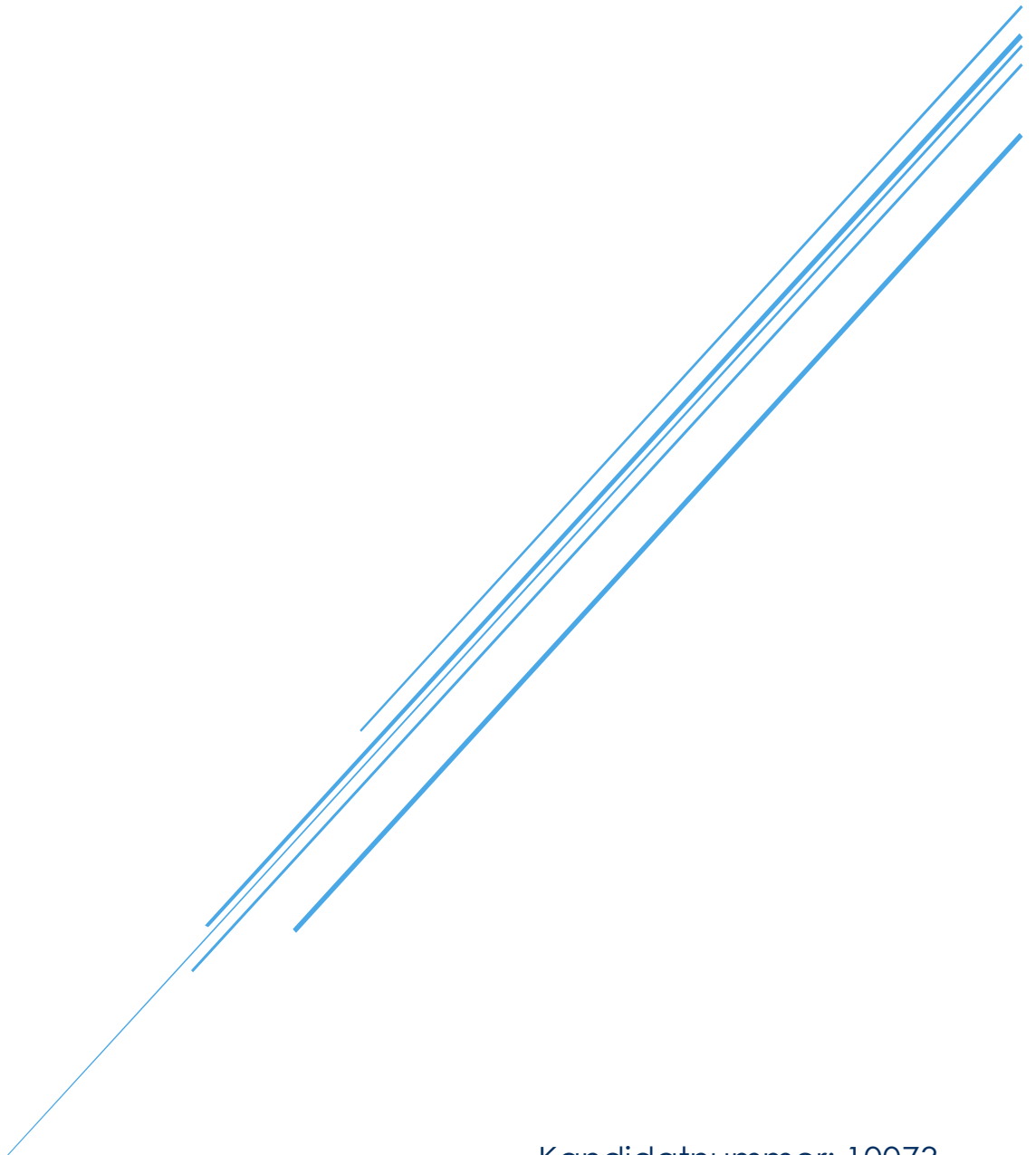


INNVANDRERE OG MAJORITETSBEFOLKNINGEN – LØNNSGAPET OG DETS ÅRSAKER

SØK2901 - Bacheloroppgave

Institutt for Samfunnsøkonomi ved NTNU



Kandidatnummer: 10073

Vår 2019

Innholdsfortegnelse

1. Innledning	2
1.1. Motivasjon	
1.2. Problemstillinger	
2. Teoretisk rammeverk og tidligere litteratur	3
2.1. Teoretisk rammeverk	
2.2. Tidligere litteratur	
3. Presentasjon av datasett og modell	7
3.1. Datamaterialet og variabler	
3.2. Deskriptiv statistikk	
3.3. Økonometrisk modell	
4. Empiriske resultater	12
4.1. Hovedresultater	
4.2. Avvik fra forventede resultater	
5. Oppsummering og konklusjon	17
6. Litteraturliste	19

1. Innledning

1.1. *Motivasjon*

Etter en periode med flere immigrasjonsbølger, fri flyt av mennesker gjennom EØS og en stadig globalisering av arbeidsmarkedet er det viktig å se på om lønnsbildet i mottakerlandene¹ er preget av diskriminering og systematisk forskjellsbehandling. Ifølge Statistisk Sentralbyrå er det selv i Norge, et av verdens antatt beste land å bo i, et lønnsgap mellom norskfødte med innvandrerbakgrunn og arbeidere uten innvandrerbakgrunn (Epland & Kirkeberg, 2018). I resten av Europa eksisterer de samme tendensene, men bakgrunnen for lønnsgapet varierer fra en studie til en annen (Tverdostup & Paas, 2018). Både statistisk diskriminering, ulik avkastning på utdanning, valg av yrke og kognitive ferdigheter er blant mange av de foreslåtte årsakene til den negative lønnsavkastningen av å være utledning.

Det som er interessant å undersøke, spesielt med tanke på at innvandrerbefolkningen i de fleste land antas å øke i fremtiden, er om størrelsen på innvandrerbefolkningen i landet kan være en forklaring på lønnsgapet mellom majoritetsbefolkningen og innvandrere. Dersom andelen innvandrere i et land viser seg å påvirke størrelsen på lønnsgapet, kan dette da knyttes opp mot en form for fremmedfrykt, eller er det andre forklaringsvariabler som kan skissere en slik sammenheng? En annen mulig forklaringsfaktor som er spennende å analysere, er om avkastningen av foreldrenes utdanning er ulik for innvandrere og majoritetsbefolkningen. Ifølge SSB tar barn av foreldre med høyere utdanning i mye større grad høyere utdanning sammenliknet med barn av foreldre med lavere grader utdanning (Statistisk Sentralbyrå, 2017). Høyere utdanning for individene antas videre å øke lønnsnivået. Samtidig kan høyt utdannede foreldre antas å ha relativt gode jobber, og kan potensielt være individets vei inn i arbeidsmarkedet. Spørsmålet som da kan undersøkes, er om utdanningen til innvandrernes foreldre i mindre grad gir en positiv avkastning på deres egen lønn.

1.2. *Problemstillinger*

- (i) Kan det observeres et lønnsgap mellom majoritetsbefolkningen og innvandrere, og hvis så: er det noen sammenheng mellom størrelsen på lønnsgapet og andelen innvandrere i befolkningen?
- (ii) Kan innvandrerenes utdanningsnivå og lesekunnskaper forklare noe av lønnsgapet?

¹ Med mottakerlandene menes de landene et individ har immigrert til.

- (iii) Kan innvandrernes foreldre sitt utdanningsnivå være en forklaringsfaktor for et eventuelt lønnsgap mellom innvandrere og majoritetsbefolkningen, og har foreldrenes utdanningsnivå forskjellig avkastning på lønn for innvandrere enn for majoritetsbefolkningen?

2. Teoretisk rammeverk og tidligere litteratur

2.1. Teoretisk rammeverk

Å definere «forskjeller» kan ofte være en komplisert jobb. Snakkes det om forskjeller som to ting som er ulike, vil begrepet omfatte det meste av menneskers forestillinger. Den vanligste bruken av ordet i relasjon til lønn er å sammenlikne gjennomsnittslønnen hos to eller flere undergrupper, og deretter bevege seg dypere inn i årsakene til et eventuelt lønnsgap. Poenget med dette er å presentere en rekke forutsetninger, og undersøke hver forutsetning enkeltvis gitt alt annet likt. På den måten kan det som anses som rasjonelle grunner til lønnsforskjeller skilles fra de årsakene som er vanskelige å begrunne med økonomiske modeller.

Den vanligste årsaksforklaringen til et lønnsgap, uavhengig av hvilke grupper som sammenliknes, er forskjeller i produktivitet eller personlige egenskaper. Dette samles under begrepet humankapital. Når lønnsforskjeller mellom innvandrere og majoritetsbefolkningen studeres, er det spesielt viktig å se på lokasjonsspesifikk humankapital. Det vil si humankapital knyttet til et spesifikt sted, og da er spesielt språk en viktig lokasjonsspesifikk egenskap som kan påvirke ansettelses- og lønnsforhold. Flere undersøkelser, som presenteres i neste del, viser at lønnsforskjellene avhenger av hvor innvandrerne er fra, og hvor mye språk og kultur skiller seg fra mottakerlandet. Samtidig åpner lokalspesifikk humankapital for et lønnsgap som kan reduseres over tid, avhengig av hvor godt innvandrerne integreres i samfunnet. Det vil da observeres et tilfelle med lønnskonvergens mellom innvandrere og majoritetsbefolkningen, grunnet en økning i det relative lønnsnivået hos innvandrerbefolkningen. Likevel har ikke forskjeller i humankapital vist seg å være den eneste forklaringen for lønnsforskjeller, da undersøkelser som korrigerer for egenskaper hos arbeiderne fortsatt viser at lønnsgapet består. Dette indikerer at integrering og konvergens på lang sikt ikke vil eliminere lønnsforskjeller mellom majoritetsbefolkningen og innvandrerbefolkningen (Grand & Szulkin, 2002).

Statistisk diskriminering er, sammen med teorier om humankapital, en bredt diskutert og anerkjent årsaksforklaring for lønnsgap. Det går ut på at forventningene til en arbeider baserer seg på en gjennomsnittsadferd knyttet til den gruppen de tilhører. Dette kan være med på å både overvurdere en ansatt, for eksempel å ansette en mindre produktiv person fordi de tilhører en gruppe som antas å være høyproduktive, eller å undervurdere en ansatt. Sistnevnte konsept kan utfolde seg på flere ulike måter, men essensen av statistisk diskriminering er likevel at personlige egenskaper (humankapital) settes til side til fordel for forutsatte egenskaper gitt undergruppen individet tilhører. Et vanlig eksempel på dette er at dersom en arbeidsgiver skal velge mellom to kandidater med like kvalifikasjoner, men som har ulike nasjonaliteter, vil arbeidsgiveren velge den kandidaten som er nærmest deres egen nasjonalitet (Moffat, 2019). Statistisk diskriminering er derfor viktig i analyser av sysselsettingen blant innvandrere og majoritetsbefolkningen, men det kan også trekkes opp mot lønnsforskjeller i den grad det fører til ulik avkastning på humankapital. I dette ligger en teori om at arbeidsgivere vil verdsette samme type egenskap, som utdanningsnivå, hos en innfødt høyere enn samme egenskap for en innvandrere. Dette kan være forskjellig avkastning på likestilte utdanninger, der den ene er fullført i mottakerlandet og den andre er fullført i et annet land, eller forskjellig avkastning på utdanninger gjennomført i mottakerlandet. Derfor kan lokasjonsspesifikk humankapital gi grunnlag for statistisk diskriminering, ved at samme egenskaper forskjellsbehandles fordi arbeidsgiveren antas å ha mer kunnskap om kvaliteten på utdanningen i mottakerlandet. Samtidig åpner statistisk diskriminering for en «ond» sirkel for minoritetsgrupper i arbeidsmarkedet. Dette er fordi teorien åpner for at statistisk diskriminering kan påvirke lønnsdannelsen på to måter: den ene gjennom arbeidsgiver, som presentert over, og den andre gjennom arbeidstakere som lar være å investere i ny humankapital fordi den forventede avkastningen på investeringen antas å være lavere enn hos andre arbeidere. Siden arbeidstakere i undergrupper ikke alltid motarbeider fordommene mot egen minoritet, fortsetter arbeidsgivere å bruke et gjennomsnitt som potensielt undervurderer produktive arbeidstakere. Slik kan statistisk diskriminering fortsette å skape et sysselsettings- og lønnsgap mellom innvandrere og majoritetsbefolkningen (Moffat, 2019).

Likevel skiller statistisk diskriminering seg fra andre teorier om diskriminering, fordi aktørene anses å være objektive og profittmaksimerende. En arbeidsgiver kan derfor være rasjonell, på tross av at de bedriver statistisk diskriminering (Moffat, 2019). Økonomisk teori om diskriminering som grunnlag for et lønnsgap, og andre barrierer i arbeidsmarkedet for minoritetsgrupper, er fortsatt relevant i dagens debatt om lønnsforskjeller. Gjerne trekkes

diskriminering frem som et argument der modellens forklaringsvariabler ikke kan begrunne hele lønnsgapet. Problemet med diskriminering i økonomisk forstand, er at det kan være vanskelig å definere, og være vanskelig å skille fra mindre betydelige fordommer. Ifølge Glenn Cain (1986), er det viktig å ikke la fordommer definere begrepet diskriminering, da fordommer er en holdning og diskriminering i større grad er resultatet av adferd og systematisk forskjellsbehandling (Cain, 1986, s.695). I nyklassisk økonomisk teori skilles det mellom diskriminering fra forbrukere, kolleger og arbeidsgivere. Essensen i de to førstnevnte ligger i at å ansette en minoritet vil påføre bedriften en kostnad. Enten ved at forbrukere anser produktet som mindreverdig og krever lavere pris for gitt produkt, eller ved at kollegaer krever en kompensasjon i lønn for å arbeide sammen med kolleger fra minoritetsgrupper. Diskriminering fra arbeidsgivere har et noe mer tvetydelig resultat, men i relasjon til lønn antas det at arbeidsgivere vil kunne sette en lavere lønnsats for arbeidere med minoritetsbakgrunn, siden etterspørselen etter minoritetsarbeidere av ulike grunner er redusert (Cain, 1986, s.713). Bakgrunnen for diskriminering kan være mange, og vil derfor være vanskeligere å kunne definere som rasjonelle. Dette er sammenliknet med statistisk diskriminering, som i større grad baserer seg på profittmaksimering enn generelle fordommer mot en minoritet. Et eksempel på dette skillet kan illustreres med lønnsgapet mellom kjønn. Statistisk diskriminering vil være å velge vekk en kvinnelig kandidat, fordi kvinner antas å få barn og derfor måtte ta permisjon en gang i løpet av ansettelsesperioden. Ren diskriminering vil være å ikke ansette kvinnen fordi arbeidsgiveren mener kvinnen ikke egner seg i et tradisjonelt mannsdominert yrke.

En siste faktor som kan påvirke lønnsgapet mellom innvandrerbefolkningen og majoritetsbefolkningen, er at de to gruppene systematisk sysselsetter seg i ulike stillinger og yrker. Da antas det at innvandrere i større grad søker seg til lavere betalte yrker og i mindre grad forsøker å klatre oppover i rangstigen. Dette kan igjen trekkes opp mot forventninger om systematisk diskriminering.

2.2. Tidligere litteratur

Økonomisk teori gjør gjerne en del forutsetninger som ikke nødvendigvis er realistiske nok til å beskrive virkeligheten. Likevel danner det grunnlaget for å teste ulike teorier om lønnsforskjeller. Generelt viser tallene at et lønnsgap mellom innvandrere og

majoritetsbefolkningen eksisterer, og derfor fokuserer de fleste studier på hvordan det allerede eksisterende lønnsgapet endres i takt med ulike forklaringsvariabler.

I Sverige og Irland har undersøkelser gitt resultater som peker på at forskjeller i humankapital forklarer noe av lønnsforskjellene mellom innvandrerbefolkningen og majoritetsbefolkningen. Ifølge Barrett og McCarthy (2007), er det en stor forskjell i lønnsgapet mellom engelsktalende og ikke-engelsktalende innvandrere relativt til majoritetsbefolkningen, med lønnsgap på henholdsvis 3% og 31% (Barrett & McCarthy, 2007, s.799). Dette viser at lokasjonsspesifikk humankapital har mye å si for lønnsnivået til innvandrerbefolkningen. I Sverige ble det også funnet forskjeller i lønnsgapet mellom vestlige og ikke-europeiske innvandrere, som kan trekkes opp mot både språkbarrierer og kulturforskjeller (Grand & Szulkin, 2002). Samme undersøkelse testet om det relative lønnsnivået hos innvandrere økte med antall år innvandrerne bodde i Sverige. Resultatet viste at lønnsgapet mellom innvandrere og sammenliknbare innfødte² reduseres i takt med antall år innvandrerens hadde bodd i mottakerlandet. Likevel ble ikke lønnsgapet helt eliminert, og forfatterne konkluderte med at forskjeller i humankapital ikke er eneste grunnen til et lønnsgap. De valgte å se på diskriminering som en mulig forklaringsfaktor for lønnsgapet gitt at innvandrerne hadde tilegnet seg lokasjonsspesifikk humankapital.

I Sverige har det også blitt studert hvorvidt andelen innvandrere som er arbeidstakere og arbeidsgivere i en bedrift kan påvirke lønnsgapet mellom innvandrere og majoritetsbefolkningen. Analysen som ble gjennomført av Tomaskovic-Devey, Hällsten, og Avent-Holt i 2015 testet hypoteser om lønnsgapet ble redusert eller ikke, når andelen innvandrere i arbeid økte og når andelen utenlandske sjefer økte. Her kom de frem til at i arbeidsplasser preget av manuelt arbeid³ vil en økt andel utenlandske arbeidere føre til et redusert lønnsgap, mens det i arbeidsplasser preget av administrativt arbeid⁴ er økningen i andelen utenlandske sjefer som bidrar til reduserte lønnsforskjeller. Undersøkelsen kom også frem til at mye av lønnsgapet kan forklares ved at innvandrere og innfødte fordeler seg ulikt i jobbmarkedet, både med tanke på yrke og stilling. Likevel avhenger effekten av dette av hvilke typer yrkessektorer som studeres. I tillegg fant forfatterne bevis for at i yrkessektorer med store forskjeller mellom innvandrere og majoritetsbefolkningen, var sannsynligheten for utnyttelse av arbeidskraft på bekostning av innvandrerne større (Tomaskovic-Devey, Hällsten,

² Med sammenliknbare, menes arbeidere som antas å ha samme gitte egenskaper utenom de som ønskes undersøkt. Dette kan være samme produktivitet, alder, utdanningsnivå o.l.

³ I kilden referert til som blå-krage-arbeidere

⁴ I kilden referert til som hvit-krage-arbeidere

& Avent-Holt, 2015). Dette kan støtte teorier om statistisk diskriminering, og vil kunne være et referansepunkt i analysen om lønnsgapet mellom innvandrere og majoritetsbefolkningen avhenger av størrelsen på innvandrerbefolkningen i et land.

3. Presentasjon av datasett og modell

3.1. Datamaterialet og variabler

I denne oppgaven brukes datamateriale fra en undersøkelse gjennomført av OECD i 2011-2012. Undersøkelsen kalles PIAAC (Programme for the International Assessment of Adult Competencies), og ble gjennomført for å se på forskjeller i kognitive ferdigheter blant befolkningen i 21 land. For å kunne sammenlikne lønnsforskjeller i ulike land, er det hensiktsmessig å bruke en PPP-justert variabel for lønn. I datasettet er dette variabelen *earnhrppp*, som viser timelønnen til et individ i amerikanske dollar justert for prisnivå i landene. I modellen brukt i denne oppgaven, defineres en ny lønnsvariabel. Den er en logaritmisk transformasjon av *earnhrppp*, og er betegnet som *lwage*.

I modellen er datasettet begrenset til å kun gjelde individer som er minst 25 år eller eldre, for å unngå feil i beregningene ved å inkludere individer som fortsatt er under utdanning. For å se på effekten av å være innvandrer brukes dummyvariabelen *migrant_first*. Dersom individet er innvandrer vil variabelen være satt til 1, og hvis ikke er variabelen satt til 0. Samtidig sammenliknes tre land med ulike innvandrerandeler i befolkningen, for å kunne besvare hypotesen om hvorvidt andelen innvandrere i et land påvirker et eventuelt lønnsgap mellom innvandrere og majoritetsbefolkningen. Landene Irland, Frankrike og Tsjekkia representerer land med ulike andeler innvandrere, med henholdsvis stor (16.5%), middels (10%) og liten (3%) prosentandel innvandrere i befolkningen. Antar da at utvalget er representativt for faktisk andel innvandrere i landene. Modellen differensierer mellom landene ved å gjøre Tsjekkia og Frankrike om til dummy-variable, og den har derfor Irland som referanseland. Dette betyr at resultatene i den estimerte modellen vil gi koeffisienter som forteller noe om variablene i Tsjekkia og Frankrike sammenliknet med verdien på de samme variablene i Irland. Ellers vil variabler for vært land inneholde enten «Ire», «Fra» eller «Tsj». I studiet om lønnsgapet er forskjellig i de ulike landene brukes variablene *Fraxmigrant* og *Tjsxmigrant*, som viser sammenhengen mellom å være innvandrer i henholdsvis Frankrike og Tsjekkia.

I tillegg til å skille mellom landene, analyseres effekten av individets utdanningsnivå og leseferdigheter på lønnsnivået deres. Utdanningsnivået representeres med variabelen *yrssqual*, som viser antall år utdanning individet har gjennomført. Variabelen for leseferdigheter, *litscore1*, er en normalfordelt variabel med gjennomsnitt rundt null. Null representerer da gjennomsnittlig poengsum for hele datasettet, og verdier på venstre side av midten vil representere leseferdigheter som er lavere enn gjennomsnittet. Tilsvarende vil verdier på høyre siden av gjennomsnittet representere høyere leseferdigheter enn snittet. En økning i utdanningsnivå vil defineres som 1 års ekstra utdanning, mens en endring i leseferdigheter vil være et standardavviks endring i ferdighetsnivå.

Ved undersøkelsen foreldrenes utdanning vil variabelen *pared* bli brukt, som er delt i indikatorverdier fra 1 til 3. 1 representer at ingen foreldre har mer enn grunnskoleutdanning, 2 representerer at minst en forelder har fullført videregående skole og påbygg, og 3 representerer at minst en forelder har fullført høyere utdanning. Her brukes også dummyvariable for høy utdanning (3) og middels utdanning (2) for å skille mellom utdanningsnivåene. Disse defineres som henholdsvis *pared_high* og *pared_mid*. Her vil lavt utdanningsnivå hos foreldrene (1) være referanselikningen til modellen, og derfor være den faste verdien som *pared_high* og *pared_mid* sammenliknes med. Til slutt fanges avkastningen av foreldrenes utdanning på innvandrernes lønn opp av interaksjonsleddet *paredxIm*, som er variabelen *pared* multiplisert med variabelen for å være immigrant, *migrant_first*.

3.2.Deskriptiv statistikk

Videre vil deskriptiv statistikk for den avhengige variabelen (lønn) presenteres, og deretter følger liknende statistikk for de ulike forklaringsvariablene.

	Alle	Alle utlendinger	Utlendinger Irland	Utlendinger Frankrike	Utlendinger Tsjekkia
Gjennomsnittlig timelønn (PPP US\$)	16,05	16,53	19,61	14,01	8,80
Minimum (PPP US\$)	0,16	2,33	5,73	4,01	2,33
Maksimum (PPP US\$)	74,47	73,88	73,88	45,71	33,35
Standardavvik	9,78	9,55	10,71	6,39	5,70
Antall observasjoner	8 174	853	447	337	69

TABELL 1

Tabell1 viser at andelen observasjoner av utlendinger utgjør en liten del av totalt antall observasjoner i datasettet. Dette kan potensielt påvirke resultatene av regresjonsanalysen, siden fordelingen av observasjonene mellom landene også er skjev. Likevel er det ikke for få observasjoner til at resultatene i analysen kan antas å være representative for virkelighetsbildet.

Videre viser tabellen at gjennomsnittlig timelønn for alle innvandrere er noe høyere enn for alle observasjonene (uavhengig av nasjonalitet), og at hverken individet med den høyeste timelønnen eller individet med den laveste timelønnen er innvandrere. I tillegg vises det at den høyeste gjennomsnittslønnen for innvandrere blant de tre landene er i Irland, og den laveste er i Tsjekkia. Samtidig er lønnsspredningen høyere i Tsjekkia enn i Irland. Frankrike ligger omtrent midt imellom Irland og Tsjekkia for alle lønnskategoriene. Dette gir allerede en indikasjon på at andelen innvandrere i et land kan ha noe å si på lønnsnivået til innvandrerbefolkningen i landet.

Videre må det presenteres data for forklaringsvariablene i modellen, og derfor er gjennomsnittsverdier for utdanningsnivå, leseferdigheter, foreldrenes utdanningsnivå og andel innvandrere for hvert av landene presentert i Tabell 2. I tabellen er verdiene for Irland, Frankrike og Tsjekkia begrenset til å gjelde innvandrerbefolkningen i disse landene.

	Alle utlendinger		Irland		Frankrike		Tsjekkia	
	Gj.snitt	Std.avv	Gj.snitt	Std.avv	Gj.snitt	Std.avv	Gj.snitt	Std.avv
Antall år utdanning (yrsqual)	13,36	4,60	15,96	2,82	10,03	4,53	13,5	3,35
Antall observasjoner utdanning	1707		860		681		166	
Foreldrenes utdanningsnivå (Pared)	1,81	0,83	2,08	0,82	1,42	0,72	1,95	0,73
Leseferdigheter (litscore1)	-0,38	1,21	-0,049	1,08	-0,82	1,22	-0,39	1,20
Innvandrerandel (prosent)	10, 2%		16, 56%		10,77 %		3,3%	

TABELL 2

Et merkbart resultat, er at andelen innvandrere over 25 år som har utdanning er mye høyere enn andelen innvandrere som mottar lønn. Det er altså en del utdannede innvandrere som står utenfor arbeidsmarkedet. Samtidig viser tabellen at Frankrike har lavere gjennomsnittsutdanning både for innvandrerne og deres foreldre, og de har lavest gjennomsnittlig leseferdigheter blant innvandrerbefolkningen. Fra tabellen er det tydelig at ingen av innvandrergruppene har poengsummer høyere enn gjennomsnittet, men i Irland er

leseferdighetene nærmere 0 enn i både Frankrike og Tsjekia. Det kan da konkluderes med at språkferdighetene for innvandrere i Irland er betraktelig høyere enn i to andre landene. Samtidig er leseferdighetene nesten dobbelt så dårlig blant franske innvandrere sammenliknet med tsjekkiske innvandrere. Når det kommer til utdanning, har irske innvandrere i gjennomsnitt noe universitetsutdanning (> 13 år), tsjekkiske innvandrere har kompetanse fra videregående skole (10-13 år) og franske innvandrere har i gjennomsnitt kun grunnutdanning. Klassifiseringen av utdanningsnivå er her gjort etter den norske modellen, og det kan variere noe innad i hvert enkelt landet.

Foreldrenes utdanning er høyest i Irland, der er gjennomsnittet marginalt over middels utdanningsnivå. I Tsjekia er utdanningen marginalt under middels utdanningsnivå, og er derfor nesten like høyt som foreldrenes utdanningsnivå i Irland. Frankrike på sin side, har lave verdier på utdanningsnivå også for foreldrene til individene i datasettet. Her er gjennomsnittet nærmere grunnskolenivå enn videregående skole.

Oppsummert er det tydelig at lønnsnivået øker for andelen innvandrere i utvalget, men at utdanningsnivået hos innvandrerbefolkningen er lavere i Frankrike enn i Tsjekia og Irland. Denne sammenlikningen gjelder også for hele befolkningen⁵, dersom gjennomsnittlig utdanningsnivå for alle observasjoner i hvert land inkluderes. Altså er det ikke kun innvandrerne som har lavere utdanning i Frankrike, men hele befolkningen preges av færre års i utdanning enn i de andre landene.

3.3. Økonometrisk modell

For å se på effekten de ulike forklaringsvariablene har på lønnen til et individ, kan det tas i bruk en økonometrisk modell for å estimere en lønnslikning. I denne analysen brukes en log-lineær-modell, det vil si at den avhengige variabelen for lønn transformeres til en logaritmsk variabel. Dette gjør det mulig å sammenlikne ulike studier på tvers av land fordi parameteren foran variabelen *migrant_first* i en log-lineær-modell kan, multiplisert med 100, tolkes som det prosentvise lønnsgapet mellom innvandrere og majoritetsbefolkningen.

⁵ Dette er ikke vist i tabellen vår, men har blitt undersøkt i STATA for å kunne sammenlikne.

I analysen vil alle kontrollvariabler være inkludert i én lønnslikning, og etter å ha estimert denne vil det gjennomføres hypotesetester for å avgjøre om variablene er signifikante for modellen. Den økonometriske modellen blir som følger:

$$lwage = \beta_1 + \beta_2 migrant_{first} + \beta_3 yrsqual + \beta_4 litscore1 + \delta_1 Frankrike + \delta_2 Tsjekia + \delta_3 FraxMigrant + \delta_4 TjsxMigrant + \gamma_1 Pared_{high} + \gamma_2 Pared_{mid} + \gamma_3 ParedxMigrant + \varepsilon_i$$

Parameteren foran interesseparameteren *migrant_first* (β_2) forteller hvor stor forskjellen i lønn mellom innvandrere og majoritetsbefolkningen er, gitt et nivå på de andre kontrollvariablene. Parameteren foran *ParedxIm* viser forskjellen i avkastning på foreldrenes utdanning for innvandrere sammenliknet med majoritetsbefolkningen, og variablene *Fraxmigrant* og *TjsxMigrant* viser størrelsen på lønnsgapet i henholdsvis Frankrike og Tsjekia sammenliknet med lønnsgapet i Irland. De andre parameterne, multiplisert med 100, gir den prosentvise endringen i lønn gitt en endring i den gjeldende kontrollvariabelen. Andre variabler som kan påvirke lønnsgapet, som ikke er inkludert i modellen, fanges opp av det stokastiske restleddet ε_i .

Målet med analysen er å finne ut er hvor stort lønnsgapet mellom innvandrere og majoritetsbefolkningen er når det tas hensyn til en rekke kontrollvariabler. Samtidig er det også et mål å teste om størrelsen på innvandrerbefolkningen i et land påvirker lønnsgapet mellom innvandrere og majoritetsbefolkningen. Til slutt skal sammenhengen mellom lønnsgapet og utdanningsnivået til immigrantene og deres foreldre analyseres. Dette vil gjennomføres ved å estimere lønnslikningen ved hjelp av minste kvadraters-metode. Ved bruk av minste kvadraters-metode, forsøkes det å estimere en lineær sammenheng mellom avhengige og uavhengige variabler som minimerer gapet mellom regresjonslinja og de faktiske observasjonene. Da kan det estimeres verdier for parameterne i lønnslikningen, og videre tolkes hva disse tallverdiene forteller om relasjonen mellom lønn, innvandrerbefolkningen og innvandrernes humankapital.

For å analysere en estimert modell med flere variabler, er det nødvendig å ha en referanselikning. Dette vil være lønnsrelasjonen dersom parameterne for kontrollvariablene er lik null. I denne modellen er Irland referanseland, og variablene for de andre landene vil derfor sammenliknes med nivåene for de samme variablene i Irland. På samme måte er lav

utdanning på foreldrene, $pared = 1$, referanse for foreldrenes utdanningsnivå. Variablene for $pared$ vil derfor stå i forhold til individer med foreldre med lavt utdanningsnivå.

4. Empiriske resultater

4.1.Hovedresultater

Først presenteres den estimerte lønnsrelasjonen, der OLS-estimeringen gjøres i STATA. I tillegg til den presenterte modellen, viser Figur1 tre restriktive modeller med færre forklaringsvariabler. Å sammenlikne de restriktive modellene med den originale lønnsrelasjonen, vil gi en indikasjon på om alle de inkluderte forklaringsvariablene er relevante for lønnen.

```
. estout est3 est4 est2 est1, stats(r2)
```

	est3 b	est4 b	est2 b	est1 b
migrant_fi~t	.0531903	.0825993	-.2123858	-.1700022
yrsqual	.0610541	.0524093	.048531	.0495082
litscore1		.0660898	.058843	.057996
Frankrike			-.184218	-.1756701
Tsjekkia			-.8584037	-.8366047
Fraxmigrant			.229695	.2186315
Tjsxmigrant			.1484019	.1348867
pared_high				.0209909
pared_mid				-.0329749
paredxIm				-.0223891
_cons	1.775391	1.883521	2.269963	2.262243
r2	.140196	.1499066	.5028277	.508106

FIGUR 1

Den estimerte lønnsrelasjonen blir følgende:

$$lwage = 2.262 - 0.170migrant_{first} + 0,050yrsqual + 0.058litscore1 - 0,176Frankrike - 0,837Tsjekkia + 0,219FraxMigrant + 0,135TsjxMigrant + 0.021Pared_{high} - 0.033Pared_{mid} - 0.022ParedxIm$$

De tre første (est3, est4 og est2) i Figur1 har alle har mindre forklaringskraft enn lønnsmodellen som analyseres i denne oppgaven. Dette vises med størrelsen på residualkvadratsummen, r^2 , som er inkludert på nederste linje i figuren. Det observeres

samtidig store forskjeller mellom forklaringskraften til modellene som tar hensyn til land og modellene som kun ser på mål for humankapital. Variablene for land og foreldrenes utdanningsnivå kan derfor antas å ha mye å si på hvor representativ lønnsmodellen er for utvalget. Videre viser modellene som kun ser på utdanning og leseferdigheter et lønnsgap mellom innvandrere og majoritetsbefolkningen som er positiv. Dette diskuteres videre i del 4.2. For de utvidede modellene er lønnsnivået som forventet negativt, med henholdsvis -21% og -17%. Det betyr at et lønnsgap mellom innvandrerne og majoritetsbefolkningen eksisterer. Korrigert for alle kontrollvariablene som oppgaven, reduseres altså lønnsgapet mellom innvandrere og majoritetsbefolkningen. Dette indikerer at noe av lønnsgapet kan forklares med forskjeller i foreldrenes utdanningsnivå.

Hvorvidt lønnsnivået er forskjellig mellom innvandrere og majoritetsbefolkningen kan også analyseres ved bruk av en t-test. Dette går ut på å teste om parameteren foran *migrant_first*, β_2 , er lik null eller ikke. Dersom nullhypotesen, H_0 , forkastes, betyr det at variabelen som indikerer at individet er innvandrer eller ikke er statistisk signifikant for individets lønnsnivå. En metode for å finne ut om nullhypotesen skal forkastes eller ikke, er ved bruk av p-verdier. P-verdien viser det laveste signifikansnivået der en nullhypotese kan forkastes. Signifikansnivået på sin side viser sannsynligheten for å feilaktig forkaste nullhypotesen. Setter signifikansnivået for testen til 5%, og setter opp følgende hypoteser:

$$H_0: \beta_2 = 0 \qquad H_A: \beta_2 \neq 0$$

Ved bruk av test-funksjonen i STATA regnes p-verdien ut til å være 0.0002. Det betyr at ved et 5% signifikansnivå kan det bekreftes at lønnsnivået er forskjellig mellom innvandrere og majoritetsbefolkningen.

Det er relevant for oppgaven studere effekten av utdanning og leseferdigheter på lønnen. Den estimerte modellen viser at dersom utdanningen øker med et år, vil lønnen øke med 5%. På samme måte vil en økning på ett standardavvik i resultatet på lesetesten øke lønnen med 5,8%. Dette er i tråd med tidlige litteratur som ser på humankapital, og spesielt lokasjonsspesifikk humankapital, som en viktig forklaring på lønnsforskjeller mellom innvandrere og majoritetsbefolkningen. Høyere utdanning og leseferdigheter antas derfor å redusere lønnsgapet. Videre brukes en F-test, som brukes når flere variabler i en modell testes for signifikans, for å se om de aktuelle variablene er statistisk signifikante for lønnsrelasjonen. Får følgende hypoteser:

$$H_0: \beta_3 = \beta_4 = 0$$

$$H_A: \text{minst én av } \beta_3, \beta_4 \neq 0$$

```
. test yrsqual = litscore1 = 0
( 1) yrsqual - litscore1 = 0
( 2) yrsqual = 0

F( 2, 7460) = 639.12
Prob > F = 0.0000
```

FIGUR 2

Figur 2 viser F-testen slik den er gjennomført i STATA. P-verdien for denne testen er lik 0.0000. Det betyr at nullhypotesen kan forkastes, og derfor bekreftes det at utdanning og leseferdigheter er statistisk signifikant for lønnen.

Videre analyseres variablene *Tsjekkia* og *Frankrike*, og fra estimeringen observeres det at lønnsnivået i begge land er lavere enn i Irland, med henholdsvis 84% og 18%. Samtidig viser estimeringen at det faktisk er et større lønnsgap mellom innvandrere og majoritetsbefolkningen i Frankrike enn i Tsjekkia. Lønnsgapet er estimert til å være 22% større i Frankrike enn i Irland, mot 13.5% større i Tsjekkia enn i Irland. Dette tilsier at størrelsen på innvandrerbefolkningen ikke nødvendigvis påvirker lønnsgapet i altfor stor grad. I tillegg er det nødvendig å teste om variablene *Fraxmigrant* og *Tjsxmigrant*, i noe grad er statistisk signifikante variabler i en lønnslikning. Tester da to nullhypoteser: først om parameterne er statistisk signifikante, og deretter om lønnsgapet eventuelt er likt i de to landene sammenliknet med Irland. Velger et 5% signifikansnivå, og sjekker om lønnsgapet er likt i de tre landene:

$$(1) H_0: \delta_1 = \delta_2 = 0 \quad H_A: \text{minst én } \delta_1, \delta_2 \neq 0$$

$$(2) H_0: \delta_1 - \delta_2 = 0 \quad H_A: \delta_1 \neq \delta_2$$

Bruker STATA for å gjennomføre testene og får følgende resultater:

$$(1) \text{ Prob} > F = 0.0000 \quad (\text{Fraxmigrant} = \text{Tjsxmigrant} = 0)$$

$$(2) \text{ Prob} > F = 0.1403 \quad (\text{Fraxmigrant} - \text{Tjsxmigrant} = 0)$$

Resultatene av testene viser at nullhypotesen i test (1) kan forkastes, men ikke i test (2). Det kan tolkes som at lønnsgapet i Frankrike og Tsjekkia er ulikt fra lønnsgapet i Irland, og det kan være et faktum at størrelsen på innvandrerbefolkningen påvirker lønnsnivået. Samtidig

viser testen at lønnsgapet i de to landene er like stort sammenliknet med Irland, noe som tilsier at andre faktorer enn kun størrelsen på innvandrerbefolkningen kan øke eller redusere lønnsgapet.

Til slutt i testes effekten av foreldrenes inntekt på individets lønn. Estimeringen gir et overraskende resultat i forholdet mellom lønnen til individer som har foreldre med videregående utdanning og individer som har foreldre med kun grunnskoleutdanning. Ifølge modellen vil et individ med foreldre som har middels utdanningsnivå ha 3% lavere lønnsnivå enn individer med foreldre som har lavt utdanningsnivå. Ikke overraskende er relasjonen med høyt utdanningsnivå hos foreldrene motsatt. Et individ med høyt utdannede foreldre har 2% høyere lønn sammenliknet med et individ med lavt utdannede foreldre. Samtidig avhenger lønnen negativt av interaksjonsleddet $ParedxIm$. Dette tyder på at avkastningen av foreldrenes utdanningsnivå er lavere for innvandrere enn for majoritetsbefolkningen. Videre brukes en F-test igjen for å undersøke om foreldrenes utdanning er relevant for lønnsrelasjonen. Velger et 5% signifikansnivå, og får følgende hypotese:

$$H_0: \gamma_1 = \gamma_2 = \gamma_3 = 0 \qquad H_A: \text{minst én } \gamma_1, \gamma_2, \gamma_3 \neq 0$$

P-verdien blir i denne testen 0.0003, og H_0 kan derfor forkastes. Det betyr at foreldrenes utdanningsnivå vil ha noe å si for lønnen til et individ. Dersom hver av variablene hadde blitt testet enkeltvis, er det kun γ_2 som hadde hatt en høy nok p-verdi til å forkaste $H_0: \gamma_i = 0$. Dette kan tolkes som at utdanningsnivået til foreldrene er en relevant faktor for lønnsnivået til individet, men at modellen ikke bekrefter at høyt utdannede foreldre påvirker individets lønn positivt. Modellen bekrefter heller ikke at avkastningen av foreldrenes utdanning er ulik for innvandrere sammenliknet med majoritetsbefolkningen. Likevel åpner resultatet av denne hypotesetesten for at disse variablene kan undersøkes videre i andre analyser.

4.2. Avvik fra forventede resultater

Basert på tidligere studier, ble det antatt at et lønnsgap mellom innvandrere og majoritetsbefolkningen eksisterer. Dette bekreftes gjennom estimeringen og hypotesetesting av lønnsrelasjonen. Før analysen var de forventede resultatene at lønnsgapet ville være størst i Tsjekkia, at innvandrere fikk mindre uttelling for foreldrenes utdanningsnivå og at leseferdigheter ville ha mye å si for lønnsgapet. Fra de empiriske resultatene er det derimot noen avvik fra disse antagelsene.

Det første oppsiktsvekkende resultatet er at dersom kontrollvariablene for land og foreldrenes utdanning fjernes, estimeres det et positivt lønnsgap mellom innvandrere og majoritetsbefolkningen. I følge disse modellene er det derfor innvandrerne som har høyest lønn. Grunnen til at modellen gir et positivt lønnsgap, er mest sannsynlig et resultat av skjev fordeling av antall innvandrere fra hvert av studielandene. Det er flest innvandrere fra Irland i datasettet, som er et land med et generelt høyere lønnsnivå enn både Tsjekkia og Frankrike. Det betyr at innvandrere i Irland kan ha høyere lønn enn majoritetsbefolkningen i både Tsjekkia og Frankrike. Det gjør at lønnsgapet i modellen som ikke tar hensyn til land blir positivt. Det positive lønnsgapet er derfor ikke nødvendigvis representativt for relasjonen mellom lønn og å ha innvandrerbakgrunn, noe som bekreftes av modellens lave forklaringskraft.

I tillegg viser esimeringen av modellen at Frankrike har større lønnsgap mellom innvandrere og majoritetsbefolkningen enn Tsjekkia. Samtidig tilsier F-testen at lønnsgapet i de to landene vil skille seg like mye fra det irske lønnsgapet mellom innvandrere og majoritetsbefolkningen. Dette tyder på at størrelsen på innvandrerbefolkningen ikke nødvendigvis har så mye å si på lønnsgapet som først antatt. Samtidig kan det hende at størrelsen på innvandrerbefolkningen kun har en positiv effekt på lønnsgapet når den er over en gitt prosentandel, men at en mindre innvandrerbefolkning ikke direkte har en negativ effekt på lønnsgapet. En annen faktor som kan ha påvirket dette resultatet, men som det er mangelfull informasjon på i datasettet, er hvor innvandrerbefolkningen kommer fra. I det teoretiske rammeverket ble det presentert at kulturgapet mellom innvandrer og mottakerlandet påvirker blant annet statistisk diskriminering og lokasjonsspesifikk humankapital. Dersom innvandrerne i Frankrike i større grad er fra fremmedkulturer enn innvandrerne i Tsjekkia, kan dette være noe som påvirker lønnsgapet mer enn den totale prosentandelen innvandrere i befolkningen. Samtidig viste datasettet at innvandrere hadde lavere leseferdigheter i Frankrike enn i Tsjekkia. Dette kan tyde på at flere innvandrere kommer fra land som skiller seg i større grad fra fransk språk og kultur, enn innvandrere i Tsjekkia skiller seg fra tsjekkisk språk og kultur. Samtidig kan det lave nivået på utdanning i Frankrike slå hardere ut på lønnsnivået til innvandrere i landet, som skaper grunnlaget for et større lønnsgap enn antatt.

Det siste overraskende resultatet, er at foreldre med middels utdanningsnivå har negativ påvirkning på individets lønn sammenliknet med foreldre med lavt utdanningsnivå. Det er vanskelig å finne noen intuitiv grunn for denne sammenhengen, men den kan være en effekt av en situasjon med et «utelatt variabel-problem». I noen tilfeller kan utelatelse av relevante

variabler i en modell gi skjeve estimater for andre inkluderte variabler i modellen. I denne modellen kan for eksempel utelatelsen av foreldrenes arbeidserfaring og individets arbeidserfaring være en mulig variabel som påvirker effekten av foreldrenes utdanning på lønnen. En forelder som startet å jobbe rett etter grunnskolen kan ha en positiv påvirkning på individets arbeidsmoral. Samtidig kan individer som følger i foreldrens fotspor og ikke tar høyere utdanning, starte i arbeid tidligere og derfor få høyere lønn gjennom lengre arbeidserfaring. Til slutt kan noe av den positive effekten av foreldrenes utdanning allerede være fanget opp av variabelen for individets egen utdanning, *yrsqual*, som kan gi et mindre representativt bilde av den faktiske effekten av foreldrenes utdanningsnivå på individets lønn.

5. Oppsummering og konklusjon

Denne oppgaven har som mål å studere lønnsgapet mellom innvandrere og majoritetsbefolkningen i ulike land, og å se på om størrelsen på innvandrerbefolkningen i et land har påvirkning på lønnsgapet. Empirien tilsier i dette tilfellet at det eksisterer et lønnsgap mellom innvandrere og innfødte, men at andelen innvandrere i befolkningen i seg selv ikke har en direkte påvirkning på lønnen. Som nevnt over vil det være mer hensiktsmessig å studere hvor innvandrerne i mottakerlandet kommer fra, da dette vil ha mer påvirkning på for eksempel statistisk diskriminering. Fremmedfrykt og ren diskriminering kan også ha en stor påvirkning på et eventuelt lønnsgap. Dette finnes det flere eksempler på, og kan bekreftes blant annet av president Trump i USA. Han har uttalt at det er mer ønskelig med norske innvandrere i USA, enn fra andre land (Bergman, 2018). Dette er et eksempel på hvordan nasjonaliteten til innvandreren kan ha mer å si på diskriminering i et land, enn hvor stor innvandrerbefolkning landet faktisk har. Dersom dette resonnementet skal trekkes opp mot diskriminering i lønn, vil det i en arbeidsplass kunne observeres lønnsforskjeller både mellom innfødte arbeidere og innvandrere, og mellom innvandrere fra ulike nasjonaliteter.

Videre ble effekten av foreldrenes utdanning på individets lønn analysert. Dette er en variabel som både direkte og indirekte kan påvirke individets lønn. Direkte ved at foreldre med høy utdanning og gode jobber kan gi barnet sitt en lettere vei inn i en godt betalt jobb, og indirekte ved at utdanningen til foreldre i stor grad påvirker om individet selv tar høyere utdanning eller ikke. Som en del av lønnsrelasjonen ble det estimerte lønnsgapet mellom innvandrere og majoritetsbefolkningen redusert når foreldrenes utdanning ble inkludert som variabel. Likevel

var de estimerte sammenhengene mellom lønn og foreldrenes utdanningsnivå noe uforventede. Grunnen til dette kan være en effekt av et «utelatt variabel-problem», og det kan argumenteres for at foreldrenes utdanning i større grad bør undersøkes gjennom andre variabler, som individets utdanning eller foreldrenes arbeidserfaring. Likevel bekrefter F-testen av modellen at foreldrenes utdanning har en påvirkningskraft for lønnen til et individ, og det vil ved senere undersøkelser være interessant å diskutere dette som et tilleggsspørsmål med andre, utvidede modeller. Samtidig kunne ikke denne modellen bekrefte at foreldrenes utdanning har ulik avkastning på lønnen til innvandrere og majoritetsbefolkningen. Dette kan være et tegn på at foreldrenes påvirkning på barnas lønn i større grad påvirker gjennom hvordan foreldrene tilegner barna ferdigheter. Foreldre i et mottakerland som ikke lærer seg språket og kun snakker morsmålet hjemme, kan gjøre at barna selv får språkvansker senere i livet. Dette vil da kunne påvirke lønnen til individene i større grad enn foreldrenes utdanning i seg selv.

Det kan argumenteres for at en utvidet modell med flere variabler som fanger opp humankapital vil være bedre til å analysere lønnsgapet enn modellen presentert i denne oppgaven. Likevel har modellen med to slike variabler som er essensielle for lønnsnivået. Leseferdigheter er en indikator på individets språkkunnskaper, og har i andre studier vist seg å være en viktig faktor på lønnsforskjeller mellom innvandrere og majoritetsbefolkningen. I denne modellen blir det likevel estimert et relativt høyt lønnsgap (17%), selv når det kontrolleres for viktige variabler på humankapital. Derfor kan det argumenteres for at diskriminering, både statistisk og personbasert, kan være en forklaringsfaktor for modellen brukt i denne analysen.

Basert på tidligere litteratur kan det ikke med sikkerhet sies at en utvidet modell ville eliminert lønnsgapet mellom majoritetsbefolkningen og innvandrere helt. I denne oppgaven har likevel noen viktige sammenhenger mellom lønn og utdanning blitt fanget opp, og det har blitt analysert hvorvidt prosentandelen innvandrere i befolkningen vil påvirke lønnsgapet. Denne sistnevnte hypotesen ble ikke bekreftet, men det ble til en viss grad vist at foreldrenes utdanningsnivå kan påvirke lønnsnivået til innvandrerne. Derfor blir konklusjonen, som i så mange andre studier før denne: Lønnsgapet mellom innvandrerne og majoritetsbefolkningen eksisterer og økt lokasjonsspesifikk humankapital har en positiv effekt innvandrerens lønn, men noen deler av lønnsgapet er fortsatt i stor grad vanskelig å begrunne uten at diskriminering av noe slag er en del av bildet.

6. Litteraturliste

Barrett, A. & McCarthy, Y. (2007). Immigrants in a Booming Economy: Analysing Their Earnings and Welfare Dependence. *LABOUR*, 21: 789-808. doi:[10.1111/j.1467-9914.2007.00389.x](https://doi.org/10.1111/j.1467-9914.2007.00389.x)

Bergman, J. (2018). Donald Trump prefers immigrants from Norway, but more Americans move to Scandinavia than vice versa. *Independent*. Hentet fra: <https://www.independent.co.uk/news/world/americas/donald-trump-shithole-countries-norway-immigration-american-emigrants-scandinavia-a8159441.html>

Cain, G.G. (1986). The Economic Analysis of Labor Market Discrimination. *Handbook of Labor Economics*, Ch.13, 693-717. Hentet fra: <http://unionstats.gsu.edu/9220/Cain%281986%29%20The%20Economic%20Analysis%20of%20Labor%20Market%20Discrimination.pdf>

Epland, J. & Kirkeberg M.I. (2018). Store inntektsforskjeller blant norskfødte med innvandrereforeldre. *SSB Analyse 18/10: Inntekt*. Hentet fra: <https://www.ssb.no/inntekt-og-forbruk/artikler-og-publikasjoner/store-inntektsforskjeller-blant-norskfodte-med-innvandrereforeldre>

Grand, C. I. & Szulkin, R. (2002), Permanent Disadvantage or Gradual Integration: Explaining the Immigrant–Native Earnings Gap in Sweden. *LABOUR*, 16: 37-64. doi:[10.1111/1467-9914.00186](https://doi.org/10.1111/1467-9914.00186)

Moffat, M. (2019). The Economics of Discrimination, An examination of the economic theory of statistical discrimination. Hentet fra: <https://www.thoughtco.com/the-economics-of-discrimination-1147202>

Statistisk Sentralbyrå (2017). Foreldrenes utdanningsnivå har mye å si. *Gjennomføring ved universiteter og høyskoler (Arkiv)*. Hentet fra: <https://www.ssb.no/utdanning/artikler-og-publikasjoner/foreldrenes-utdanningsniva-har-mye-a-si>

Tomaskovic-Devey, D., Hällsten & M., Arent-Holt, D. (2015), Where do Immigrants Fare Worse? Modeling Workplace Wage Gap Variation with Longitudinal Employer-Employee Data. *American Journal of Sociology* 120, no-4: 1095 – 1143. doi: <https://doi.org/10.1086/679191>

Tverdostup, M. & Paas, T. (2018). Skills heterogeneity and immigrant-native wage gap in the European countries. *Journal of Economic Literature* Vol. VII(2), pp. 119-142. , hentet fra: <http://www.iises.net/international-journal-of-economic-sciences/publication-detail-6942>