

Bacheloroppgave i Samfunnsøkonomi

Våren 2019

Lønnsforskjeller mellom innvandrere og
majoritetsbefolkning i to land med ulik
økonomisk vekst

Antall ord: 10 473

13.mai 2019

Institutt for økonomi/NTNU

Innholdsfortegnelse

1. Innledning.....	3
1.1 Problemstillinger.....	3
2. Teoretisk grunnlag.....	4
2.1 Humankapital – generell og landsspesifikk.....	4
2.2 Utdanning og avkastning.....	5
2.3 Utdanning og lønnsavkastning for innvandrere.....	7
2.4 Innvandrere i Italia og arbeidssituasjonen deres.....	8
2.5 Land med ulik økonomisk vekst.....	8
3. Metode.....	9
3.1 Den kvantitative metoden.....	10
3.2 Datasettet og PIAAC-undersøkelsen.....	10
3.3 Variabelbeskrivelse.....	11
3.3.1 Avhengig variabel.....	11
3.3.2 Uavhengig variabel.....	11
3.3.3 Kontrollvariabler.....	11
3.4 Deskriptiv statistikk.....	12
3.4.1 Norge.....	12
3.4.2 Italia.....	13
4. Økonometrisk modell.....	14
4.1 Empirisk strategi.....	14
4.2 Valg av funksjonsform.....	15
4.3 Hypotesetesting.....	15
4.4 Regresjonsmodell for Norge.....	17
4.4.1 Modell 1.....	18
4.4.2 Modell 2.....	19
4.4.3 Modell 3.....	19
4.4.4 Modell 4.....	20
4.4.5 Oppsummering.....	21
4.5 Regresjonsmodell for Italia.....	21
4.5.1 Modell 1.....	22
4.5.2 Modell 2.....	22
4.5.3 Modell 3.....	23
4.5.4 Modell 4.....	24
4.5.5 Oppsummering.....	24
4.6 Regresjonsmodell for Norge og Italia sammen.....	24
4.6.1 Modell 1.....	25
4.6.2 Modell 2.....	26
4.6.3 Modell 3.....	26
4.6.4 Modell 4.....	27
5. Diskusjon.....	28
5.1 Hvordan vil det å være innvandrere påvirke lønnsnivået?	28
5.2 Er det bedre å være innvandrere i det ene eller andre landet?	29
5.3 Styrker og svakheter ved modellene og oppgaven.....	30
6. Konklusjon.....	32
7. Litteraturliste.....	33

1. Innledning

Arbeidsmarkedet er et sentralt punkt for samfunnsdeltakelse, og påvirker i stor grad hvor knyttet man er til det sosiale, kulturelle og politiske i et land. I denne oppgaven ønsker jeg å studere hvordan det å være innvandrere spiller inn på arbeidssituasjon og lønnsnivå, samt hvilke effekter utdanning, leseferdigheter, tallforståelse og helsenivå har på lønn for innvandrere kontra majoritetsbefolkningen. I tillegg til å studere denne sammenhengen, ønsker jeg også å si noe om det er større forskjeller knyttet til lønn og arbeidssituasjon for innvandrere i land med ulik økonomisk vekst. Landene jeg har valgt ut for å representere situasjoner med ulik økonomisk vekst, er Norge og Italia.

Kan det tenkes at det er bedre å innvandre til et land med stor økonomisk vekst, eller er dette uvesentlig for lønnsnivået? Er avkastningen på lønn for innvandrere avhengig av om de har tatt utdanningen sin i hjemlandet eller landet de har immigrert til? Dette, og flere spørsmål, vil jeg forsøke å besvare gjennom denne oppgaven. Etersom lønnsnivå og arbeidssituasjon er svært knyttet til hverandre, vil jeg bruke disse begrepene litt om hverandre videre i oppgaven.

1.1 Problemstillinger

Her beskriver jeg kort de ulike problemstillingene jeg har valgt for oppgaven.

- 1) Observeres det et lønnsgap mellom innvandrere og majoritetsbefolkningen i Norge og Italia, og hvor stort er dette?
- 2) I hvor stor grad kan det eventuelle lønnsgapet forklares av forskjeller i utdanning, leseferdighet og tallforståelse?
- 3) Kan det eventuelle lønnsgapet forklares av forskjeller i helsestatus?
- 4) Er lønnsgapet mellom innvandrere og majoritetsbefolkningen større i et av de to landene? Kan dette forklares av forskjeller i den økonomiske situasjonen?

Videre i oppgaven vil det først følge en teoridel, og deretter et metodekapittel. Så vil jeg presentere den økonometriske modellen, tolke modellene, og deretter diskutere funnene i sammenheng med det teoretiske grunnlaget, før jeg vil avslutte med en konklusjon.

2. Teoretisk bakteppe

I denne delen av oppgaven vil jeg gjøre rede for begreper og teori som er relevant for temaet som jeg ønsker å studere. Teoridelen vil først og fremst dreie seg rundt teorier om hvordan det å være innvandrere påvirker arbeidssituasjonen, og hvordan avkastningen av utdanning hos innvandrere er sammenliknet med majoritetsbefolkningen. Her vil jeg komme inn på begreper som humankapital og inntektsutvikling. I slutten av teorikapittelet vil jeg kort gjennomgå teori for økonomisk vekst, og forsøke å si om lønn, arbeidssituasjon og utdanning preges av om innvandrere har flyttet til et land med høy eller lav økonomisk vekst.

Det teoretiske materialet jeg bruker i denne oppgaven vil i stor grad basere seg på teorier om risiko for arbeidsledighet så vel som teori om lønnsvariasjon, fordi disse to begrepene i stor grad sammenfaller med hverandre. Om en teori tilsier at en gruppe innvandrere har større risiko for å være arbeidsledig, så vil dette også bety at denne gruppen vil oppleve lavere inntektsnivå.

2.1 Humankapital – generell og landspesifikk

Arbeidsmarkedet er, som nevnt innledningsvis, et sentralt punkt for samfunnsdeltakelse, og er i tillegg å gi inntekt til befolkningen, også viktig for involvering både sosialt og politisk. Derfor er det svært viktig med et arbeidsmarked som favner om alle samfunnsgrupper, og som ikke diskriminerer. Likevel viser flere studier fra skandinaviske land at det finnes betydelige forskjeller i inntekt mellom majoritetsbefolkning og innvandreregrupper (Wiborg, 2006:277). Dette gjelder også til tross for at innvandrerne kommer til det nye landet med god utdanning, eller fullfører en høy utdanning i landet de immigrerer til (Ibid.).

Det finnes flere teoretiske forklaringer på ulikhetene i inntekt blant folkegrupper, og blant disse har vi teorien om humankapital. Hvor godt man integreres i arbeidsmarkedet, og hvor mye inntekt man oppnår, vil avhenge av flere forhold. Det vil være relevant å se på forskjeller i inntekt som forklart av både arbeidstakers ressurser, samt arbeidsgivers preferanser, og jeg vil fokusere mest på førstnevnte (Wiborg, 2006:278). En arbeidstakers ressurser består av blant annet sosiale nettverk, motivasjon og utdanning, og vil sammen være med på å påvirke dens muligheter i arbeidslivet. Ofte bruker man begrepet humankapital for å vurdere en persons produktivitet, som direkte spiller inn på muligheten for å få seg jobb, samt inntektsnivået. Humankapital er viktig av den enkle grunn at arbeidskraft som kjøpes og selges ikke bare avhenger av det kvantitative behovet, men vel så mye av kvalitet (Ibid:279). Dette betyr at jo høyere nivå på humankapital en person har, dess mer produktiv er han/hun, og jo mer vil denne

personen tjene. Altså vil likt nivå på humankapital gi lik inntekt, uavhengig av opprinnelse, kjønn, alder og sosioøkonomisk status, ideelt sett.

Et steg videre fra begrepet humankapital, finner vi det som Chiswick (1978) kaller landspesifikk humankapital (Wiborg, 2006:277). Chiswick argumenterer for at også språklige ferdigheter og samhandlingsegenskaper, i tillegg til utdanning, er relevante i vurderingen av en person sin produktivitet, og derfor må dette også legges til. Teorien om landspesifikk humankapital baserer seg altså på at innvandrere vil ha en sterkere inntektsutvikling enn majoritetsbefolkningen, ettersom arbeidserfaring og tilpasning i det nye landet vil bedre deres vilkår i arbeidslivet (Wiborg, 2006:277). Ettersom innvandrerne blir bedre tilpasset i landet de har immigrert til, og får tilegnet seg språk og annen kulturspesifikk kompetanse, desto likere blir befolkningsgruppene, og ulikhetene i inntekt vil avta over tid (Ibid.). Som en følge av dette vil tiden man har bodd i det landet man ønsker å arbeide, påvirke graden av produktivitet, men dette gjelder mest av alt når botiden i landet er kort. Gitt at en innvandrer har bodd i det landet det immigreres til i lang tid (10-15år), så vil ikke lenger forskjellene i landspesifikk humankapital eksistere i særlig grad, og dermed vil denne personen ha like forutsetninger i arbeidsmarkedet som en person fra dette landet, når man kontrollerer for humankapital (Wiborg, 2006:279).

Hittil er det altså realistisk å anta at innvandrere og majoritetsbefolkning i et land har relativt like muligheter for lik lønn og arbeid, gitt at nivået på utdanning og andre ferdigheter er lik, samt at innvandrerene har bodd i landet lenge nok til å mestre språket relativt bra. Hvis vi imidlertid begynner å studere sosiale nettverk og dets påvirkning på mulighetene i arbeidsmarkedet, finnes det teorier som tilsier at det er vanskeligere for innvandrere å konkurrere på et likt nivå (Wiborg, 2006:280). Sosiale nettverk brukes hyppig som en snarvei inn i en bestemt jobb, fordi det å ha kontakter gjør at man lettere kan vise egne ferdigheter for en eventuell arbeidsgiver (Ibid.).

2.2 Utdanning og avkastning

Så vel som å si noe om inntektsforskjellene mellom innvandrere og majoritetsbefolkning i et land, ønsker jeg også å studere om eventuelle forskjeller kan forklares av utdanning og ferdigheter. Vil likt nivå på utdanning gi lik lønn og mulighet i arbeidslivet, eller finnes det avvik basert på etnisitet og hvor man har tatt utdanningen sin? For å besvare dette vil jeg først

gjøre rede for sammenhengen mellom utdanning og lønn, samt si noe om avkastning på utdanning.

Det er godt dokumentert og forsket på at det er en positiv sammenheng mellom utdanningsnivå og inntekt, og at personer som har valgt høyere utdanning får igjen for dette (Raaum, 1999:3). Hvis man ser på utdanning fra et rent økonomisk perspektiv, så vil utdanning regnes som en investering. Til tross for at man går glipp av inntekt mens man studerer, samt minsker den potensielle arbeidstiden, så vil utdanning bidra til at man har flere valgmuligheter og er bedre sikret fra arbeidsledighet (Ibid.). Til tross for at man kan stadfeste at det er en positiv korrelasjon mellom utdanning og inntekt, så er det vanskelig å måle den faktiske avkastningen man får av utdanningen man tar (Raaum, 1999:5). Likevel finnes det noen argumenter som viser sammenhengen mellom utdanning og dens avkastning ganske godt. Først og fremst ser man at utdanning gir mennesker produktiv kunnskap (humankapital), som jeg var inne på i forrige delkapittel, som brukes til å produsere varer eller tjenester. Etersom man i større eller mindre grad blir betalt basert på produktiviteten sin, så vil utdanning gi økt inntekt (Ibid:6). Det andre argumentet som sier noe om sammenhengen mellom utdanning og inntekt, baserer seg på antakelsen om at utdanningskarrierer ikke er tilfeldig fordelt, men avhenger av en rekke valg påvirket av den oppveksten man har hatt. Dette gjør at folk som har stort utbytte av skolegang vil velge å ta en høy utdanning, samtidig som at de som velger å ta høy utdanning ofte har et stort lønnspotensiale (Ibid.). Det finnes imidlertid store individuelle forskjeller blant de som tar høyere utdanning, så det er ikke alltid like lett å tallfeste det man tror er en sterk sammenheng mellom utdanning og lønn (Ibid). Videre er det også utfordrende å si noe om hvilken vei sammenhengen mellom produktivitet og utdanning, og dermed inntekt, går. Det er nemlig ikke alltid slik at utdanning gjør folk produktive, men heller at produktivitet bestemmer hvor lang utdanning en person velger å ta (Raaum, 1999:7).

Mikroøkonomisk teori om at individer er nyttemaksimerende, tilsier at hvert individ vil velge det utdanningsnivået som maksimerer inntekten sin, noe som innebærer at de vil velge å utdanne seg til det punktet der lønnsgevinsten av utdanningen er lik kostnaden knyttet til å ikke arbeide (Raaum, 1999:7). Det finnes imidlertid problemer med dette rent praktisk, ettersom individer gjerne ikke regner på den faktiske nytten/kostnaden av utdanning, og at de nødvendigvis ikke vet hvor nyttig utdanningen de tar er, før den er ferdig. I tillegg til slike informasjonsbegrensninger, blir regnestykket også komplisert ettersom avkastningen på utdanning ikke nødvendigvis er lik for hvert år man utdanner seg. Spesielt ser man at

inntektseffekten av høyere utdanning på 1-4 år ikke er spesielt stor, men at mennesker med dette utdanningsnivået likevel i gjennomsnitt tjener bedre enn de som kun har fullført videregående skole (Ibid:4). Har man imidlertid fullført en mastergrad eller liknende, så viser studier at inntektseffekten er relativt høy (Ibid.).

2.3 Utdanning og lønnsavkastning for innvandrere

Nå som jeg har gjort rede for sammenhengen mellom utdanning og lønn helt generelt, ønsker jeg å si noe om innvandrere i Norge sin avkastning av utdanning. Generelt sett ser man en trend om at de som har gjennomført høyere utdanning i Norge, stort sett stiller med like gode muligheter i arbeidslivet og har samme lønnsnivå som majoritetsbefolkningen (Raaum, 1999:40). Dette virker som en rimelig antakelse ettersom disse innvandrerne gjerne scorer relativt godt på landspesifikk humankapital, som omtalt i delkapittel 2.1. Altså virker det som at høyere utdanning gjennomført i Norge er en god investering for innvandrere, som gjør at de stiller på lik linje med resten av befolkningen. Ser man videre på situasjoner der innvandrerne har tatt utdanningen sin i hjemlandet før de flyttet til Norge, så finnes det noen lønnsforskjeller. Disse kan forklares av flere årsaker, som for eksempel at utdanning i et land ikke samsvarer med nødvendige egenskaper og ferdigheter i en jobb i et annet land. Likevel ser det ut til at utdannelser på mastergrads-/profesjonsnivå gir lik lønn uansett hvor man har studert.

En tydelig tendens i studier om utdanning og inntekt, viser at innvandrere fra ikke-vestlige områder scorer jevnt over dårligere på lønnsnivå enn andre innvandrergrupper. Dette skyldes blant annet at det gjennomsnittlige utdanningsnivået til denne gruppen er lavere, samt at disse søker jobb i yrker med lavere lønnsnivå. (Raaum, 1999:41). I tillegg observerer man at innvandrere fra den tredje verden er overrepresentert i yrker som ansetter ufaglærte (Ibid.). En annen forklaring på at mennesker fra den ikke-vestlige verden opplever større lønnsforskjeller sammenliknet med andre innvandrere, kan skyldes at det er større kulturforskjeller for disse når de kommer til Norge. Dette gjør det vanskeligere å score høyt på landspesifikk humankapital, og dermed tar det lengre tid før lønnsforskjellene jevner seg ut.

Det finnes flerfoldige utfordringer knyttet til å si noe om hvorfor innvandrere gjennomsnittlig har en lavere inntekt enn majoritetsbefolkningen i et land, og dette skyldes kanskje mest av alt operasjonelle utfordringer. Mangel på informasjon og mulighet til å beregne ”nyttens” av en utdanning i et annet land, gjør det vanskelig å tallfeste sammenhenger, og dette vil jeg komme mer tilbake til i metodekapittelet.

2.4 Innvandrere i Italia og arbeidssituasjonen deres

Italia – landet jeg har valgt ut for å representere en økonomisk situasjon preget av lav økonomisk vekst, har de siste tiårene opplevd høy grad av immigrasjon. Immigrasjonen kan beskrives som et dilemma, hvor avveiningen står mellom å oppfylle ønsker på markedet versus på det politiske og sosiale plan. (Ambrosini, 2013). Den kraftige innvandringen til Italia skyldes både lett tilgjengelige grenser over Middelhavet, samt stor arbeidskraftsetterspørsel. Flere byer blir mer og mer multi-etniske, en realitet som ledelsen i landet i mange år har nektet for. Det virker svært utfordrende å stabilisere landet som preges av store ulikheter og heterogenitet (Ibid.). På en side har man sett en stor økning i utdanningsnivået blant unge italienerne de siste årene, noe som legger til rette for at disse kan jobbe i høyere stillinger enn tidligere (Ambrosini, 2013). Dette åpner opp for stor etterspørsel etter arbeidskraft i yrker og sektorer hvor det fortsatt trengs manuell arbeidskraft. Disse sektorene er ikke lenger attraktive for flertallet av majoritetsbefolkningen, og sånn sett kan man se på den store immigrasjonen til Italia som nødvendig for å dekke arbeidskraften som trengs i lavinntektsyrker. Dette skaper imidlertid et stort skille mellom majoritetsbefolkningen og innvandrerne, der innvandrerne, til tross for å være akseptert i arbeidsmarkedet, kun er godkjent på det sosiale planet så lenge de holder seg nederst på rangstigen hva gjelder arbeid. Arbeidsledigheten blant innvandrerne kan altså antas å være relativt lav, noe som er bra, men andelen som jobber i yrker som krever liten eller ingen kompetanse, er veldig høy (Ibid.). Uten å trekke for tidlige konklusjoner, kan man likevel tenke seg at det finnes inntektsforskjeller mellom majoritetsbefolkningen og innvandrerne i Italia, som følge av at innvandrerne jobber i lavinntekts-bransjer.

2.5 Valg av land med ulik økonomisk vekst

For å kunne studere om det er lønnsforskjeller for innvandrere som immigrerer til land med ulik økonomisk vekst, så har jeg valgt meg ut to land i Europa som er ulike i så måte. Jeg har benyttet et datasett for økonomisk vekst distribuert av University of Groningen (2019), og brukt dette til å plukke ut to land som også finnes i det datasettet som jeg bruker i den økonometriske modellen. Valget falt på Norge og Italia – to europeiske land, som ingen kan sies å være helt på bunn av økonomisk utvikling hvis vi ser på verdensbasis, men som er såpass ulike at det vil være interessant å kunne avdekke forskjeller for innvandrerne som bor i de to landene. Nedenfor har jeg lagd en tabell som viser utviklingen i BNP per capita fra 1990 til 2011:

Land	BNP p.c. 1990	BNP p.c. 2011
Norge	29164.39	80623.66
Italia	24710.5	35151.28

Her ser vi tydelig at landene skiller seg veldig fra hverandre når det kommer til vekst i økonomien, og til tross for at de startet på et relativt likt nivå i 1990, så skiller det nå over dobbelt så mye mellom Italia og Norge.

Økonomisk vekstteori dreier seg om hvilke faktorer som trengs for å gi utvikling og vekst, samtidig som at teori om økonomisk vekst også sier noe om effektene av at et land sin økonomi blir større og bedre. Begreper som økt velstand, lavere dødsrate, lavere fødselsrate, høyere utdanningsnivå, bedre politiske institusjoner og et bedre arbeidsmarked knyttes gjerne sammen med økonomisk vekst (Smith & Todaro, 2015:41). For noen av disse faktorene er det vanskelig å si hvilken vei sammenhengen går. Er det slik at økonomisk vekst gir bedre politiske institusjoner, eller er det motsatt? Og det samme kan absolutt sies å gjelde for utdanning og arbeidsmarkedet også. For det virker logisk å tenke at med økt utdanningsnivå i befolkningen, samt et åpent og velfungerende arbeidsmarked, så legger man til rette for høyere produktivitet, som vil resultere i økonomisk vekst. Samtidig argumenteres det for at med et relativt høyt nivå på BNP per innbygger, så har man gode muligheter for å lykkes på andre områder, som med utdanning og arbeidsmarked (Ibid: 59, 69). Det er denne linken jeg vil bruke videre i oppgaven. Det kan nemlig være rimelig å anta at land med høyere økonomisk vekst, slik Norge har sammenliknet med Italia, har et bedre utdannings- og arbeidsmarkedssystem, som gjør at folk her har et generelt høyere utdannings- og lønnsnivå. Dette vil kunne ha positiv innvirkning på innvandrere, og muligens gjøre det lettere å være innvandrer i Norge enn i Italia, hvis vi ser på det rent økonomisk.

3. Metode

Denne delen av oppgaven vil innledes ved å gjøre rede for den kvantitative metoden, og deretter vil jeg presentere datasettet og variablene som jeg skal benytte meg av i analysedelen. Jeg vil også gjøre en liten diskusjon rundt datasettets styrker og svakheter, samt utfordringene som kan oppstå i forbindelse med en slik forskningsmetode. Jeg har valgt å bruke kun ett datasett for hele oppgaven, og vil komme nærmere inn på valget lenger ned i dette kapitlet. Avslutningsvis i metoddelen presenteres deskriptiv statistikk for variablene.

3.1 Den kvantitative metoden

Ettersom jeg i denne oppgaven vil studere hvordan det å være innvandrere versus majoritetsbefolkning i et land påvirker ens muligheter i arbeidsmarkedet, og påfølgende inntektsnivå, vil dette best kunne studeres ved hjelp av den kvantitative metoden. Den kvantitative metoden defineres og kjennetegnes av undersøkelser som tar for seg en stor mengde data og som forsøker å generalisere et fenomen (Ringdal, 2013:104). Lønn i seg selv er svært kvantifiserbart, men andre variabler som skal brukes i oppgaven kan by på større operasjonelle utfordringer. Ringdal argumenterer likevel for at sosiale fenomener ofte er så stabile at de kan måles og beskrives kvantitativt (Ibid.). Det finnes imidlertid utfordringer ved den kvantitative metoden, og blant disse finner man at det er vanskelig å sikre god og lik informasjonstilgang fra alle områdene og landene man ønsker å studere (Bamberger, 2000:12). I tillegg erfarer man også at noen sosiale og samfunnsrelaterte fenomener ikke kan operasjonaliseres og måles like lett som andre (Ibid.).

3.2 Datasettet og PIAAC-undersøkelsen

I denne oppgaven har jeg valgt å benytte et datasett fra PIAAC-undersøkelsen i regi av OECD, som ble gjennomført i 2011-2012 (Bjørkeng, 2013:7). PIAAC-undersøkelsen hadde som formål å danne et bilde på voksne menneskers ferdigheter på tre sentrale områder: leseferdighet, tallforståelse og problemløsning i IKT (Bjørkeng, 2013:7). Det finnes utfordringer med å få resultatene fra undersøkelsen til å være sammenliknbare mellom de ulike landene, men dette er gjort i så stor grad det er mulig. Datasettet inneholder observasjoner fra 24 land, men jeg har som nevnt valgt meg ut Italia og Norge for denne oppgaven. Valget av datasett baserer jeg på at det har nødvendige observasjoner for å kunne studere lønnsforskjeller mellom innvandrere og majoritetsbefolkningen, og at det også inneholder tilstrekkelige observasjoner om andre forhold som gjerne knyttes til å forstå årsakene til ulikheter i inntektsnivå.

Datasettet inneholder observasjoner fra folk mellom 16 og 65 år, men i oppgaven vil jeg kun inkludere observasjoner som gjelder for folk som er 25 år eller eldre. Dette skyldes at ettersom jeg forsøker å si noe om lønnsforskjeller, som i stor grad har sammenheng med utdanning, så er det mest hensiktsmessig å kun inkludere observasjoner fra folk som stort sett er ferdigutdannede. Man vil nødvendigvis få noen mangler av dette, siden det finnes endel folk under 25 som er ferdigstudert, men likevel vil dette valget gjøre analysen mest riktig.

3.3 Variabelbeskrivelse

Her vil jeg beskrive og gjøre rede for de variablene som inkluderes i den økonometriske modellen, slik at analysen videre blir oversiktlig.

3.3.1 Avhengig variabel

Den avhengige variabelen jeg har valgt er den som er kalt *earnhrppp* i datasettet, som jeg videre i oppgaven vil omtale som *timelønn*. Jeg har valgt denne kontinuerlige variabelen fordi den angir timelønnen til en person, justert for PPP. Altså vil den lønna som oppgis være justert for prisforskjeller mellom landene, slik at man får et så ”likt” sammenlikningsgrunnlag som mulig. Jeg har valgt å logtransformere variabelen slik at man kan se på prosentvis økning i timelønnen, noe som gjør koeffisientene i økonometrimodellene veldig lett å tolke. Mer om logtransformasjonen kommer senere.

3.3.2 Uavhengig variabel

Jeg har valgt *migrant_first* som den uavhengige variabelen i oppgaven, og dette er en dummyvariabel som følgelig har to kategorier, der 1 = innvandrere og 0 = majoritetsbefolkning. Denne variabelen er definert for de som er første-generasjonsinnvandrere, noe som betyr at dersom du eller minst en av foreldrene dine er født i utlandet, så regnes man som migrant. Videre i oppgaven vil jeg omtale denne variabelen som *innvandrere*.

3.3.3 Kontrollvariabler

Jeg har valgt ut en god del kontrollvariabler, som jeg tilsammen tror vil gi et tilstrekkelig bilde på forhold som spiller inn på lønnsnivået til en innvandrere kontra majoritetsbefolkning, med hensyn på problemstillingene jeg har valgt. Utvalget av disse kontrollvariablene baserer seg også på det jeg vurderte som mest interessant å studere, for det kan jo alltid inkluderes flere forhold.

Den første kontrollvariabelen jeg har med er *yrqual*, en utdanningsvariabel som viser det høyeste antall år en respondent har studert. Denne variabelen strekker seg fra et minimumsnivå på utdanning på 7 år til maksimum på 21 år. Variabelen vil omtales som *utdanning*. Videre har jeg tatt med to av de tre variablene som sier noe om resultatet på PIAAC-undersøkelsen, følgelig variablene for leseferdigheter og tallforståelse. Årsaken til at jeg har droppet variabelen om IKT-forståelse skyldes at datasettet for Italia ikke hadde med observasjoner på denne

variabelen. I datasettet har man to ulike varianter på ferdighetstestene, og i den deskriptive statistikken så vises statistikk for ”rådataene” til de to variablene, hvor leseferdighetene og tallforståelsen til respondentene beregnes på en skala fra 0 til 500 (Bjørkeng, 2013:14). I den økonometriske modellen derimot, har jeg valgt ut den standardiserte versjonen av variablene, henholdsvis *litscore1* og *numscore1*. At variablene er standardiserte betyr at man har justert de slik at gjennomsnittet er 0 og standardavviket er 1, og dette gjøres for å lette den analytiske tolkningen videre. Videre i oppgaven vil jeg omtale ferdighetsvariablene som henholdsvis *leseferdigheter* og *tallforståelse*. I tillegg til disse, har jeg også inkludert variabelen *i_q08*, en variabel som beskriver helsesituasjonen til respondenten. Denne variabelen hadde i utgangspunktet 5 kategorier for helsen til respondenten, men jeg har valgt å generere den om til en dummyvariabel, slik at man kun har to kategorier med henholdsvis ”god” eller ”dårlig” helse. Den nye variabelen har jeg kalt *helse*, der verdien 1 gir god helse, og 0 betyr dårlig helse. Den siste kontrollvariabelen jeg har med er en dummyvariabel som heter *Norway*. Denne variabelen vil kun benyttes når jeg ønsker å sammenlikne lønnsforskjellene i Norge og Italia, så den er kodet slik at Norge=1 og Italia=0.

3.4 Deskriptiv statistikk for Norge og Italia

Nedenfor illustreres deskriptiv statistikk for henholdsvis Norge og Italia, først kontinuerlige variabler og deretter dummyvariabler. Under tabellene vil jeg gjøre en rask gjennomgang av statistikken, for å gi et inntrykk av hvordan forholdene er i de to landene når det kommer til timelønn, utdanning, ferdigheter, etnisitet og helsenivå. Jeg har valgt å sortere statistikken slik at man ser verdiene for innvandrere og majoritetsbefolkningen hver for seg. Dette har jeg gjort ettersom det er eventuelle forskjeller mellom gruppene som i hele oppgaven vil være relevant å studere.

3.4.1 Norge

Tabell 3.1

Variabler	Innvandrere		Majoritetsbefolkning	
	Gjennomsnitt	Standardavvik	Gjennomsnitt	Standardavvik
Timelønn	22.09952	8.349453	25.98364	8.523635
Utdanning	14.75244	3.073183	14.65127	2.340415
Leseferdighet	248.2835	62.46656	288.0702	41.40415
Tallforståelse	239.5552	73.19652	290.8638	47.45738

I tabell 3.1 vises deskriptiv statistikken for henholdsvis innvandrere og majoritetsbefolkning, slik at man kan se på eventuelle observerte forskjeller fra spørreundersøkelsen. Først og fremst ser vi at innvandrere i snitt tjener nesten 4 dollar mindre i timen enn majoritetsbefolkningen. Videre ser vi at innvandrene scorer jevnt over lavere på ferdighetstestene, men at gjennomsnittlig utdanningsnivå er så og si likt hos de to gruppene. Ved å se på standardavvikene, ser vi en tydelig tendens til at det er større spredning i ferdighetsnivået til innvandrene.

Tabell 3.2

Helse	Antall	Andel	Etnisitet	Antall	Andel
God helse	3299	0.795	Majoritetsbefolkning	3465	0.871
Dårlig helse	852	0.205	Innvandrere	513	0.129
Totalt: 4151			Totalt: 3978		

I tabell 3.2 har jeg laget en oversikt over de to dummyvariablene som jeg vil bruke videre i oppgaven. Av denne ser vi at nesten 80% av befolkningen regner helsen sin som relativt god i Norge, og når det gjelder befolkningsfordelingen så ser vi at ca. 13% av respondentene er innvandrere.

3.4.2 Italia

For Italia har jeg lagd eksakt tilsvarende tabeller som for Norge, og nedenfor vil jeg ta en rask gjennomgang av de viktigste observasjonene i tabellene.

Tabell 3.3

Variabler	Innvandrere		Majoritetsbefolkning	
	Gjennomsnitt	Standardavvik	Gjennomsnitt	Standardavvik
Timelønn	11.09867	5.78518	15.79131	8.136222
Utdanning	11.55219	3.588383	11.59942	4.101197
Leseferdighet	226.996	49.23663	255.1731	98.97003
Tallforståelse	232.8316	55.81878	252.2269	55.453

Tabell 3.5 viser en statistikk hvor innvandrene og majoritetsbefolkningen er skilt fra hverandre, og i likhet med statistikken for Norge, så ser vi også her at innvandrene jevnt over scorer dårligere på timelønn og ferdigheter. På utdanningsvariabelen ser vi at innvandrene i

Italia ligger på tilsvarende utdanningsnivå som majoritetsbefolkningen. Her ser vi også en tendens til at det er større spredning i ferdighetsnivåene blant majoritetsbefolkningen.

Den siste tabellen viser statistikk for dummyvariablene, og her ser vi at nesten 80% regner helsen sin som god, tilsvarende som vi så for Norge. I tillegg ser vi at under 10% av befolkningen i Italia er innvandrere.

Tabell 3.4

Helse	Antall	Andel	Etnisitet	Antall	Andel
God helse	3271	0.799	Majoritetsbefolkning	3762	0.927
Dårlig helse	824	0.201	Innvandrere	297	0.073
Totalt: 4095			Totalt: 4059		

4. Økonometrisk modell

I denne delen av oppgaven vil jeg gjøre rede for den økonometriske modellen som brukes til å estimere sammenhenger mellom lønnsnivå og det å være innvandrer. Jeg vil diskutere valg av den empiriske strategien og forutsetninger som tilhører den, samt valg av funksjonsform. Deretter vil jeg presentere regresjonsmodellene og tolke disse.

4.1 Empirisk strategi

Jeg benytter metoden *Ordinary Least Squares* (OLS) til å estimere sammenhengen mellom det å være innvandrer versus majoritetsbefolkning og timelønn. Denne metoden gjør at den regresjonslinja som man estimerer er slik at residualene er så små som mulig (Thomas, 2005:266). Med residualer menes avstanden fra faktiske verdier på avhengig variabel til predikerte verdier, så OLS-metoden estimerer verdier slik at summen av de kvadrerte residualene er minimert (Ibid.). Med valg av OLS som metode, ligger det noen antakelser til grunn, og noen av disse vil jeg kort gjøre rede for her. Først og fremst er OLS en metode som estimerer en lineær sammenheng mellom variablene, og det bør derfor sikres at de variablene som brukes gir en relativt lineær sammenheng med lønn. Om ikke dette er tilfellet, vil det være hensiktsmessig å transformere variablene slik at den riktige sammenhengen blir fanget opp. I tillegg forutsetter man at det ikke er multikollinearitet mellom kontrollvariablene, samt at restleddet ikke korrelerer med den avhengige variabelen. Med multikollinearitet menes det at kontrollvariablene ikke burde korrelere med hverandre i særlig grad, fordi dette forstyrrer den faktiske effekten som hver og en av kontrollvariablene har på timelønn. I en OLS-regresjon

legger man også til grunn antakelsen om at det ikke er noen ko-variasjon mellom restleddet og den avhengige variabelen, altså at $\text{Co}(Y, \varepsilon_i) = 0$. De nevnte forutsetningene kan være vanskelig å oppfylle, og hvorvidt modellen blir preget av brudd på disse, vil diskuteres senere i oppgaven.

4.2 Hypotesetest

Når man lager en regresjonsmodell, så ønsker man å teste om den estimerte sammenhengen stemmer, og dette gjøres ved hjelp av en hypotesetest. En hypotesetest innebærer at man definerer en nullhypotese og en alternativhypotese for sammenhengen mellom to variabler. Nullhypotesen defineres som at det ikke er noe sammenheng mellom de to variablene, altså at $\beta_i = 0$, mens alternativhypotesen antar en sammenheng mellom variablene som enten er $\beta_i >$ eller $<$ eller $\neq 0$. Hypotesetesten vil deretter avsløre hvilken av hypotesene man kan beholde, og dermed verifisere eller avkrefte den estimerte sammenhengen mellom variablene som ble gitt av den økonometriske modellen. Jeg vil benytte to ulike hypotesetester i denne oppgaven – såkalt t-test og F-test. F-testen benytter jeg når jeg vil teste sammenhengen til flere variabler samtidig, mens t-testen brukes når jeg bare sjekker den estimerte sammenhengen mellom to variabler. De to hypotesetestene har følgende formel for å regne ut testobservatoren:

$$TS = \frac{b - \beta}{s} \sim t - \text{fordeling}$$

$$TS = \frac{(R^2 - R_1^2)/h}{(1 - R^2)/(N - k)} \sim F - \text{fordeling}$$

Dersom absoluttverdien til testobservatoren (TS) er større enn kritisk verdi for fordelinga, gitt et valgt signifikansnivå, så kan nullhypotesen forkastes. I t-testene benytter jeg at stata også regner ut testobservatoren, den verdien som står i parentes under koeffisientene i modellene.

4.3 Valg av funksjonsform

Jeg har valgt å benytte meg av en log-lineær funksjonsform fordi dette både er vanlig og hensiktsmessig å gjøre når man har lønn som den avhengige variabelen. Ved bruk av en logtransformert lønnsvariabel vil den estimerte koeffisienten foran de andre variablene tilsvare prosentvis økning i lønn uavhengig av målenivå på den avhengige variabelen. Å logtransformere lønnsvariabelen har også flere positive virkninger, blant annet at den tar vekk ekstremverdier på variabelen. I tillegg vil logtransformasjonen gi en lineær sammenheng

mellom den avhengige og de uavhengige variablene, en forutsetning for OLS. Funksjonsformen for regresjonsmodellene vil være som følgende:

$$1) \quad l_{timelønn} = \beta_1 + \beta_2 innvandreri + \beta_3 X_i + \varepsilon_i$$

Funksjonen ovenfor med logtransformert lønnsvariabel gjør at tolkningen av β_2 vil være det prosentvise lønnsgapet som følger av å være innvandrer versus majoritetsbefolkning. β_1 er konstantleddet til modellen, og ε_i er det stokastiske restleddet som fanger opp effekten av alle andre forhold som kan tenkes å påvirke timelønnen, men som ikke eksplisitt er inkludert her. Leddet $\beta_3 X_i$ illustrerer alle kontrollvariablene som jeg etterhvert inkluderer i modellene, og deres effekt på timelønnen registrert av β_i .

For å få tilstrekkelige tall på sammenhengen mellom å være innvandrer og timelønn, samt sjekke for andre relevante forhold, har jeg lagd 4 ulike modeller som beskriver forhold som kan tenkes å påvirke sammenhengen mellom å være innvandrer eller majoritetsbefolkning og timelønn. Disse gjelder for hvert av landene separat. I tillegg har jeg 4 modeller som forsøker å avsløre forskjeller mellom å være innvandrer i Norge eller Italia, med sammenslått datasett. Alle modellene illustreres nedenfor:

Modeller for Norge og Italia separat

$$1) \quad l_{timelønn} = \beta_1 + \beta_2 innvandreri + \varepsilon_i$$

$$2) \quad l_{timelønn} = \beta_1 + \beta_2 innvandreri + \beta_3 utdaninng + \beta_4 leseferdighet + \beta_5 tallforståelse + \varepsilon_i$$

$$3) \quad l_{timelønn} = \beta_1 + \beta_2 innvandreri + \beta_3 utdaninng + \beta_4 leseferdighet + \beta_5 tallforståelse + \beta_6 helse + \varepsilon_i$$

$$4) \quad l_{timelønn} = \beta_1 + \beta_2 innvandreri + \beta_3 utdaninng + \beta_4 leseferdighet + \beta_5 tallforståelse + \beta_6 helse + \beta_7 (utdanning - gjennomsnittlig utdanning) * innvandreri + \varepsilon_i$$

Modeller for Norge og Italia sammenslått

$$1) \quad l_{timelønn} = \beta_1 + \beta_2 innvandreri + \varepsilon_i$$

$$2) \quad l_{timelønn} = \beta_1 + \beta_2 innvandreri + \beta_3 Norway + \beta_4 innvandreri * Norway + \varepsilon_i$$

$$3) \text{ ltimelønn} = \beta_1 + \beta_2 \text{innvandrer}_i + \beta_3 \text{Norway} + \beta_4 \text{innvandrer} * \text{Norway} + \beta_5 \text{utdanning} + \beta_6 \text{leseforståelse} + \beta_7 \text{tallforståelse} + \beta_8 \text{helse} + \epsilon_i$$

$$4) \text{ ltimelønn} = \beta_1 + \beta_2 \text{innvandrer}_i + \beta_3 \text{Norway} + \beta_4 \text{innvandrer} * \text{Norway} + \beta_5 \text{utdanning} + \beta_6 \text{leseforståelse} + \beta_7 \text{tallforståelse} + \beta_8 \text{helse} + \beta_9 \text{innvandrer} * \text{helse} + \epsilon_i$$

Etterhvert i diskusjonen av modellenes styrker og svakheter, vil jeg komme inn på hvilke variabler i tillegg til de overnevnte som kunne vært relevante å ta med for å gjøre analysen av sammenhengen mellom timelønn og om man er innvandrer eller ei, mer riktig. Dersom potensielt innflytelsesrike variabler blir utelatt fra en analyse, så trosser man på mange måter antakelsen om at det ikke er noe ko-variasjon mellom timelønn og restleddet. Her vil jeg imidlertid argumentere for at det ikke var mulig med hensyn på plassbegrensningen å inkludere alle variabler av interesse. Så til tross for at flere forhold, som for eksempel arbeidserfaring og yrkesretning, med stor sannsynlig korrelerer med lønnsnivået, ble det tatt en vurdering basert på begrensning av oppgavens omfang, som gjorde at disse ikke ble inkludert.

I gjennomgangen av regresjonsmodellene vil jeg hovedsakelig kommentere tre verdier – b-koeffisienten, R^2 og signifikansstjernene. Disse verdiene sier noe om henholdsvis effekten av en marginal økning i uavhengig variabel/kontrollvariablene på den avhengige, forklaringskraften til modellen, samt om funnene er signifikante og dermed kan sies å gjelde på et populasjonsnivå. Sistnevnte vil jeg også undersøke ved hjelp av t- og F-tester, og i t-testene benytter jeg verdiene som står i parentes som testobservatorer, og dette vil jeg komme nærmere innpå senere i oppgaven.

4.4 Regresjonsmodell for Norge

Her presenterer jeg 4 ulike modeller som beskriver effekten av de uavhengige variablene på lønnsnivået i Norge, og jeg vil gå gjennom modell 1-4 nedenfor.

TABELL 4.1

	Modell 1 Timelønn	Modell 2 Timelønn	Modell 3 Timelønn	Modell 4 Timelønn
Innvandrer	-0.1760011*** (-10.18)	-0.1190602*** (-7.13)	-0.1192254*** (-7.13)	-0.1192195*** (-7.07)
Utdanning		0.0379076*** (15.35)	0.0377891*** (15.24)	0.0377919*** (13.99)

Leseferdighet		-0.0644364*** (-5.55)	-0.064069*** (-5.50)	-0.0640965*** (-5.50)
Tallforståelse		0.1306125*** (10.93)	0.1301068*** (10.85)	0.1301074*** (10.85)
Helse			0.008517 (0.51)	0.0085152 (0.51)
Utdanning* Innvandrere				0.0000153 (-0.00)
Konstant	3.209341*** (529.87)	2.62363*** (71.76)	2.617885*** (68.47)	2.617843*** (63.49)
N	3020	3019	3019	3019
R ²	0.0332	0.2054	0.2054	0.2054

t-statistikk i parentesene

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.4.1 Modell 1

$$1) \text{ timelønn} = \beta_1 + \beta_2 \text{innvandrere}_i + \varepsilon_i$$

I modell 1 har jeg kun estimert sammenhengen mellom timelønn og det å være innvandrer versus majoritetsbefolkning i Norge. Av den estimerte b-koeffisienten ser vi at det å være innvandrer vil gi en lavere timelønn på ca. 17.6% kontra hvis man er en del av majoritetsbefolkningen. Ved hjelp av en t-test sjekker vi om denne sammenhengen kan verifiseres:

$$H_0: \beta_2 = 0 \text{ mot } H_1: \beta_2 \neq 0$$

$$TS = -10.18 \text{ og velger forkastningsnivå} = 0.05$$

$$t_{0.05/2, 3020} = -1.96$$

→ - $TS < -t$, altså kan vi forkaste nullhypotesen

Dette betyr at funnet om at det er et lønnsgap mellom innvandrere og majoritetsbefolkningen er statistisk signifikant, noe vi også ser av signifikansstjernene. Modellen har en R²-verdi på 0.0332, noe som betyr at 3.3% av variansen og effekten på timelønn kan forklares av den uavhengige variabelen.

4.4.2 Modell 2

$$2) \text{ timelønn} = \beta_1 + \beta_2 \text{innvandreri} + \beta_3 \text{utdaninng} + \beta_4 \text{leseferdighet} + \beta_5 \text{tallforståelse} + \varepsilon_i$$

I den neste modellen har jeg inkludert utdanningsnivå, samt leseferdigheter og tallforståelse. Den første umiddelbare endringen som er verdt å legge merke til er at forklaringskraften nå har økt til litt over 20%, noe som indikerer at utdanning og ferdigheter i stor grad spiller inn på lønnsnivået i Norge. Videre ser vi også at det fortsatt er en negativ sammenheng mellom det å være innvandrere og timelønn, men denne effekten er svakere her (ca. 12%) enn i forrige modell. Det betyr at det å ta utdanning og score høyere på ferdighetstestene demper effekten av å være innvandrere på timelønn, altså kan lønnsgapet forklares av forskjeller i utdanning og ferdighet. For å kunne stadfeste at denne sammenhengen stemmer, vil det være relevant å gjøre en F-test for å se om utdanning og ferdigheter har en signifikant effekt på timelønn:

$$H_0: \beta_2 = \beta_3 = \beta_4 = 0 \text{ mot } H_1: \beta_2 \text{ og/eller } \beta_3 \text{ og/eller } \beta_4 \neq 0$$

$$TS = \frac{(R^2 - R_1^2)/h}{(1 - R^2)/(N - k)} = \frac{(0.2054 - 0.0332)/3}{(1 - 0.2054)/(3020 - 5)} = 217.8$$

$$F_{3, 3015} = 1$$

→ $TS > F$, altså kan vi forkaste nullhypotesen

Av dette ser vi at utdanning og ferdigheter har en signifikant effekt på timelønn, og av b-koeffisientene kan vi se størrelsen på de. Effekten av ett års ekstra utdanning på timelønn ligger på ca. 4%, men det kan tenkes at denne effekten kan være sterkere/svakere avhengig av utdanningsnivået. Videre ser vi at et standardavviks økning i leseferdighet vil gi 6.4% lavere timelønn, men at et standardavviks økning i tallforståelse vil gi 13% høyere timelønn. Alle funnene i modell 2 er statistisk signifikante på et 0.01-nivå.

4.4.3 Modell 3

$$3) \text{ timelønn} = \beta_1 + \beta_2 \text{innvandreri} + \beta_3 \text{utdaninng} + \beta_4 \text{leseferdighet} + \beta_5 \text{tallforståelse} + \beta_6 \text{helse} + \varepsilon_i$$

Modell 3 inkluderer i tillegg til utdanning og ferdigheter også en variabel som sier noe om effekten av helsenivå på timelønn. Helsevariabelen er dummykodet slik at de som har ”god” helse scorer 1, mens de med ”dårlig” helse scorer 0. I modellen ser vi at det er et lønnsgap mellom de med god og dårlig helse på nesten 1%, i favør de med god helse. Dette funnet er imidlertid ikke signifikant, og man kan dermed ikke hevde at denne sammenhengen gjelder

utenom de observasjonene som er gjort i undersøkelsen. Dette kan vi også undersøke ved hjelp av en t-test:

$$H_0: \beta_6 = 0 \text{ mot } H_1: \beta_6 \neq 0$$

$$TS = 0.51 \text{ og velger forkastningsnivå} = 0.05$$

$$t_{0.05/2, 3019} = 1.96$$

→ $TS < t$, altså beholdes nullhypotesen

Vi kan ikke forkaste nullhypotesen om at lønnsgapet mellom innvandrere og majoritetsbefolkningen er uavhengig av helsenivået. Ellers i modellen holder de andre b-koeffisientene seg relativt like, og det dermed ikke noe tegn til at det å ha god helse demper effekten av å være innvandrer på timelønnen.

4.4.4 Modell 4

$$4) \text{ } l_{\text{timelønn}} = \beta_1 + \beta_2 \text{innvandrer}_i + \beta_3 \text{utdanning} + \beta_4 \text{leseferdighet} + \beta_5 \text{tallforståelse} + \beta_6 \text{helse} + \beta_7 (\text{utdanning} - \text{gjennomsnittlig utdanning}) * \text{innvandrer} + \varepsilon_i$$

I den siste regresjonsmodellen har jeg inkludert et samspillsledd mellom det å være innvandrer og utdanning. Samspillsleddet er på formen $(\text{utdanning} - \text{gjennomsnittlig utdanning}) * \text{innvandrer}$, en formulering som gir en nyttig tolkning av koeffisienten β_7 . Hadde man bare inkludert et vanlig interaksjonsledd så ville tolkningen blitt lønnsforskjell for innvandrere og majoritetsbefolkning med 0 års utdanning, og dette er jo observasjoner som jeg allerede har utelatt fra datasettet. B_2 har en verdi på -0.119 som kan tolkes slik at innvandrere har en lavere timelønn enn majoritetsbefolkningen på 11.9%, også når man kontrollerer for et gjennomsnittlig utdanningsnivå. Dette funnet underbygges også av en t-test, som man ser nedenfor:

$$H_0: \beta_2 = 0 \text{ mot } H_1: \beta_2 \neq 0$$

$$TS = -7.07 \text{ og velger forkastningsnivå} = 0.05$$

$$t_{0.05/2, 3019} = -1.96$$

→ $-TS < -t$, altså kan vi forkaste nullhypotesen

B-koeffisienten på -0.119 er så og si lik som i forrige modell, altså endres ikke lønnsgapet av at man kontrollerer for gjennomsnittlig utdanningsnivå. Videre ser vi at β_7 har en minimal

verdi, som tilsier at når utdanninga avviker fra gjennomsnittet så har det ingen signifikant effekt på lønna. Også dette funnet kan verifiseres av en t-test:

$$H_0: \beta_7 = 0 \text{ mot } H_1: \beta_7 \neq 0$$

$$TS = -0.00 \text{ og velger forkastningsnivå} = 0.05$$

$$t_{0.05/2, 3019} = -1.96$$

→ - TS > - t, altså beholdes nullhypotesen

Forklaringskraften ligger fortsatt på 20.5%, og de andre b-koeffisientene har holdt seg stabile.

4.4.5 Oppsummering

Kort oppsummert indikerer regresjonsmodellene at det er et lønnsgap på alt mellom 11.9 og 17.6% mellom innvandrere og majoritetsbefolkningen i Norge, i favør majoritetsbefolkningen. Den negative sammenhengen for innvandrerne blir dempet av utdanningsnivå og andre ferdigheter. Helsenivået til respondentene kan ikke sies å være særlig relevant for å si noe om timelønnen til innvandrere versus majoritetsbefolkningen, ettersom b-koeffisienten var veldig liten og funnene ikke er signifikante på et respektabelt nivå. Ved å kontrollere for gjennomsnittlig utdanningsnivå, så vi at innvandrerne fortsatt har lavere lønn på ca. 12%, men at utdanningsnivå over eller under gjennomsnittet ikke har noen signifikant effekt på lønna. Variasjonen og effekten på timelønn kan i svært liten grad forklares av om man er innvandrer eller ikke, mens faktorer som utdanning og ferdigheter synes viktigere for å forklare variasjoner i lønnsnivået.

4.5 Regresjonsmodell for Italia

Regresjonsmodellen for Italia er helt tilsvarende oppsettet for Norge, og jeg vil på samme måte som i delkapittel 4.2 gå igjennom modellene hver for seg.

TABELL 4.2

	Modell 1 Timelønn	Modell 2 Timelønn	Modell 3 Timelønn	Modell 4 Timelønn
Innvandrer	-0.3338657*** (-9.13)	-0.2705122*** (-7.79)	-0.2691339*** (-7.75)	-0.2667795*** (-7.36)
Utdanning		0.0370774*** (13.53)	0.037275*** (13.56)	0.0379472*** (13.36)

Leseferdighet		-0.0075743 (-0.44)	-0.0073172 (-0.43)	-0.0072009 (-0.42)
Tallforståelse		0.0670986*** (3.85)	0.0671257*** (3.85)	0.067018*** (3.84)
Helse			-0.0243237 (-0.87)	-0.0245377 (-0.88)
Utdanning* Innvandrere				-0.0089519 (-0.95)
Konstant	2.656055*** (251.46)	2.175428 *** (62.55)	2.194004*** (53.72)	2.185751*** (52.34)
N	1846	1846	1846	1846
R ²	0.0432	0.1895	.01898	0.1902

t-statistikk i parentesene

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.5.1 Modell 1

$$1) \text{ } l_{\text{timelønn}} = \beta_1 + \beta_2 \text{innvandrere}_i + \varepsilon_i$$

I modell 1 ser vi et estimert lønnsgap på hele 33% mellom innvandrere og majoritetsbefolkningen i Italia, en forskjell som er mye større enn i Norge. At man observerer et lønnsgap mellom de to folkegruppene kan også kontrolleres for ved hjelp av en t-test:

$$H_0: \beta_2 = 0 \text{ mot } H_1: \beta_2 \neq 0$$

$$TS = -7.07 \text{ og velger forkastningsnivå} = 0.05$$

$$t_{0.05/2, 1860} = -1.96$$

→ - TS < -t, altså kan vi forkaste nullhypotesen

Variansen og effekten på timelønn forklares 4% av om man er innvandrere eller ei.

4.5.2 Modell 2

$$2) \text{ } l_{\text{timelønn}} = \beta_1 + \beta_2 \text{innvandrere}_i + \beta_3 \text{utdanning} + \beta_4 \text{leseferdighet} + \beta_5 \text{tallforståelse} + \varepsilon_i$$

I modell 2 så ser vi at effekten av å være innvandrere har blitt dempet som følge av at utdannings- og ferdighetsvariablene er inkludert. Nå estimeres en negativ sammenheng mellom det å være innvandrere og timelønn på ca. 27%, et signifikant funn. For å kunne besvare problemstillingen om utdanning og ferdigheter er med på å forklare lønnsgapet, må man først og fremst sjekke om det finnes en statistisk effekt på lønna av utdanning og ferdigheter, ved hjelp av en F-test:

$$H_0: \beta_2 = \beta_3 = \beta_4 = 0 \text{ mot } H_1: \beta_2 \text{ og/eller } \beta_3 \text{ og/eller } \beta_4 \neq 0$$

$$TS = \frac{(R2^2 - R1^2)/h}{(1 - R2^2)/(N - k)} = \frac{(0.1895 - 0.0432)/3}{(1 - 0.1895)/(1846 - 5)} = 110.77$$

$$F_{3, 1841} = 1$$

→ $TS > F$, altså kan vi forkaste nullhypotesen

I tillegg til at vi nå har at utdanning og ferdigheter gir signifikant effekt på lønna, er det også relevant å si noe om størrelsen på disse effektene. Vi ser at ett års høyere utdanning gir en lønnsøkning per time på rundt 3%. I likhet med modellen for Norge ser vi også at i Italia så vil et standardavviks økning i leseferdigheter gi en liten negativ redusering i timelønn, men dette funnet er ikke statistisk signifikant. Et standardavviks økning i tallforståelse derimot gir en anslått prosentvis timelønnsøkning på ca. 6%, og dette er statistisk signifikant. Forklaringskraften økes til over 18% når kontrollvariablene ble lagt til.

4.5.3 Modell 3

$$3) \text{ } l_{\text{timelønn}} = \beta_1 + \beta_2 \text{innvandreri} + \beta_3 \text{utdaninng} + \beta_4 \text{leseferdighet} + \beta_5 \text{tallforståelse} + \beta_6 \text{helse} + \varepsilon_i$$

I den neste modellen har jeg inkludert helsevariabelen, og man ser at denne variabelen ikke endrer effekten av å være innvandreri på timelønn i særlig grad. Den andre b-koeffisientene holder seg også like som i forrige modell. B_6 , koeffisienten til helsevariabelen, tilsier at det er et lønns gap på 2% mellom de med god og dårlig helse, i favør de med dårlig helse. Dette virker kanskje litt ulogisk, men funnet er ikke signifikant og kan dermed ikke sies å gjelde utover de observasjonene som er i denne undersøkelsen. Antakelsen om at dette funnet var ulogisk underbygges av en t-test nedenfor:

$$H_0: \beta_6 = 0 \text{ mot } H_1: \beta_6 \neq 0$$

$$TS = 0.51 \text{ og velger forkastningsnivå} = 0.05$$

$$t_{0.05/2, 3019} = 1.96$$

→ $TS < t$, altså beholdes nullhypotesen

Det er ingen signifikant effekt av å ha god/dårlig helse på timelønn. Forklaringskraften til modell 3 endres ikke av særlig betydning av at helsevariabelen er inkludert.

4.5.4 Modell 4

$$4) \text{ timelønn} = \beta_1 + \beta_2 \text{innvandrer}_i + \beta_3 \text{utdaninng} + \beta_4 \text{leseferdighet} + \beta_5 \text{tallforståelse} + \beta_6 \text{helse} + \beta_7 (\text{utdanning} - \text{gjennomsnittlig utdanning}) * \text{innvandrer} + \varepsilon_i$$

I den siste regresjonsmodellen for Italia har jeg igjen inkludert et interaksjonsledd mellom innvandrer og utdanning, for å kunne si noe om avkastning av utdanning på lønn er ulik for de to befolkningsgruppene. Av estimatene ser vi at det β_7 har en verdi på -0.0089, som indikerer at når utdanningsnivået avviker fra gjennomsnittet, så tjener innvandrer litt mindre enn majoritetsbefolkningen. Dette funnet er imidlertid ikke statistisk signifikant, som vi ser av t-testen nedenfor:

$$H_0: \beta_7 = 0 \text{ mot } H_1: \beta_7 \neq 0$$

$$TS = -0.00 \text{ og velger forkastningsnivå} = 0.05$$

$$t_{0.05/2, 3019} = -1.96$$

→ - TS > - t, altså beholdes nullhypotesen

Det er også relevant å nevne at b-koeffisienten for innvandrervariabelen, forblir som den har vært, noe som indikerer at når man kontrollerer for gjennomsnittlig utdanningsnivå, så er effekten av å være innvandrer versus majoritetsbefolkning på timelønn den samme, slik vi også så for Norge. Forklaringskraften til modellen er på 19%. Det finnes med andre ord mange andre faktorer som ikke er inkludert her som spiller inn på variasjonen og effekten på timelønn.

4.5.5 Oppsummering

Kort oppsummert så har vi sett de samme effektene og tendensene på lønnsgap mellom innvandrere og majoritetsbefolkningen i Italia som vi så for Norge. Det virker imidlertid til at den rene effekten av å være innvandrer, altså når man ikke kontrollerer for andre forhold, er sterkere i Italia. Det vil bli interessant å se om flere ulikheter mellom de to landene kan avdekkes i den neste regresjonsmodellen, hvor jeg kombinerer de to landene i samme datasett.

4.6 Regresjonsmodell for Norge og Italia

Hensikten med en regresjonsmodell hvor datasettene for Norge og Italia er slått sammen, er å kunne undersøke om det finnes forskjeller i timelønn som følge av at man bor i, og er innvandrer i ulike land. Om det finnes estimerte forskjeller her, vil det være interessant å knytte det opp

mot økonomisk vekstteori for å si noe om det kan være forskjeller mellom å være innvandrere til land med ulik økonomisk vekst/nivå.

TABELL 4.3

	Modell 1 Timelønn	Modell 2 Timelønn	Modell 3 Timelønn	Modell 4 Timelønn
Innvandrere	-0.1882385*** (-9.74)	-0.3338657*** (-14.61)	-0.2723992*** (-12.76)	-0.2596608*** (-5.44)
Norway		0.5532855*** (55.53)	0.4643134*** (47.48)	0.4642953*** (47.47)
Innvandrere*Norway		0.1578646*** (5.05)	0.1472442*** (5.11)	0.1470516*** (5.10)
Utdanning			0.0372573*** (24.98)	0.0372547*** (24.98)
Leseferdighet			-0.0266788*** (-3.12)	-0.0266861*** (-3.12)
Tallforståelse			0.0892803*** (10.19)	0.089272*** (10.18)
Helse			-0.0083949 (-0.63)	-0.0071784 (-0.52)
Innvandrere*helse				-0.0140814 (-0.30)
Konstant	2.898995*** (471.64)	2.656055 *** (402.30)	2.179*** (100.93)	2.17798*** (99.63)
N	6712	6712	6712	6711
R ²	0.0139	0.3621	0.4634	0.4634

t-statistikk i parentesene

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.6.1 Modell 1

$$1) \text{ ltimelønn} = \beta_1 + \beta_2 \text{innvandrere}_i + \varepsilon_i$$

Her har jeg i likhet med de separate regresjonsmodellene for hvert land, kun inkludert dummyvariabelen for om man er innvandrere eller ikke, for å si om det er en generell trend for landene når det kommer til timelønn og hvilken befolkningsgruppe man tilhører. Vi ser at det er estimert et lønns-gap på 18%, slik at de som er innvandrere forventes å ha lavere timelønn enn majoritetsbefolkningen. Dette virker rimelig, ettersom det er en verdi som ligger mellom funnene for hvert av landene separat. I denne modellen ser vi at forklaringskraften kun er på ca. 1%, altså er det mange andre faktorer som spiller inn på timelønn.

4.6.2 Modell 2

$$2) \textit{timelønn} = \beta_1 + \beta_2 \textit{innvandreri} + \beta_3 \textit{Norway} + \beta_4 \textit{innvandreri} * \textit{Norway} + \varepsilon_i$$

Videre i modell 2 har jeg inkludert en dummyvariabel som kalles Norway. Denne er kodet slik at dersom man bor i Norge scorer man 1 og dersom man bor i Italia scorer man 0. Denne variabelen, samt interaksjonsleddet mellom de to dummyvariablene innvandreri og Norway, gjør det mulig å si om effekten på timelønn er ulik for befolkningsgruppene i de to landene. I denne modellen ser vi umiddelbart at β_2 , som her tolkes som lønnsforskjellen mellom innvandrere og majoritetsbefolkningen i Italia, er på hele 33%. Dette betyr at lønnsgapet mellom folkegruppene i Italia er mye større enn i Norge. Av koeffisienten for Norway-variabelen (β_3) ser vi at majoritetsbefolkningen i Norge tjener rundt 55% mer enn majoritetsbefolkningen i Italia, og dette funnet er statistisk signifikant. Det er hittil en klar tendens til at lønnsnivået generelt er høyere i Norge enn i Italia, både for innvandrere og majoritetsbefolkningen. Videre ser vi på β_4 at lønnsgapet mellom innvandrere og majoritetsbefolkningen er nesten 16% lavere i Norge enn i Italia, altså at forskjellen i inntekt for innvandrere i Norge er mindre enn i Italia. Dette funnet har jeg testet nedenfor:

$$H_0: \beta_4 = 0 \text{ mot } H_1: \beta_4 \neq 0$$

$$TS = 5.05 \text{ og velger forkastningsnivå} = 0.05$$

$$t_{0.05/2, 6712} = 1.96$$

→ $TS > t$, altså forkastes nullhypotesen

Altså har vi en statistisk signifikant observasjon på at det er forskjeller i lønnsgapet mellom innvandrere og majoritetsbefolkningen i Italia og Norge, i favør innvandrerne i Norge. Verdien på R^2 er nå på ca. 36%, altså kan man si at variasjonen og effekten på timelønn i relativt stor grad avhenger av om man er fra Norge eller Italia, samt i hvilket av de landene man er innvandrer.

4.6.3 Modell 3

$$3) \textit{timelønn} = \beta_1 + \beta_2 \textit{innvandreri} + \beta_3 \textit{Norway} + \beta_4 \textit{innvandreri} * \textit{Norway} + \beta_5 \textit{utdanning} + \beta_6 \textit{leseforståelse} + \beta_7 \textit{tallforståelse} + \beta_8 \textit{helse} + \varepsilon_i$$

I modell 3 har jeg inkludert kontrollvariablene utdanning, leseferdighet, tallforståelse og helse, for å se om disse faktorene kan være med på å forklare de observerte lønnsgapene som finnes. Den estimerte effekten av å være innvandrer synker til 27%, som nå er det estimerte lønnsgapet

mellom innvandrere og majoritetsbefolkningen. Også effekten av å være norsk, samt effekten av å være innvandrere i Norge dempes når man kontrollerer for utdanning, ferdigheter og helse. Dette tilsier at jo høyere utdanning og ferdighetsnivå man har, jo mindre relevant er hvilket land man bor i for timelønnen. Om det er slik at de observerte lønnsgapene avhenger av ferdigheter, utdanning og helse, har jeg testet ved hjelp av en F-test.

$$H_0: \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0 \text{ mot } H_1: \beta_5 = \beta_6 = \beta_7 = \beta_8 \neq 0$$

$$TS = \frac{(R^2 - R_1^2)/h}{(1 - R^2)/(N - k)} = \frac{(0.4634 - 0.3621)/4}{(1 - 0.4634)/(6712 - 8)} = 110.77$$

$$F_{4, 6704} = 1$$

→ $TS > F$, altså kan vi forkaste nullhypotesen

Av β_5 ser vi at et års høyere utdanning gir ca. 3.7% høyere lønn. Det er en svak negativ sammenheng mellom leseferdigheter og timelønn, samt en litt sterkere positiv sammenheng mellom tallforståelse og timelønn. Helsevariabelen er svakt negativ, noe som tilsier at god helse ikke nødvendigvis gir høyere timelønn, men dette funnet er ikke statistisk signifikant. Altså kan man konkludere med at særlig utdanning og ferdigheter kan forklare lønnsgapet mellom innvandrere og majoritetsbefolkningen, men at effekten av helsenivå er mer usikker. I likhet med tidligere modeller som inkluderer kontrollvariablene, økes forklaringskraften også her. Ca. 46% av variasjonen og effekten på timelønn kan forklares av de inkluderte variablene.

4.6.4 Modell 4

$$4) \text{ } l_{\text{timelønn}} = \beta_1 + \beta_2 \text{innvandrere}_i + \beta_3 \text{Norway} + \beta_4 \text{innvandrere} * \text{Norway} + \beta_5 \text{utdanning} + \beta_6 \text{leseforståelse} + \beta_7 \text{tallforståelse} + \beta_8 \text{helse} + \beta_9 \text{innvandrere} * \text{helse} + \varepsilon_i$$

I den siste modellen har jeg inkludert et interaksjonsledd mellom innvandrere og helsenivå, for å kunne avdekke om det finnes forskjeller i effekt på timelønn for innvandrere med god versus dårlig helse. Jeg starter å se på β_9 -koeffisienten, som har en verdi på -0.259, som kan tolkes slik at blant de med dårlig helse så tjener majoritetsbefolkningen 26% bedre enn innvandrerne. β_9 , koeffisienten for interaksjonsleddet, viser oss at lønnsgapet mellom innvandrere og majoritetsbefolkningen er 1.4% større for de med god enn de med dårlig helse. Dette funnet er ikke statistisk signifikant, og en t-test kunne ikke avvise nullhypotesen:

$$H_0: \beta_9 = 0 \text{ mot } H_1: \beta_9 \neq 0$$

$$TS = -0.30 \text{ og velger forkastningsnivå} = 0.05$$

$$t_{0.05/2, 6711} = -1.96$$

→ $-TS < -t$, altså beholdes nullhypotesen

Dermed kan vi ikke si at god helse nødvendigvis har en positiv effekt på lønnsgapet for innvandrere, til tross for at logisk tenking tilsier at god helse gir en lettere arbeidssituasjon og lavere risiko for å være sykemeldt/arbeidsledig.

5. Diskusjon

I denne delen av oppgaven vil jeg forsøke å knytte det teoretiske grunnlaget opp mot funnene mine i de økonometriske modellene, for å kunne besvare problemstillingene enda bedre. Ettersom modellene estimerte stort sett de samme tendensene i Italia og Norge når det kommer til sammenhengen mellom timelønn og det å være innvandrer, vil jeg i det første delkapittelet diskutere hvorvidt funnene fra modell 4.1 og 4.2 passer inn med teori jeg har brukt. Videre vil jeg forsøke å knytte ulikhetene rundt det å være innvandrer i Norge versus Italia opp mot teorien om økonomisk vekst, for å forsøke å besvare om det er bedre for lønnsnivået til innvandrere å bo i det ene eller andre landet. Avslutningsvis i diskusjonskapittelet vil jeg gjøre rede for styrkene og svakhetene i de økonometriske modellene, som jeg var litt inne på tidligere i oppgaven.

5.1 Hvordan vil det å være innvandrer spille inn på lønnsnivået?

Hovedproblemstillingen for oppgaven dreide seg rundt om man observerer lønnsforskjeller mellom innvandrere og majoritetsbefolkningen. Gjennom regresjonsmodellene, både for Italia og Norge, har vi sett en tydelig tendens til at det å være innvandrer er negativt for lønnsnivået. For Norge ble det anslått et lønnsgap som varierte mellom 18 og 12%, og for Italia var denne effekten noe høyere. Disse funnene sammenfaller i stor grad med tidligere litteratur på feltet, og det ble i teorikapittelet argumentert for at dette lønnsgapet har opphav i såkalt "landspesifikk humankapital". Dette stemmer godt med funnene i den økonometriske modellen, ettersom at når det ble kontrollert for utdanning og ferdigheter, ble lønnsgapet til innvandrerne redusert. Altså kan man anta at utdanning, leseferdigheter og tallforståelse spiller positivt inn på lønnsnivået til både innvandrere og majoritetsbefolkning, men at lønnsgapet fortsatt er betydelig til tross for dette. Den landspesifikke humankapitalen, som kan innebære sosiale nettverk, språkkunnskaper osv., virker sentral for å forklare inntektsforskjellene. Interaksjonsleddet mellom innvandrer og utdanning ble inkludert i Modell 4, i forsøk på å

avsløre ulik avkastning av utdanning mellom de to folkegruppene. Her kunne jeg imidlertid ikke finne noen relevante signifikante funn.

Tanken bak det å inkludere en helsevariabel, var å sjekke om det finnes forskjeller i helsenivå som kan brukes til å forklare inntektsgapet som finnes mellom innvandrere og majoritetsbefolkningen. I teoridelen har jeg ikke eksplisitt forklart hvilken sammenheng det å ha god helse har å si for lønna, men det er naturlig å anta at god helse i stor grad reduserer risikoen for å være arbeidsledig. Har man god helse så burde man derfor ha gode forutsetninger for å være i arbeid, og dermed ha høyere lønn enn de som har dårlig helse. I Modell 3 fikk jeg imidlertid resultater som tilsier at det ikke er noen statistisk sammenheng mellom det å ha god helse og høyere lønn. I tillegg holdt koeffisienten til innvandrervariabelen seg lik som i modellen før, slik at man ikke kan si at lønnsgapet mellom folkegruppene kan forklares av forskjeller i helsenivå.

5.2 Er det bedre eller verre å være innvandrer i det ene eller andre landet?

Den siste problemstillingen min for oppgaven dreide seg rundt om det er fordeler eller ulemper knyttet til å være innvandrer i et land med god versus dårlig økonomi/økonomisk vekst. Teorigrunnlaget her tilsa at det er en toveis-sammenheng mellom økonomisk vekst og forhold som utdanning, ferdigheter, helsenivå osv. Jeg valgte å benytte den linken som tilsier at et land med god økonomisk vekst/god økonomi har bedre forutsetninger for å kunne gi befolkningen sin gode utdanningsmuligheter, et velfungerende arbeidsmarked og bedre helsesituasjon. Faktorer som alle drar i retning av et høyere lønnsnivå. I den økonometriske modellen hvor Italia og Norge var sammenslått kunne vi også se at dette synes å stemme. Det ble estimert et større lønnsgap mellom folkegruppene i Italia enn i Norge, og til tross for at lønnsgapet ble dempet da man kontrollerte for utdanning, ferdigheter og helse, var forskjellen fortsatt betydelig. Modellene kunne imidlertid ikke avdekke om disse forskjellene skyldtes dårligere helsenivå blant befolkningen i Italia. Med bakgrunn i det teoretiske rammeverket, kan man kanskje argumentere for at grunnen til at lønnsgapet er større i Italia enn i Norge, skyldes de interne problemene i Italia hvor den store arbeidsinnvandringen gjør at innvandrerne havner nederst på "rangstigen" i samfunnet. Dette igjen kan forklares av at Italia, med en dårlig økonomi, ikke har like store muligheter til å legge til rette for en god integrering av innvandrerne. Det faktum at det er et større lønnsgap mellom folkegruppene i Italia, kan man argumentere for at skyldes de interne problemene som følger med å ikke ha like stort økonomisk

spillerom som Norge har, noe som nødvendigvis vil gå utover arbeidsmarkedet og samfunnsstrukturen.

5.3 Styrker og svakheter i oppgaven

I metode- og økonometrikapittelet ble det gjort rede for utfordringer knyttet til den kvantitative metoden, datasettet og OLS. Først og fremst byr den kvantitative metoden på utfordringer, ettersom det å kvantifisere sosiale og samfunnsmessige fenomener ikke alltid er like lett. Dette henger også tett sammen med utfordringer knyttet til datasettet. Lønnsvariabelen jeg har valgt å bruke er i seg selv kvantifiserbar, men når det kommer til mål på utdanning, ferdigheter og helsenivå så må man stille seg kritisk til hvor valide observasjonene er. Er det rimelig å anta at mål på utdanning og ferdigheter i to ulike land er fullt sammenliknbare, eller er det avvik mellom de som forstyrrer en faktisk sammenheng med lønn? Helsevariabelen i datasettet valgte jeg å generere om til en dummyvariabel, for å lette den videre analysen. Her valgte jeg det slik at de som hadde svart 1,2 eller 3 fikk verdien god helse, mens de som stod på 4 eller 5 ble regnet som mennesker med dårlig helse. Man kan imidlertid spørre seg hvilke forskjeller som finnes mellom de som svarte 3 eller 4, og om en slik deling mellom de verdiene gir et helt riktig bilde på helsesituasjonen til en person eller ei. Uansett så jeg dette som en bedre løsning enn å beholde de 5 nivåene for helse. Man kan altså argumentere for at den økonometriske analysen kan være preget av operasjonelle utfordringer knyttet til variablene.

Videre ønsker jeg å diskutere litt rundt utfordringene med OLS som metode, ettersom man ved bruk av enhver metode legger til grunn noen forutsetninger som skal holdes for at regresjonen skal være riktig. To av forutsetningene for OLS var at det ikke fantes betydelig multikollinearitet, samt at det ikke er noen ko-variasjon mellom den avhengige variabelen og restleddet. Multikollinearitet beskriver, som nevnt tidligere, en situasjon der kontrollvariablene i modellen korrelerer med hverandre. Utfordringen med dette er at en slik intern korrelasjon forstyrrer den enkelte kontrollvariabelens effekt på den avhengige variabelen. Særlig er det grunn til å tro at det finnes en slik multikollinearitet mellom de to variablene fra ferdighetstestene, leseferdighet og tallforståelse. Nedenfor illustreres en korrelasjonsmatrise som viser dette:

Korrelasjonsmatrise	Leseferdighet	Tallforståelse	Utdanning
Leseferdighet	1.0000		
Tallforståelse	0.8523	1.0000	
Utdanning	0.3854	0.4058	1.0000

Av denne matrisen ser vi at det er en sterk korrelasjon på 0.85 mellom leseferdighet og tallforståelse. Dette betyr at blant hvis man scorer høyt på den ene variabelen, så er det 85% sannsynlighet for at man også scorer høyt på den andre. Dette i seg selv er ikke noe problem, men en slik multikollinearitet kan som sagt forstyrre den rene effekten som hver av variablene har på timelønn. Også utdanning korrelerer med ferdighetsmålene, men ikke i like stor grad.

Også forutsetningen om at restleddet ikke skal ko-varierte med lønnsvariabelen står sentralt i OLS-metoden. Det er derfor viktig at modellene inkluderer alle variabler som kan tenkes å påvirke effekten på lønn, slik at ikke koeffisientene blir feilaktig anslått. I modell 1 for Norge ble det estimert en effekt av å være innvandrere på timelønn på nesten 18%. Da vi inkluderte noen kontrollvariabler så sank denne imidlertid til ca. 12%. Dette vitner om at modell 1 ikke var spesielt presis til å estimere sammenhengen mellom innvandrere og timelønn, fordi det ikke ble kontrollert for andre viktige forhold. Her har altså restleddet ”absorbert” de andre forholdene som spiller inn, og sånn sett gjort at modellen har estimert en for stor b-koeffisient. Dette problemet vil ofte være til stede i en regresjonsmodell, fordi det er utfordrende å kontrollere for alle forhold som kan spille inn på det man studerer. En svakhet i denne oppgaven er derfor at fordi jeg har begrenset problemstillingen til å dreie seg rundt kontrollvariablene utdanning, ferdigheter og helse, og dermed utelatt forhold som arbeidserfaring og yrkesretning, så kan det antas at b-koeffisientene blir feil estimert. Med hensyn til denne oppgavens omfang og plassbegrensning, må man imidlertid akseptere at ikke alle faktorer som kan påvirke innvandrernes lønn kunne inkluderes.

Helt avslutningsvis i dette delkapittelet ønsker jeg å si noe om hvilke andre forhold som videre forskning på feltet med fordel kan studere. Her vil jeg først og fremst trekke frem at variabler som sier noe om arbeidserfaring vil kunne være med på å forklare de observerte lønnsforskjellene. I tillegg vil en variabel som skiller mellom hvilket yrke folk jobber i, kunne vise om en større andel av innvandrere jobber i lavinntektsyrker. Dette igjen kan ha sammenheng med om innvandrerne har tatt utdanningen sin i Norge eller utlandet, som min økonometriske modell heller ikke tok hensyn til. Kort oppsummert så er det mange forhold som er med på å forklare lønnsforskjeller mellom innvandrere og majoritetsbefolkningen, og som med fordel kan studeres nærmere. Med hensyn til både plass- og tidsbegrensning så fikk jeg ikke studert alle forhold så nøye som jeg kunne tenke meg. På dette feltet virker det som om at jo mer man vet og begir seg ut på, desto flere nye problemstillinger dukker opp.

6. Konklusjon

Denne oppgaven kan konkludere med at det er betydelige inntektsforskjeller mellom innvandrere og majoritetsbefolkningen i både Norge og Italia, en effekt som synes sterkere i Italia. Lønnsgapet dempes når man kontrollerer for utdanning og ferdigheter, noe som vitner om at generell humankapital har god effekt på lønnsnivået. Dette stemmer godt overens med tidligere litteratur på feltet, som argumenterte for at ferdigheter og utdanning jevner ut lønnsforskjeller. Landspesifikk humankapital som nevnt i teorikapittelet, men som jeg ikke direkte undersøkte i den økonometriske modellen, kan også antas å påvirke lønnsnivået. Helsestatus på sin side kunne imidlertid ikke brukes til å forklare lønnsforskjeller.

En sammenlikning av Norge og Italia kunne avsløre at lønnsgapet er større i Italia enn i Norge, og det ble argumentert for at dette skyldes at Norge med god økonomisk vekst over lang tid, har bedre forutsetninger for å sikre en god integrering av innvandrerne i utdanningsinstitusjoner, så vel som i arbeidsmarkedet. Dette er i stor grad med på å dempe lønnsgapet, selv om man også i Norge observerer betydelige lønnsforskjeller.

Litteraturliste

Ambrosini, Maurizio (2013). Immigration in Italy: Between Economic Acceptance and Political Rejection. *Journal of International Migration and Integration*, 14(1), 175-194. Hentet fra <https://link.springer.com/article/10.1007/s12134-011-0231-3>

Bamberger, Michael (2000). *Integrating Quantitative and Qualitative Research in Development Projects* (The World Bank 6/2000). Hentet fra <http://elibrary.worldbank.org/doi/pdf/10.1596/0-8213-4431-5#page=13>

Bjørkeng, Birgit (2013). *Ferdigheter i voksenbefolkningen*. Statistisk sentralbyrå (42). Hentet fra https://www.researchgate.net/profile/Bengt_Lagerstrom/publication/273998965_Ferdigheter_i_voksenbefolkningen_Resultater_fra_den_internasjonale_undersoekslen_om lese-og_tallforstaelse_PIAAC/links/5511b8c40cf270fd7e316d99/Ferdigheter-i-voksenbefolkningen-Resultater-fra-den-internasjonale-undersoekslen-om-lese-og-tallforstaelse-PIAAC.pdf

Feenstra, Robert C., Inklaar, Robert and Timmer, Marcel P. (2015). The Next Generation of the Penn World Table. *American Economic Review*, 105(10), 3150-3182. Hentet fra www.ggdc.net/pwt

Raaum, Oddbjørn (1999). Inntektseffekter av utdanning i Norge – en litteraturoversikt. *Frischsenteret*, 1999(2), 1-57. Hentet fra https://www.frisch.uio.no/publikasjoner/pdf/arbnot99_02.pdf

Ringdal, Kristen (2013). *Enhet og mangfold*. Fagbokforlaget: Bergen.

Smidt, Stephen C. & Todaro, Michael P. (2015). *Economic Development* (utgave 12). Pearson Education: United Kingdom.

Thomas, R. L. (2005). *Using Statistics in Economics*. McGraw-Hill: United Kingdom

Wiborg, Øyvind (2006). Suksess i arbeidsmarkedet blant høyt utdannede innvandrere. *Sosiologisk Tidsskrift*, 14(3), 276-297. Hentet fra <https://www.duo.uio.no/bitstream/handle/10852/12905/1/wiborg2006sositidsskriftA5.pdf>