraft for vårt forskningsspørsmål.

Oppgavens uavhengige variabler er hentet fra samme OECD panel-datasett som vår avhengige variabel, og er justert for å kun omhandle relevante observasjoner. Følgende variabler kan kort forklares som følgende:

Alder: Deltakeren ble bedt om å dokumentere alder. Disse rangerte fra 16 til 65 år.

Antall år skolegang: Denne variabelen tar for seg totalt antall skoleår vedkommende har studert.

Tallforståelse og leseforståelse: To sentrale dokumentasjoner i OECD's PIAAC undersøkelse. Disse to variablene tar for seg respondentens nivå til å bearbeide og ta til seg oppgaver som omhandler personens tallforståelse og leseferdigheter.

Erfaring: Denne variabelen henviser til hvor mye erfaring deltageren har opparbeidet seg i arbeidslivet. I tillegg til c_q09 (*erfaring*), bruker vi og variabelen c_q09^2 . Denne inkluderes senere i regresjonsanalysen ettersom erfaring er avtakende, og vil ikke lineært øke i løpet av en lengre periode. Dette er ikke presentert i "Tabell 3".

Sektor: Siste uavhengige variabel tar for seg hvilken sektor individet arbeider i, og er tildelt tre forskjellige verdier der 0 er gitt "privat sektor", 1 er gitt "offentlig sektor" og 2 er tildelt "en veldedighetsorganisasjon".

Følgende tabell illustrerer gjennomsnittet og standardavvik for menn og kvinner samlet, for menn alene, og kvinner, i de to forskjellige landene oppgaven ønsker å ta for seg. Variablenes originale merkinger fra PIAAC-undersøkelsen er vedlagt i tabellen:

Uavhengige variabler:

	Total		Menn		Kvinner	
	Japan	Norge	Japan	Norge	Japan	Norge
Alder (age_r)	43,66 (11,04)	43,54 (10,65)	43,42 (11,24)	43,58 (10,75)	43,91 (10,84)	43,51 (10,54)
Antall år skolegang (yrsqual)	13,44 (2,28)	14,93 (2,3)	13,59 (2,52)	14,79 (2,37)	13,27 (1,99)	15,07 (2,23)
Tallforståelse (pvnum1)	292,60 (42,88)	290,96 (50,09)	298,14 (44,70)	297,02 (51,84)	286,79 (40,09)	284,74 (47,45)
Leseforståelse (pvlit1)	299,95 (38,26)	287,72 (43,63)	301,16 (39,41)	288,41 (45,38)	298,66 (36,97)	287,00 (41,76)
Erfaring (c_q09)	20,13 (11,22)	21,47 (11,08)	22,56 (11,97)	22,48 (11,61)	17,59 (9,74)	20,43 (10,41)
Sektor (d_q03)	0,21 (0,48)	0,45 (0,53)	0,19 (0,45)	0,31 (0,49)	0,22 (0,51)	0,59 (0,53)

Tabell 3: Deskriptiv statistikk for uavhengige variabler.

Selv om de uavhengige variablene har som hensikt å teste, og potensielt forklare vårt forskningsspørsmål, er det fortsatt et par sentrale trekk vi kan hente ut fra tabellen. Vi kan se at resultatene for Japan og Norge fremstår som veldig jevne, og det er hovedsakelig bare små verdier som skiller de to landene. Den største kontrasten på landsbasis er en høyere verdi for tall- og leseforståelse fra Japan. Det er også veldig små forskjeller internt mellom kvinner og menn i Japan når det kommer til disse variablene.

Innenfor utdanning er det norske kvinner som topper gjennomsnittet, med norske menn tett bak. En av de største kontrastene vi observerer er skillet mellom erfaring blant japanske menn og japanske kvinner. Hvor norske menn og kvinner ligger relativt nærmere, skiller japanske menn og kvinner seg betydelig mer.

En annen viktig faktor å påpeke er at det er norske kvinner som strekker seg mest mot den offentlige sektoren, i kontraster til japanske menn som ligger nærmest den private sektoren.

3.3 Databehandling

Svarene som er hentet ut fra PIAAC-undersøkelsen er frembrakt ved bruk av et selvutfyllingsskjema. Representantene i undersøkelsen blir stilt to forskjellige variasjoner av spørsmål; lukkede og åpne. Innsamlet data er hovedsakelig prydet av svar generert fra lukkede spørsmål, hvor respondenten besvarer et eller flere av faste svaralternativer. I noen tilfeller blir respondenten spurt et åpent spørsmål, hvor han/hun fritt kan svare hva de ønsker, ofte i form av en spesifikk sum eller et spesifikt sted.

Når vi utfører databehandlingen for denne oppgaven, er det flere ting vi må ta hensyn til. Et av punktene er at vi ser på fordelingen av demografiske variabler som kjønn, alder og landsdel. I Kristen Ringdals bok "Enhet og mangfold" påpeker han at det er vanlig at kvinner er noe underrepresentert i norske intervjuundersøkelser (Ringdal, 2018, s.279). PIAAC er gjennomført av OECD. Vi observerer at mangelen på representasjon av kvinner ikke er et problem i oppgavens data. Videre er aldersfordelingen stabil, og landsdelsfordelingen er ikke tatt med i undersøkelsen.

Videre tar vi for oss hvordan manglende informasjon i datamatriksen svekker utvalgets representativitet. Vi må altså ta hensyn til at for hver person som ikke svarer på et spørsmål, blir det et hull i datamatriksen (Ringdal, 2018, s.280). Vårt utvalg av data har i flere tilfeller variabler med manglende verdier, men dette påvirker ikke analysen ettersom relevante variabler og kontroll-verdier er tilgjengelig. Dette medfører at forskningsresultatene er reproduerbare og pålitelige.

Dette bringer oss videre til nytteverdien eller gyldigheten (validiteten) av resultatene. I en kvantitativ forskningsartikkel handler dette om at våre funn faktisk beskriver virkeligheten for hele populasjonen, og at resultatene ikke skyldes andre ytre faktorer som det ikke er

kontrollert for, eller ren tilfeldighet (Bjørnnes & Gjevjon, 2019). Ettersom oppgaven tar for seg variabler som er redegjort for og innsamlet på en legitim måte av OECD, kan vi konkludere med at ren tilfeldighet eller ytre faktorer ikke påvirker resultatet.

4. Teoretisk rammeverk og modeller

I kapittelet vil vi gå inn på regresjonsmodellen som vi bruker i bacheloroppgaven samt de tilhørende hypotesene. Hensikten er å gi en innføring av metodene og prinsipper som legger grunnlag for videre analyse. Multippel lineær regresjon (MLR) benyttes for å undersøke lønnsforskjellene mellom kvinner og menn, i Norge og Japan. I oppgaven benytter vi oss av den avhengige variabelen lønn og de uavhengige variablene utdanning, arbeidserfaring, sektor og de kognitive ferdighetene leseforståelse og tallforståelse. Videre skal vi benytte oss av en T-test for å undersøke om forskjellene ved gjennomsnitt mellom gruppene er statistisk signifikante. T-testen gjennomføres i sammenheng med hypotesene, og vil hjelpe oss ved valget om å beholde eller avkaste.

For å analysere innsamlet data, og undersøke hvilke faktorer som påvirker lønnsnivået, tar vi først utgangspunkt i en økonometrisk modell. Vi definerer den forenklete lønnslikningen (1):

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

I ligningen tar man utgangspunkt i funksjonen for timelønn, for kvinner. Videre representerer x humankapital variabler som ferdigheter, utdanning, arbeidserfaring, sektor, osv. Som antatt tidligere i oppgaven vil variablene ha positiv effekt på timelønnen til individet. Videre, for å analysere avkastningen av de uavhengige variabler på den avhengig variable, bruker vi en log-lineær forklaringsmodell som tar for seg x 's påvirkning på Y . Vi kan med det presentere følgende ligning for regresjonsanalysen:

$$(2) LHRPPP = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + \varepsilon_i$$

Ligningen kan kort utredes følgende:

$$LHRPPP = \text{Log} - \text{Hourly PPP} (\hat{y}).$$

β = Koeffisient av uavhengig variabel.

α = Konstantleddet / skjæringspunktet når andre variabler er 0.

ε = Støyleddet. Tar for seg utvendige faktorer.

Vi prioriterer å bruke en logaritme variant av den klassiske regresjonsmodellen ettersom dette gjør resultatene enklere å sammenligne på tvers av landene. En logaritme variant oppgir resultatene som en prosentvis endring av avhengig variabel, basert på innflytelsen av de uavhengige variablene. Vi bruker i regresjonsanalysen den justerte verdien av timelønn “PPP”, fremfor å bruke landenes individuelle valuta NOK og Yen. Variabelen “PPP” ble i Stata generert (gen =) til en logaritme-variabel (ln) og ble gitt navnet “LHRPPP”. Videre i punkt 4. skal vi redegjøre for hva som ligger bak regresjonsanalysen som vi i denne oppgaven skal ta til bruk, og forklare hva en “*Ordinary Least Square*” (OLS) analyse er. OLS vil videre i oppgaven bli presentert som “MKM”.

4.1 Minste kvadraters metode

Minste kvadraters metode (MKM) er den enkleste formen for lineær regresjon. Hensikten er å finne den minste summen mellom det kvadrerte avviket (y) og den predikerte verdien (\hat{y}).

Når man minimerer det kvadrerte avviket, finner man estimatene av parameterne a og b , altså stigningstallet og skjæringspunktet. Den generelle likningen for populasjonen er gitt ved ($Y = \alpha + \beta x + \varepsilon$), der Y representerer den avhengige variabelen, x den uavhengige variabelen. α er konstantledd. Mens β -koeffisienten representerer stigningstallet til regresjonslinjen. ε er lagt inn for tilfeldige feil (Wooldridge,2020).

For å finne MKM bruker man følgende formel:

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

Y illustrerer den faktiske verdien med hensyn på observasjon i .

\hat{Y} illustrerer den predikerte verdien med hensyn på observasjon ut ifra estimat på α og β .

I observasjonen vil avviket mellom faktisk timelønn kontra predikert timelønn kalles for residualen.

Videre bestemmer MKM koeffisientene slik at resultatet av kvadratisk residual er så lite som mulig, mens støyledet har et gjennomsnitt på null. Ut ifra koeffisientene kan det oppstå en antakelse om at det eksisterer en kausal sammenheng mellom den avhengige og de uavhengige variablene, men vi må ta til hensyn for at kausalitet kan ikke bare sluttes fra regresjonskoeffisienter alene; det kan være utelatte variabler eller medvirkende faktorer, osv. Det blir da naturlig å utføre hypotesetester om koeffisientene, ved å benytte 5% signifikansnivå (Wooldridge, 2020). Ved bruk av MKM legger vi til grunn diverse antagelser. Først antar vi at forklaringsvariablene (verdiene for X_{i1} og X_{i2}), er de samme i hvert utvalg. Videre antar vi at et individs i er fast over observasjonene.

Antagelser for MKM-estimatorer (Wooldridge, 2020):

1. Null forventning av residualer, $E(\epsilon_i | X) = 0$, restleddet har forventning lik 0.
2. $Var(\epsilon_i | X) = \sigma^2$
3. $Cov(\epsilon_i | X_{ij}) = 0$, for alle i, j . Ingen samvariasjon mellom restleddet og tilhørende verdi på kontrollvariabler, i dette tilfellet betegnet som X_{ij} .
4. $\epsilon_i \sim N(0, \sigma^2)$, normalfordeling av residualen
5. Ingen sammenheng mellom kontrollvariablene.

5. Regresjonsanalyse

Vi ønsker i denne delen av oppgaven å analysere og diskutere regresjonsanalysen produsert i dataprogrammet Stata. Etter å ha behandlet og sortert relevant data fra OECD, utfører vi ved bruk av en logaritmisk regresjonsmodell som diskutert i del 4., fem variasjoner av en regresjonsmodell for avhengig variabel “LHRPPP”. I hver iterasjon av modellen vil vi inkludere en eller to ekstra uavhengige variabler for å øke forklaringskraften til modellen, de uavhengige variablene vi vil bruke er $gender_b$, som er en dummy variabel ment til å vise effekten av kjønn, der kvinne er satt som 1 og mann er satt som 0. De resterende uavhengige variablene er også forklart under 3.2.1, der $yrsqual$ viser totalt antall år med utdanning hos en deltaker. Variablene $pnvnum1$ og $pvlit1$ står for deltakerens resultater i de ulike testene som kartlegger tallforståelse og leseforståelse. I modell 4 legger vi til c_q09 og c_q09^2 , og de viser henholdsvis antall års erfaring i arbeid og de kvadrerte verdiene, der sistnevnte har som hensikt å vise om erfaring har en ikke-lineær effekt på lønn. Den siste uavhengige variabelen som legges til i modell 5 er d_q03 , som viser hvordan valg av sektor påvirker lønnsforskjellen mellom kvinner og menn.

Disse fem modellene presenteres følgende og blir utført i Stata for å produsere kommende “Regresjonsanalyse 1” og “Regresjonsanalyse 2”:

$$(1)LHRPPP = \alpha + \beta gender_b + \varepsilon_i$$

$$(2)LHRPPP = \alpha + \beta gender_b + \beta yrsqual + \varepsilon_i$$

$$(3)LHRPPP = \alpha + \beta gender_b + \beta yrsqual + \beta pnvnum1 + \beta pvlit1 + \varepsilon_i$$

$$(4)LHRPPP = \alpha + \beta gender_b + \beta yrsqual + \beta pnvnum1 + \beta pvlit1 + \beta c_q09 + \beta c_q09^2 + \varepsilon_i$$

$$(5)LHRPPP = \alpha + \beta gender_b + \beta yrsqual + \beta pnvnum1 + \beta pvlit1 + \beta c_q09 + \beta c_q09^2 + \beta d_q03 + \varepsilon_i$$

Vi kan påpeke at fra og med modell (4) er variabelen c_q09^2 inkludert. Denne variabelen er nødvendig for å sette et tak på den positive effekten av erfaring. Dette begrunnes ved at erfaring ikke kan ha en pågående stigende effekt i løpet av alle år, men vil på et tidspunkt avta.

5.1 Regresjonsanalyse – kvinner i Japan

Vi tar først for oss den genererte analysen for kvinner i Japan. Ved å utføre de fem modellene blir vi gitt følgende analyse:

Regresjonsanalyse - Kvinner Japan:

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
gender_b	-0.413*** (0.0266)	-0.380*** (0.0260)	-0.336*** (0.0262)	-0.298*** (0.0263)	-0.303*** (0.0264)
yrsqual		0.0758*** (0.00558)	0.0561*** (0.00619)	0.0594*** (0.00614)	0.0566*** (0.00622)
pvnum1			0.00509*** (0.000583)	0.00374*** (0.000580)	0.00380*** (0.000580)
pvlit1			-0.00341*** (0.000640)	-0.00155** (0.000648)	-0.00164** (0.000648)
c_q09				0.0339*** (0.00369)	0.0337*** (0.00369)
sq_c_q09				-0.000500*** (8.46e-05)	-0.000502*** (8.46e-05)
d_q03					0.0734*** (0.0269)
Constant	2.741*** (0.0184)	1.707*** (0.0782)	1.485*** (0.110)	0.873*** (0.119)	0.841*** (0.120)
Observations	3,323	3,322	3,322	3,310	3,309
R-squared	0.068	0.117	0.139	0.182	0.183

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Regresjonsanalyse 1: Betragtning av alle uavhengige variabler, utredet i Stata.

Ved bruk av innsamlet data og programvaren Stata kan vi illustrere et klart bilde over hvordan oppgavens uavhengige variabler påvirker timelønn. Et av oppgavens sentrale funn, og som tidligere diskuteres, er at å være kvinne medfører en kritisk reduksjon på timelønn med -0.413 satt opp mot motparten mann. Dette kan vi se i “Model (1)” hvor kjønn er eneste variabelen som analyseres.

Videre inkluderer vi våre uavhengige variabler for å få en oversikt over hvilken effekt disse har på timelønn. Fra modell 1 - 5 vil og introduksjonen av flere variabler påvirke utfallet på tidligere variabler, og vi kan med det se at det negative innspillet på timelønn av å være kvinne minsker frem til "Model (4)", før så i "Model (5)" hvor sektor inkluderes, øker gapet igjen. Vi kan se at alle uavhengige variabler medfører en positiv økning på timelønn, sett bort fra *sq_c_q09* og *pvlit1* som begge har negative koeffisienter. Som tidligere nevnt er *sq_c_q09* negativ for å forhindre en urealistisk vekst fra erfaring. For "*pvlit1*" (leseferdigheter) kan potensielt denne negative effekten forklares i at yrker som baserer seg på leseferdigheter, fremfor tallferdigheter som ofte assosieres med ingeniør, data eller økonomi, fremkommer i lavere lønnet yrker.

5.2 Regresjonsanalyse- kvinner i Norge

Videre bruker vi samme modell og fremgangsmåte for å analysere innsamlet data for Norge. Dette gir oss en regresjonsanalyse med litt andre verdier enn analysen i punkt 5.1:

Regresjonsanalyse - Kvinner Norge:

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
gender_b	-0.153*** (0.0160)	-0.164*** (0.0149)	-0.138*** (0.0151)	-0.132*** (0.0143)	-0.121*** (0.0146)
yrsqual		0.0734*** (0.00307)	0.0616*** (0.00330)	0.0484*** (0.00319)	0.0511*** (0.00327)
pvnum1			0.00204*** (0.000324)	0.00122*** (0.000307)	0.00114*** (0.000307)
pvlit1			-0.000689* (0.000366)	0.000535 (0.000351)	0.000617* (0.000351)
c_q09				0.0316*** (0.00206)	0.0314*** (0.00206)
sq_c_q09				-0.000487*** (4.78e-05)	-0.000478*** (4.77e-05)
d_q03					-0.0503*** (0.0140)
Constant	3.189*** (0.0112)	2.125*** (0.0457)	1.891*** (0.0567)	1.621*** (0.0553)	1.645*** (0.0557)
Observations	3,628	3,627	3,627	3,625	3,624
R-squared	0.025	0.157	0.178	0.277	0.280

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Regresjonsanalyse 2: Betragtning av alle uavhengige variabler, utredet i Stata.

Selv om vi i denne analysen og observerer et lønnsgap, kan vi se at lønnsgapet mellom kvinner og menn i Norge er betydelig mindre enn lønnsgapet i Japan. Regner vi sammen de logaritmiske koeffisientene kan vi si at det er cirka ett 26% større lønnsgap mellom kvinner og menn i Japan når variabelen kjønn er det eneste som påvirker vår avhengige variabel. I motsetning til analysen i 5.1, kan vi se at lønnsgapet minsker helt fra modell 1 til 5. Vi ser videre at nesten alle uavhengige variabler har en positiv effekt på timelønn for kvinner i Norge. En som skiller seg ut er d_q03 , altså sektor. Ser vi på “Tabell 3” under punkt 3.2 kan vi observere at norske kvinner er i høyeste grad de som jobber mest mot offentlig sektor.

5.3 Resultater

Sammenligner vi resultatene kan vi som tidligere påpekte at lønnsgapet mellom kvinner og menn i Japan er større enn gapet i Norge. Ettersom begge regresjonsanalysene forteller oss at innenfor OECD's genererte data befinner det seg et lønnsgap mellom kvinner og menns timelønn i begge land, kan vi med dette besvare oppgavens hovedproblemstilling; Ja, det eksisterer en lønnsforskjell mellom kvinner og menn i Norge og i Japan.

Videre ønsker vi å se på hvordan flere uavhengige variabler påvirker timelønn, og om disse kan forklare lønnsgapet. Vi har under "*Tabell 3*" i punkt 3.2 redegjort for at verdiene for menn og kvinner på flere punkter er relativt lik innad i landene. Skal de uavhengige variablene forklare og fremstå som hovedgrunnen for de ulike lønnsgapene, burde koeffisienten fremstått som negativ og med det hatt en negativ effekt på timelønn. For "*Regresjonsanalyse 1*" og "*Regresjonsanalyse 2*" kan vi observere at i de fleste tilfeller har de ulike variablene positive koeffisienter fra Model 2 til 5. Vi kan med dette besvare vårt delproblemstilling og si at avkastningen på utdanning, erfaring, kunnskaper og ulik sektor ikke kan begrunne lønnsgapet i Norge og i Japan.

Videre tester vi om disse resultatene er statistisk signifikant.

6. Signifikanstesting av uavhengige variabler.

I denne delen av oppgaven skal vi ta for oss om de presenterte variablene i regresjonsanalysen er statistisk signifikant, og om de faktisk påvirker timelønn for kvinner i Norge og Japan. Om variablene er statistisk signifikant, betyr dette at det er en lineær sammenheng, og at det med det er en påvirkning på utfallet av Y, når en tar hensyn til den uavhengige variabelen. Om utfallet av testen peker mot at de uavhengige variablene X ikke er signifikant, betyr dette at det ikke er tilstrekkelig med bevis som peker mot at X har et lineært forhold med Y, og med det at den ikke medfører en reduksjon eller økning i Y.

6.1 Er kjønn en signifikant faktor for avkastning på timelønn?

Vi har i denne oppgaven utført regresjonsanalysene ved bruk av dataprogrammet Stata. Under utførelsen av regresjonsanalysene, beregner og oppgir Stata selv om våre benevnede variabler er statistisk signifikante eller ikke. Dette markeres i “*Regresjonsanalyse 1*” og “*Regresjonsanalyse 2*” ved tegnene “***”. Stata har i dette tilfellet presisert at tre stjerner betyr at variablene er statistisk signifikante helt ned til 0.01, noe som tilsvarer en test av 1% signifikansnivå. Selv for dette ønsker vi å utrede og presentere en test for vår første del av del-problemstillingen.

Vi utfører en T-test for å se om variablene er statistisk signifikant, og begynner med å presentere en nullhypotese og en alternativ hypotese:

$$H^0: \beta^1 = 0$$

$$H^0: \beta^1 \neq 0$$

Nullhypotesen tilsier at det ikke er en lineær sammenheng mellom predikerte Y og variabelen “gender_b” i vår analyse. Selv om koeffisienten for norske kvinner tilsier en påvirkning på -0.153 på timelønn, er ikke denne signifikant overfor populasjonen. Den alternative hypotesen hevder at det er en lineær sammenheng, og at påvirkningen fra variabelen er signifikant. Vi regner ut verdien ved bruk av følgende ligning:

$$TS = \frac{\widehat{\beta}_1 - \beta_1}{se(\widehat{\beta}_1)} = \frac{\widehat{\beta}_1 - 0}{se(\widehat{\beta}_1)}$$

Ved bruk av denne ligningen tester vi om den ikke-predikerte variablene X har en lineær sammenheng med Y. I det tilfellet vil vi gi β_1 verdien 0, ettersom testen antar at den faktiske helningen på koeffisienten er 0. Den mest utbredte metoden i økonometri er ved bruken av et 5% signifikansnivå, noe som tilsvarer et “confidence level”, videre omtalt som konfidensnivå, på 95%. Dette betyr at vi vil 95% av tilfellene forvente å oppnå samme resultat av korrelasjonen for den lineære sammenhengen.

Ettersom vi har verdiene for koeffisienten og SE (*standard error*), trenger vi verdien for 0 i nullhypotesen. Ettersom vi bearbeider utregningen med et slingringsmonn på 0.05, da 5% signifikansnivå, må vi beregne Degrees of Freedom, frihetsgrader, videre omtalt som DF, og beregne ved bruk av en T-tabell vår kritiske t-verdi. DF beregnes ved bruk av følgende ligning:

$$DF = n - k - 1$$

I denne ligningen tilsier “*n*” antall observasjoner. Vi kan se i “*Tabell 1*” og “*Tabell 2*” at vi har ulike mengder observasjoner for norske og japanske kvinner, på 3628 og 3323. For “*k*” i ligningen setter vi inn antall variabler brukt i regresjonsanalysen. Vi har brukt 7 uavhengige variabler i vår regresjonsanalyse, og vil så beregne DF i to følgende ligninger:

$$DF = 3628 - 7 - 1 = 3620$$

$$DF = 3323 - 7 - 1 = 3315$$

Vi kan ved bruk av t-tabellen i *Introductory Econometrics* (Wooldridge, 2020, s.786) se at DF over 1000 oppnår samme verdi. Ved et signifikansnivå på 5% kan vi se at den kritiske t-verdien for en two-tailed test er: 1.962. Vi kan nå ta i bruk verdiene for å kalkulere om den uavhengige variabelen “*gender_b*” er statistisk signifikant på en 5% signifikansnivå:

T-test, Kvinner Japan:

$$TS = (-0.413 - 0)/0.0266 = -15.52631577$$

T-test, Kvinner Norge:

$$TS = (-0.153 - 0)/0.0160 = -9.5625$$

Vi kan med dette se at verken av verdiene ligger innenfor konfidensintervallet -1.962 og 1.962, og vi kan med det konkludere at den uavhengige variabelen er statistisk signifikant for både kvinner i Norge og Japan, og vi kan med det forkaste nullhypotesen og konkludere med at det er en lineær sammenheng.

6.2 Observasjon av signifikansnivå for resterende uavhengige variabler

Som nevnt i 6.1 redegjør Stata selv for signifikansnivået under utførelsen av regresjonsanalysen. Stata presenterer dette ved følgende verdier:

- *** $p < 0.01$: Tilsier et konfidensnivå på 99%
- ** $p < 0.05$: Tilsier et konfidensnivå på 95%
- * $p < 0.1$: Tilsier et konfidensnivå på 90%

I denne delen av oppgaven vil vi ikke utføre spesifikke utregninger for hver enkel variabel ettersom vi kan observere i regresjonsanalysen hvilket signifikansnivå hver variabel oppfyller, og om de med det er statistisk signifikant og har en lineær sammenheng opp mot vår avhengig variabel “timelønn”.

Vi kan først og fremst påpeke at påvirkningen av variabelen “gender_b”, altså kjønn, er statistisk signifikant i alle ledd av modellen for kvinner i Norge og Japan, dette med et konfidensnivå på 99%.

Vi leser av regresjonsanalysen produsert av Stata at nesten alle våre predikerte verdier er statistisk signifikant helt opp til et konfidensnivå på 99%, noe som betyr at de uavhengige variablene som brukes i vår analyse har en lineær sammenheng med timelønn. Det eneste tilfellet som skiller seg ut finner vi i “*Regresjonsanalyse 2*” for norske kvinner, der variabelen “pvlit1” er signifikant på en 95% konfidensnivå når den først introduseres i Model 3, men presenteres som statistisk ikke-signifikant videre i Model 4, og signifikant kun på et

konfidensnivå på 90% for Model 5. Vi tester dette ved samme fremgangsmåte som presentert i 6.1:

pvlit1 Model 4 - Norske kvinner:

$$TS = (0.000535 - 0)/0.000351 = 1.524216524$$

pvlit1 Model 5 - Norske kvinner:

$$TS = (0.000617 - 0)/0.000351 = 1.757834758$$

Vi kan ved disse beregningene se at begge resultatene befinner seg innenfor intervallet på -1.962 og 1.962, og vil med det regnes som statistisk ikke-signifikant for et signifikansnivå på 5%. Som Stata påpeker, opprettholder variabelen statistisk signifikant i "Model (5)" til et signifikansnivå på 10%, altså et konfidensnivå på 90%. Ved bruk av t-tabellen kan vi se at for en "two-tailed" test ligger DF 1.645 og -1.645.

Vi vil da forkaste nullhypotesen for Model 5, men beholde den for Model 4. Dette sier at det ikke er en lineær sammenheng mellom vår avhengig variabel og "pvlit1". Ettersom denne variabelen er signifikant i Model 3 og Model 5, kan vi ta hensyn til at resultatet for Model 4 kan skyldes andre faktorer som utvalgsfeil (sampling error), variabelens relevans eller utvalgsstørrelsen.

Ved bruk av egen utregning og Stata's analyse kan vi si at den største andelen av variablene brukt i vår analyse er statistisk signifikante og har med det en påvirkning på timelønn for kvinner i Norge og Japan.

7. Avslutning

I denne delen av oppgaven skal vi presentere en oppsummering av oppgaven og gitte konklusjoner mot den presenterte problemstillingen og delproblemstillingen forskningsprosjektet har undersøkt. Etter OECD publiserte sin artikkel “Japan Policy Brief: Inequality” i april 2015 om Japans svake rangering blant OECD landene, ønsket oppgaven å utforske om dette medførte et lønnsgap blant kjønnene, og om det kunne forklares på manglende, eller lav, avkastning på faktorer som utdanning, erfaring og ulike kunnskaper. Samtidig ble data fra Norge analysert i det samme forskningsspørsmål for å utrede et grunnlag for sammenligning. Oppgavens datamateriale ble innhentet fra OECD sin offisielle PIAAC undersøkelse og ga oppgaven grunnlag for å analysere data fra samme kilde. For å analysere påvirkningen våre uavhengige variabler hadde på timelønn, ble en logaritmisk regresjonsanalyse utført i dataprogrammet Stata, og to individuelle regresjonsanalyser ble utarbeidet. Disse presenterte resultater for kvinner i Japan og Norge, og ga datagrunnlag for å legge en konklusjon for oppgavens problemstilling. De ulike resultatene ble i programmet Stata testet som statistisk signifikante, men manuelle beregninger og diskusjon rundt resultatene ble i tillegg utført.

7.1 Konklusjon

Ved hjelp av “*Regresjonsanalyse 1*” og “*Regresjonsanalyse 2*” kan vi for det første konkludere med at det er et lønnsgap mellom mann og kvinne, både i Norge og i Japan. Hvis vi ser på avhengig variabel “timelønn” i Norge, vil vi se at dersom denne kun påvirkes av variabelen kjønn, vil det medføre en reduksjon på -15,3%. Ser vi på resultatene for Japan får vi et enda større skille med -41,3%. Vi kan med dette konkludere overfor vår hovedproblemstilling og si at det finnes lønnsforskjeller mellom kvinner og menn i Norge, og i Japan. Som henvist til i OECD sin rapport kan vi og se at skillet er betydelig større i Japan enn i Norge, selv for at det eksisterer innad i begge landene.

Videre stiller oppgaven spørsmålet om dette skillet kan forklares i form av forskjell i avkastning på våre uavhengige variabler. Under punkt 3.2.1 fremlegges deskriptiv statistikk fra PIAAC som redegjør for resultatene av våre uavhengige variabler. Vi kan der se at forskjellene mellom kjønnene i begge land er relativt like. Når vi ser hvordan disse utspiller seg i de ulike regresjonsanalysene, kan vi konkludere med at disse har en positiv effekt for begge land, og kan med det ikke forklare hvorfor et lønnsgap eksisterer.

Vi kan med dette henvise til at lønnsgapet eksisterer på bakgrunn av andre faktorer. I denne oppgaven har vi under punkt 2. redegjort for ulike teorier som kan forklare hvorfor lønnsgapet i Japan er i stor grad større enn lønnsgapet i Norge, selv for relativ lik avkastning på utdanning, erfaring og kunnskaper. Det vil til større grad fremstå som sannsynlig at lønnsgapet i Japan er større på bakgrunn av preferanseteori og et ulikt syn på humankapital. Dette kan reflekteres rundt på bakgrunn av tradisjonelle og kulturelle fremgangsmåter og tankeganger innenfor det japanske systemet.

7.2 Kritikk

Oppgaven har til sin beste evne utført databehandlingen og analysen på en grundig måte. Videre kunne oppgaven ha tatt for seg flere relevante variabler for å observere flere innvirkninger på resultatet. Regresjonsanalysene kunne blitt utført i samme datasett, og ikke i to individuelle, men med identiske fremgangsmåter. Dette kunne latt seg gjøre ved å implementere en dummy-variabel for de to ulike landene. Oppgavens forskere kunne også gjort seg mer kjent med japansk arbeidskultur og politikk.

7.3 Videre forskning

Videre forskning for å analysere lønnsgapet i Japan, sammenlignet med lønnsgapet i Norge, burde ta for seg flere samfunnsøkonomiske og politiske teorier. Fremtidige forskningsprosjekt burde legge mer fokus på japansk struktur og arbeidsliv for å bedre forklare lønnsgapet, ettersom statistisk data ikke kan forklare problemet. Om en empirisk undersøkelse av data videre ønskes å utforskes, anbefales det å ta hensyn til et bredere og større utvalg av data.

8. Litteraturliste

1. Becker, G. S. (1962). *Investment in human capital: A theoretical analysis*. *Journal of Political Economy*, 70*(5, Part 2), 9-49. <https://www.jstor.org/stable/pdf/1829103.pdf>
2. Bjørnnes, A. K., & Gjevjon, E. R. (2019). *Kvalitet i kvantitativ metode - et innblikk*. *Sykepleien Forskning*, 14*(78806), e-78806. <https://doi.org/10.4220/Sykepleienf.2019.78806>
3. Carson, S. G., & Kosberg, K. (2018). *Etikk- teori og praksis* (4th ed.)*. Cappelen Damm.
4. *Encyclopedia Britannica*. (2024, April 21). Norway. <https://www.britannica.com/place/Norway>
5. Fløtre, I. A. (2024). *Slik kan lønnsforskjellen mellom kvinner og menn forklares*. Statistisk sentralbyrå. <https://www.ssb.no/arbeid-og-lonn/lonn-og-arbeidskraftkostnader/statistikk/lonn/artikler/slik-kan-lonnsforskjellen-mellom-kvinner-og-menn-forklares>
6. Grini, K. H., & Fløtre, I. A. (2023). *Lønnsgapet i Norge*. Statistisk sentralbyrå. https://www.ssb.no/arbeid-og-lonn/lonn-og-arbeidskraftkostnader/artikler/lonnsgapet-i-norge/_attachment/inline/d5758cea-dd26-434e-b0c0-b73bc999b855:559d2fc99dde9e0cfc7adf68c20d0d2b23becdae/RAPP2023-04_medlenker.pdf
7. Hakim, C. (2000). *Work-lifestyle choices in the 21st century: Preference theory**. Oxford University Press.
8. Holden, S. (2016). *Makroøkonomi* (1st ed.)*. Cappelen Damm Akademisk.
9. Masai, Y., Latz, G., & Hijino, S. (2024, February 23). *Economy of Japan*. *Encyclopaedia Britannica**. <https://www.britannica.com/money/economy-of-Japan>

10. OECD. (2015). *Inequality, greater gender equality for more inclusive growth*. <https://www.oecd.org/policy-briefs/japan--greater-gender-equality-for-more-inclusive-growth.pdf>
11. OECD. (2016). *PIAAC-Undersøkelse Norge og Japan*. <https://webfs.oecd.org/piaac/puf-data/CSV/>
12. OECD. (2023). *Joining forces for gender equality*. <https://www.oecd.org/japan/Gender2023-JPN-En.pdf>
13. Ringdal, K. (2018). **Enhet og mangfold - samfunnsvitenskapelig forskning og kvantitativ metode** (4th ed.). Vigmostad & Bjørke AS.
14. Statista. (2023, desember). *Japan: Distribution of gross domestic product (GDP) across economic sectors from 2011 to 2021 [Dataark]*. Hentet fra <https://www.statista.com/statistics/270093/distribution-of-gross-domestic-product-gdp-across-economic-sectors-in-japan/>
15. Statista. (2023, desember). *Norway: Share of economic sectors in the gross domestic product (GDP) from 2012 to 2022 [Dataark]*. Hentet fra <https://www.statista.com/statistics/327233/share-of-economic-sectors-in-the-gdp-in-norway/#:~:text=This%20statistic%20shows%20the%20share,sector%20contributed%20about%2042.15%20percent.>
16. Svennebye, L. (2013). *Kjøpekraftspariteter, prisnivåindekser og prisnivåjustering: En brukerveiledning*. I Statistisk sentralbyrå. https://www.ssb.no/priser-og-prisindekser/artikler-og-publikasjoner/_attachment/143082?_ts=141a6c28560
17. Wooldridge, J. M. (2020). *Introductory to Econometrics: A Modern Approach* (7. utg.). Cengage Learning.
18. Zahidi, S. (2023, 20. juni). *Global Gender Gap Report 2023*. World Economic Forum. Hentet fra <https://www.weforum.org/publications/global-gender-gap-report-2023/in-full/>

