

# Lønnsforskjeller mellom innvandrere og majoritetsbefolkning i Norge

Bacheloroppgave i Samfunnsøkonomi

Veileder: Bjarne Strøm

Mai 2019

**NTNU**  
Norges teknisk-naturvitenskapelige universitet  
Fakultet for økonomi  
Institutt for samfunnsøkonomi



# Lønnsforskjeller mellom innvandrere og majoritetsbefolkning i Norge

Bacheloroppgave i Samfunnsøkonomi  
Veileder: Bjarne Strøm  
Mai 2019

Norges teknisk-naturvitenskapelige universitet  
Fakultet for økonomi  
Institutt for samfunnsøkonomi





## Innhold

|                                                    |    |
|----------------------------------------------------|----|
| 1. Innledning .....                                | 1  |
| 1.1 Problemstilling .....                          | 1  |
| 2. Teori .....                                     | 2  |
| 2.1 Innledning.....                                | 2  |
| 2.2 Teoretisk rammeverk.....                       | 2  |
| 2.3 Tidligere litteratur.....                      | 3  |
| 3. Metode.....                                     | 5  |
| 3.1 Innledning.....                                | 5  |
| 3.2 Datamaterialet, metoder og forutsetninger..... | 5  |
| 3.3 Definisjon av variablene:.....                 | 5  |
| 3.4 Utledning av modeller .....                    | 6  |
| 3.5 Valg av funksjonsform .....                    | 9  |
| 4. Deskriptiv statistikk.....                      | 10 |
| 5. Resultater .....                                | 14 |
| 6. Konklusjon .....                                | 20 |
| 7. Appendiks .....                                 | 22 |
| 8. Litteraturliste .....                           | 23 |

# 1. Innledning

Innvandring og integrering er et yndet tema i samfunnsdebatten i Norge. Høyresiden i politikken er for en strengere innvandringspolitikk enn venstresiden. Siden Erna Solbergs regjering tok over i 2013, har innvandringen i Norge gått ned fra 75 789 personer i 2013 til 52 485 personer i 2018.<sup>1</sup> Selv om innvandringspolitikken har blitt strengere de siste årene, er det fortsatt mange innvandrere som kommer til Norge på grunn av konflikt og krig i deres hjemland. 5. mars 2019 var det registrert 422 839 sysselsatte innvandrere i Norge.<sup>2</sup> Dette gjør at innvandrere utgjør en viktig del av arbeidsstyrken i Norge. Likevel har innvandrere langt lavere yrkesdeltakelse enn den øvrige befolkningen. Dette gjelder i størst grad flyktninger fra ikke-vestlige land.<sup>3</sup>

Det er stor konsensus om at integrering av innvandrere er viktig. I den sammenheng er det interessant å se på hvor mye innvandrerne tjener i forhold til majoritetsbefolkningen. Norge er et av verdens mest utviklede land, men er vi gode på integrering av innvandrere? Innvandrere er en sammensatt gruppe mennesker som kommer til Norge med helt forskjellige forutsetninger og mål. Disse skal så ta del i arbeidsmarkedet, med ulike språk og kompetansenivå.

## 1.1 Problemstilling

Vi har valgt følgende problemstillinger for denne oppgaven:

- Hovedproblemstilling: Tjener innvandrere mindre i gjennomsnitt enn majoritetsbefolkningen i Norge?
- I hvilken grad kan den eventuelle forskjellen i lønnsgapet forklares av leseferdigheter, arbeidserfaring, antall år utdanning og bransje?
- Tilleggsspørsmål:
  - Har innvandrere lik avkastning på utdanning som majoritetsbefolkningen?

---

<sup>1</sup> <https://www.ssb.no/innvandring-og-innvandrere/faktaside/innvandring>

<sup>2</sup> <https://www.ssb.no/arbeid-og-lonn/statistikker/innvregsys>

<sup>3</sup> <https://www.nho.no/tema/arbeidsliv/artikler/innvandring-og-integrering/>

## 2. Teori

### 2.1 Innledning

I dette kapittelet skal vi se på hva som kan forklare lønnsforskjellen mellom innvandrere og majoritetsbefolkningen.

### 2.2 Teoretisk rammeverk

Lønnsforskjeller blant innvandrere og i majoritetsbefolkning kan ha mange forklaringsfaktorer. I vår oppgave ser vi på følgende:

- Språkforståelse:

Dette er viktig både for å kunne få første jobb og for å kunne jobbe seg opp til bedre stillinger med bedre lønn. Språkferdigheter kan derfor være med å forklare lønnsnivå.

- Erfaring:

Det er naturlig å anta at med mer jobberfaring, vil man tjene mer på grunn av ansiennitet og kompetanse.

- Utdanningsnivå:

Grad av utdanning kan være med å forklare hvorfor noen tjener mer enn andre.

- Bransjer:

Mellom ulike bransjer er det store lønnsforskjeller. Det at innvandrere dominerer enkelte lavtlønnede bransjer, kan være en av årsakene til at innvandrere tjener dårligere enn majoritetsbefolkningen.

Andre forklaringsfaktorer:

Diskriminering:

I vårt datasett er det ikke tatt høyde for diskriminering. Selv om Norge er et utviklet land, har forskning vist at diskriminering kan bidra til at det er vanskeligere for Ali enn Ola Nordmann å få et jobbintervju<sup>4</sup>.

Kjønn:

Om man er kvinne eller mann kan være med å påvirke lønnen man får for gitt arbeid. Det er også store ujevnheter når det gjelder hvilke sektorer menn og kvinner jobber i. I tillegg er det flere kvinner som har deltidsstillinger<sup>5</sup>. Dette er naturligvis faktorer som kan forklare lønnsforskjeller. I denne oppgaven vil vi ikke se på kjønn som forklaringsfaktor.

## 2.3 Tidligere litteratur

I 2019 er det per dags dato 765 108 innvandrere i Norge. Ifølge lønnsstatistikk for innvandrere laget av SSB, utgjør lønn for innvandrere 86 % av lønn for majoritetsbefolkningen i Norge. Den gjennomsnittlige månedslønnen for majoritetsbefolkningen var 44 200 kroner i september 2016, mens innvandrere hadde en gjennomsnittlig månedslønn på 38 000 kroner.<sup>6</sup> SSB definerer innvandrere som personer som er født i utlandet av to utenlandske foreldre og fire utenlandske besteforeldre.<sup>7</sup> I lønnsstatistikken tallene i denne oppgaven er hentet fra, er det i overkant av 400 000 arbeidstakere med innvandrerbakgrunn.

Det er mange faktorer som påvirker arbeidstakeres lønn. Lønnsforskjeller i samme yrke kan skyldes grad av erfaring, hvor mye ansvar man har og hvilke arbeidsoppgaver man har. Hvor lenge en innvandrer har vært bosatt i Norge kan være avgjørende for lønnsnivået. De som har bodd relativt lenge i Norge, er ofte bedre integrert i samfunnet og behersker språket bedre. Hvor

---

<sup>4</sup> [https://www.idunn.no/st/2014/04/diskriminering\\_i\\_arbeidslivet\\_-\\_resultater\\_fra\\_randomiserte](https://www.idunn.no/st/2014/04/diskriminering_i_arbeidslivet_-_resultater_fra_randomiserte)

<sup>5</sup> <https://www.ssb.no/befolkning/artikler-og-publikasjoner/store-kjonnforskjeller-i-arbeidslivet>

<sup>6</sup> <https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/lavere-lonn-blant-innvandrere>

<sup>7</sup> <https://www.ssb.no/befolkning/artikler-og-publikasjoner/slik-definerer-ssb-innvandrere>



godt man er integrert i samfunnet og hvilke språkkunnskaper man har kan påvirke arbeidsoppgaver, og dermed ha betydning for lønnsnivået.<sup>8</sup>

De vanligste årsakene til innvandring er flukt, familiegjening og arbeidsinnvandring. Flyktninger har andre forutsetninger for å konkurrere i arbeidsmarkedet enn arbeidsinnvandrere. Arbeidsinnvandrere kommer stort sett fra Norden, Øst-Europa og Vest-Europa, og de kommer raskt i arbeid fordi de ofte besitter kompetanse som er etterspurt. Flyktninger kommer tradisjonelt fra Afrika og enkelte deler av Asia. Flyktninger besitter ulik grad av kompetanse og utdanningsnivå, og mange sliter med å komme i arbeid. Det å få disse i jobb er en viktig del av integreringsprosessen.<sup>9</sup>

En av de viktigste forutsetningene for å få jobb i Norge, er utdanning. Dette gjelder uavhengig av om man er innvandrer eller ikke. Dermed har de med lavt utdanningsnivå færre jobbmuligheter i arbeidsmarkedet. Høyere utdanning gir som regel tilgang til bedre betalte jobber. Innvandrere fra Norden og Vest-Europa har relativt likt utdanningsnivå som majoritetsbefolkningen, derfor har de mange av de samme jobbene. Innvandrere fra Øst-Europa har ofte mer yrkesfaglig utdanning. Mange av disse jobber derfor som håndverkere og renholdere. Afrikanere og asiater har ofte lav utdanning, og mange av disse har ikke fullført grunnskole. Derfor er de ofte sysselsatt i yrker der det kreves lite formell kompetanse, her er lønnen naturlig nok lav.<sup>10</sup>

Oppsummert er noen av de viktigste årsakene til lønnsforskjeller mellom innvandrere og majoritetsbefolkningen i Norge språkferdigheter, antall års arbeidserfaring, utdanningsnivå og bransje.

---

<sup>8</sup> <https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/lavere-lonn-blant-innvandrere>

<sup>9</sup> <https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/lavere-lonn-blant-innvandrere>

<sup>10</sup> <https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/lavere-lonn-blant-innvandrere>

## 3. Metode

### 3.1 Innledning

I denne delen skal vi først redegjøre for datamaterialet som er brukt i oppgaven. Videre vil vi etablere og forklare relevante metoder, forutsetninger, begreper og modeller. Til slutt skal vi velge funksjonsform og definere variabler.

### 3.2 Datamaterialet, metoder og forutsetninger

#### **Datamaterialet**

For å angripe problemstillingen bruker vi datasettet «data2901-Norway.dta», som er en del av PIAAC-undersøkelsen. Datasettet gir blant annet informasjon om inntekten til majoritetsbefolkning og innvandrere. Det inneholder data fra en undersøkelse gjennomført mellom august 2011 og april 2012. I den norske delen av undersøkelsen var det med om lag 5000 personer i alderen 16-65 år. Dette er over 30 observasjoner og datasettet regnet derfor som stort. Ved store datasett sier sentralgrenseteoremet at utvalgsgjennomsnittet er tilnærmet normalfordelt uavhengig av fordelingen til populasjonen.<sup>11</sup>

### 3.3 Definisjon av variablene:

**Earnhr:** Timelønn i lokal valuta, her i norske kroner.

**Learnhr:** Variabelen earnhr på logaritmisk form.

**Migrant\_first:** Dummy for innvandrere. Innvandrere defineres her som at personen er født i utlandet og at minst en av foreldrene også er det. Verdien til variabelen gir tallet 1 hvis disse forutsetningene er oppfylt, og 0 hvis de ikke er det.

**Litscore1:** Er de «råe» testscoreresultatene standardisert, med gjennomsnitt 0 og standardavvik 1.

**Erfaring:** Antall år med betalt arbeid i løpet av livet.

**Erfaring2:** Erfaring på kvadrert form, har denne med fordi det er mulig at erfaring er ikke-lineær.

**Yrsqual:** Høyeste utdannelsesnivå oppnådd beregnet i antall år med utdanning.

---

<sup>11</sup> Thomas 2005: 126

**Isc01c:** Forskjellige bransjer med nummer 2-10, der 2 er referansegruppe for de andre.

2: Lovgivere og sjefer

3: Profesjonelle

4: Teknikere

5: Butikkmedarbeidere, kontormedarbeidere

6: Servicearbeidere

7: Jordbrukere og fiskere

8: Håndverkere

9: Fabrikkarbeidere og maskinarbeidere

10: Grunnleggende arbeid

## 3.4 Utledning av modeller

### Økonometri

I økonometrien bruker man data for å tallfeste økonomiske sammenhenger. Dette gjøres gjennom å formulere økonomiske modeller og estimere verdier på parametere. Videre angir man usikkerhetene på parameteranslagene og utfører hypotesetesting. Til slutt vurderes resultatene kritisk og treffer en konklusjon.<sup>12</sup>

### Enkel regresjon

I enkel lineær regresjonsanalyse ser man på sammenhengen mellom endogene Y og en eller flere eksogene variabler. Y er utfallsvariabel og avhengig som betyr at den blir forklart av de uavhengige variablene i modellen.

$$Y = \alpha + \beta X + \varepsilon \quad (1)$$

Her er  $\alpha$  og  $\beta X$  ukjente populasjonsparametere hvor  $\alpha$  er konstantleddet og  $\beta$  er stigningstallet.

Vi har antatt en lineær konsumfunksjon.

Vi kan videre uttrykke (1) for observasjon i

---

<sup>12</sup> Thomas 2005: 248

$$Y_i = \alpha_i + \beta X_i + \varepsilon_i, \text{ der } i = 1, 2, 3 \dots, n \quad (2)$$

Det kan være vanskelig eller umulig å innhente tilstrekkelig data for å beregne  $Y$  i (2). Derfor utledes en modell som gir en forventet verdi av  $Y$ .

$$E(Y) = \alpha + \beta X \quad (3)$$

I økonomiske sammenhenger vil det alltid være forstyrrende forhold som ikke fanges opp i modellen. Dette resulterer i en differanse mellom  $Y$  og  $E(Y)$ , som kalles et støyledd<sup>13</sup>.

$$\varepsilon = E(Y) - Y \quad (4)$$

Derfor vil også alle økonomiske sammenhenger alltid være stokastiske. Som nevnt er populasjonsparametere ukjent og man utarbeider en utvalgsregresjonslikning

$$\hat{Y}_i = a + bX_i + e_i, \text{ der } i = 1, 2, 3 \dots, n \quad (5)$$

$a$  og  $b$  representerer estimater for  $\alpha$  og  $\beta$ .  $\hat{Y}_i$  utgjør en estimert modell av  $Y$ . I våre analyser vil vi bruke OLS for å beregne  $Y$ , og lage en likning tilsvarende regresjonslikning (5). OLS handler om å predikere  $Y$  best mulig ved hjelp av en regresjonslikning.

### **Minste kvadraters metode/Ordinary least squares (OLS)**

Målet er å estimere lønnsforskjellen mellom majoritetsbefolkningen og innvandrere. Dette gjøres ved å bruke Minste kvadraters metode (OLS). Vi bruker et datasett til å estimere  $\alpha$  og  $\beta$ , de ukjente parameterne i utvalgsregresjonslikningen. Denne metoden er en estimeringsmetode og den klart mest kjente for å utarbeide en regresjonslinje<sup>14</sup>. Metoden kalkulerer sammenhengen mellom en eller flere forklaringsvariabler samt responsvariabel.

OLS finner en sammenheng mellom variablene som minimerer variansen. Det OLS hovedsakelig gjør er å lage en rett linje i et punktdiagram hvor punktene er spredd.

$$\text{Min} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n (Y_i - a - bX_i)^2 \quad (6)$$

---

<sup>13</sup> Thomas 2005: 261

<sup>14</sup> Thomas 2005: 266

Når man minimerer summen av kvadrert avvik mellom faktiske observasjoner og predikerte verdier, kommer man frem til disse formlene for den beste tilpassede regresjonslinje

$$a = \bar{Y} - b\bar{X} \quad (7)$$

$$b = (\sum y_i x_i) / (\sum x_i^2) \quad (8)$$

Hvor (8) er stigningstallet til utvalgsregresjonslinjen.

### **Determinasjonskoeffisienten, R-squared ( $R^2$ )**

$R^2$  viser hvor godt regresjonslinjen beskriver datasettet. Den beskriver hvor stor andel av variasjonen som forklares ved modellen.  $R^2$  er alltid et tall mellom 0 og 1, og jo nærmere 1 jo bedre føyning. Om den viser seg å være 1 er det 100% perfekt føyning og all variasjon vil la seg forklare.

Hvis,

$$\text{SST, total variasjon} = \sum (Y_i - \bar{Y})^2,$$

$$\text{SSR, restvariasjon(residualvariasjon)} = \sum (Y_i - \hat{Y}_i)^2 = \sum (e_i)^2$$

$$\text{SSE, forklart variasjon} = \sum (\hat{Y}_i - \bar{Y})^2$$

$$R^2 = \text{SSE} / \text{SST} \quad (9)$$

(9) er forklart variasjon ift. total variasjon.

### **Hypotesetesting**

Når man har et utvalg og ser på om populasjonsparameterne er lik eller forskjellig fra en bestemt verdi, snakker man om en statistisk hypotesetest. Det er vanlig å ta utgangspunkt i en nullhypotese,  $H_0$ , og en alternativhypotese,  $H_a$ . Eksempel på en hypotese:

$$H_0: \mu_1 = \mu_2.$$

$$H_a: \mu_1 \neq \mu_2.$$

$H_a$  må dekke alle andre alternativer, som vist over.<sup>15</sup>

---

<sup>15</sup> Thomas 2005: 126

Det sier seg selv at  $H_0$  og  $H_a$  ikke samtidig kan være korrekt. Man kan enten forkaste  $H_0$  og dermed akseptere  $H_a$ , eller beholde  $H_0$ .

**F-test** er en statistisk test som har en F-distribusjon under en nullhypotese og er vanlig å bruke i regresjonsanalyser. F-testen blir oftest brukt når man sammenligner statistiske modeller som har blitt tilpasset et datasett. Formålet er å reflektere populasjonen som utvalget er hentet fra, best mulig. Vi kalkulerer F-testen i Stata. Testen indikerer om man burde støtte eller avvise nullhypotesen, som vist over. Det er et kriterium for å kunne gjennomføre F-testen, at populasjonen er antatt tilnærmet normalfordelt. I tillegg må utvalget være uavhengig.<sup>16</sup>

### 3.5 Valg av funksjonsform

Vi har valgt å bruke log-lineær funksjonsform, fordi vi tror at det vil være mest anvendelig for denne oppgaven. Denne funksjonsformen gjør oss i stand til å tolke lønnsforskjellen som tilnærmet prosentvist lønnsgap, uavhengig av målnivå på avhengig variabler.  $\beta$  representerer koeffisientvektoren til kontrollvariabelen. I likningen under er kontrollvariabelen `migrant_first` og  $\beta_1$  er den tilhørende koeffisientvektoren.<sup>17</sup>

$$\text{Learnhr} = \alpha + \beta_1 \text{migrant\_first} + \varepsilon$$

$100 * \beta_1$  = prosentvis forskjell i gjennomsnittslønn mellom innvandrere og majoritetsbefolkning. I denne modellen er `learnhr` den endogene variabelen, mens  $\alpha$  og  $\beta_1 \text{migrant\_first}$  er eksogene.  $\varepsilon$  er summen av restleddet og vil alltid være null ved OLS-estimering.

---

<sup>16</sup> <https://courses.lumenlearning.com/boundless-statistics/chapter/one-way-anova/>

<sup>17</sup> Thomas 2005: 415

## 4. Deskriptiv statistikk

| NORGE                    | Alle      | Innvandrere | Majoritetsbefolkning |
|--------------------------|-----------|-------------|----------------------|
| Gjennomsnittlig timelønn | 226,49 kr | 196,88 kr   | 230,68 kr            |
| Minimum                  | 52,31 kr  | 62,50 kr    | 52,31 kr             |
| Maksimum                 | 572,63 kr | 538,46 kr   | 572,63 kr            |
| Standardavvik            | 76,33     | 74,47       | 75,67                |
| Antall                   | 3023 stk. | 374 stk.    | 2649 stk.            |

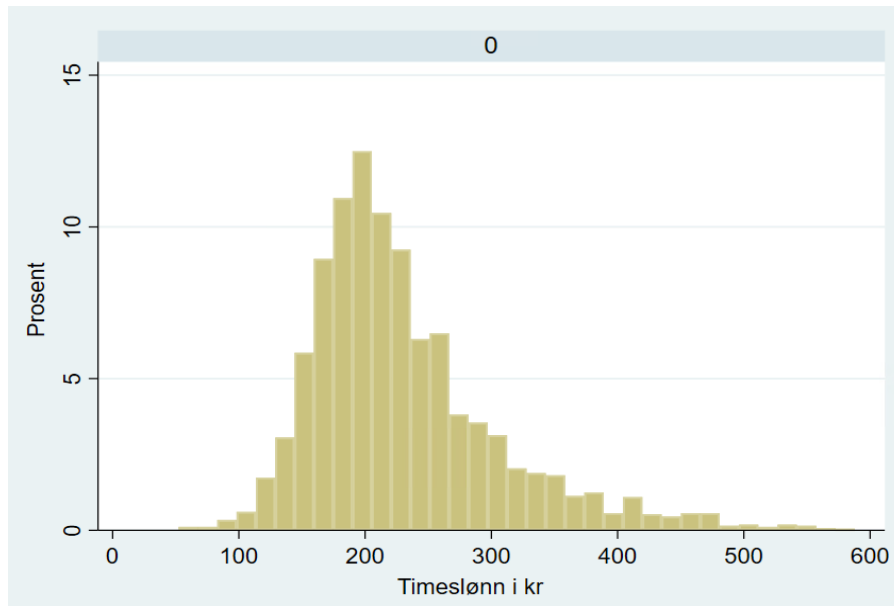
*Tabell 1: Deskriptiv statistikk uten kontrollvariabler*

### **Kommentar til tabellen:**

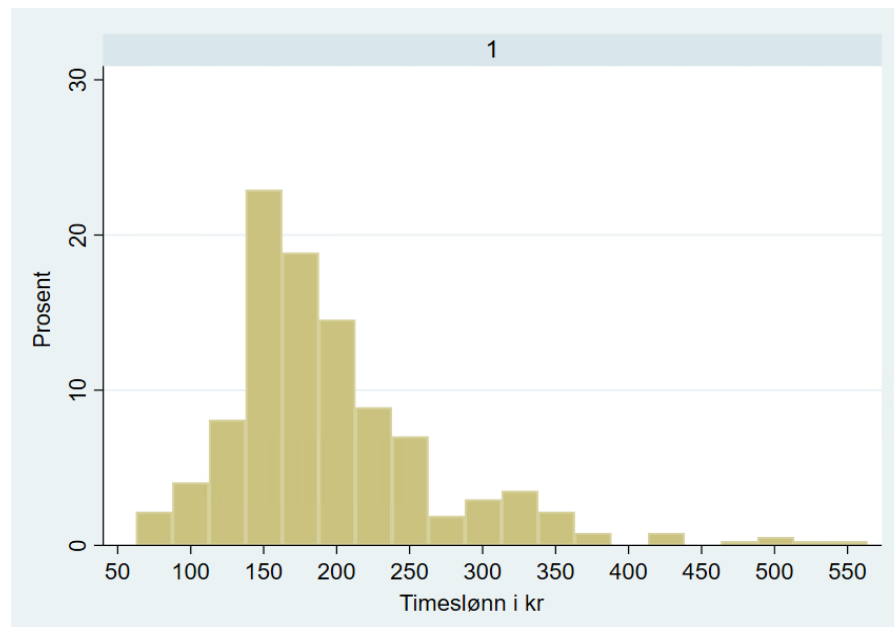
Tabell 1 viser at gjennomsnittlig timelønn for en norsk lønnstaker er 230,68 kr i timen.

Gjennomsnittlig timelønn for innvandrere er 85,37 % av den gjennomsnittlige timelønnen til majoritetsbefolkningen og ligger på 196,88 kr. Tabellen viser også et litt større lønnsavvik (standardavvik) for majoritetsbefolkning enn for innvandrene.

Figur 1: Lønnsfordeling majoritetsbefolkning



Figur 2: Lønnsfordeling innvandrere



**Kommentar til figurene:**

Figur 1 og 2 viser at typetallet for timelønn i majoritetsbefolkningen er omtrent 200 kr.

Tilsvarende for innvandrere er omtrent 150 kr.



|                                 | Alle         |                | Majoritetsbefolkning |                | Innvandrere  |                |
|---------------------------------|--------------|----------------|----------------------|----------------|--------------|----------------|
|                                 | Gjennomsnitt | Standard-avvik | Gjennomsnitt         | Standard-avvik | Gjennomsnitt | Standard-avvik |
| Utdanning                       | 14,66        | 2,45           | 14,65                | 2,34           | 14,74        | 3,07           |
| Språkkunnskaper /leseforståelse | 0,0280       | 1,01           | 0,1398               | 0,90           | -0,7245      | 1,36           |
| Bransje 3(profesjonelle)        | 250,45       | 74,86          | 247,46               | 74,58          | 248,47       | 85,24          |
| Bransje 4(teknikere)            | 242,77       | 75,82          | 240,73               | 76,42          | 221,90       | 72,45          |
| Bransje 6(Servicearbeidere)     | 179,64       | 55,15          | 166,17               | 59,18          | 159,73       | 44,97          |
| Bransje 10(Grunnleggende yrker) | 162,81       | 50,21          | 165,67               | 62,29          | 147,57       | 39,63          |
| År med erfaring                 | 21,69        |                | 22,54                |                | 15,79        |                |

Tabell 2: Deskriptiv statistikk med kontrollvariabler

### Kommentar til tabellen:

**Utdanning:** Vi ser av tabell 2 at gjennomsnittlig utdanning i Norge er 14,66 år. Vi ser at det er små forskjeller på gjennomsnittlig antall år utdanning mellom innvandrere og majoritetsbefolkning.

**Språkkunnskaper (litscore1):** Vi ser at innvandrere har en språkforståelse som er under gjennomsnittet med -0.7245, mens majoritetsbefolkningen ligger over gjennomsnittet med 0,1398.

**Bransjer:** Vi har valgt å se på fire ulike bransjer og lønnen i disse bransjene. Den første bransjen er profesjonelle. I denne bransjen kan vi se at innvandrere har høyere gjennomsnittslønn (248,47 kr) enn majoritetsbefolkningen (247,46 kr). Dette er den eneste bransjen hvor innvandrere har gjennomsnittslønn høyere enn majoritetsbefolkningen. I de tre andre bransjene vi ser på, ligger snittlønnen til majoritetsbefolkningen godt over lønnen til innvandrerne. De andre yrkene vi ser på, er teknikere, servicemedarbeidere og grunnleggende yrker. Her er gjennomsnittslønnen til majoritetsbefolkningen høyere enn innvandrernes gjennomsnittslønn.

**Erfaring:** Vi ser av tabellen at gjennomsnittlig antall år med erfaring er mye lavere for innvandrere enn for majoritetsbefolkningen. Majoritetsbefolkningen har gjennomsnittlig erfaring på 22,54 år, mens innvandrerne kun har 15,79 år erfaring i snitt. Her er forskjellen på majoritetsbefolkning og innvandrere relativt stor, dette kan være en grunn til lønnsforskjeller.

## 5. Resultater

$$(A): \text{learnhr} = \alpha + \beta_1 \text{migrant\_first}$$

$$(B): \text{learnhr} = \alpha + \beta_1 \text{migrant\_first} + \beta_2 \text{litscore1}$$

$$(C): \text{learnhr} = \alpha + \beta_1 \text{migrant\_first} + \beta_2 \text{litscore1} + \beta_3 \text{erfaring} + \beta_4 \text{erfaring2}$$

$$(D): \text{learnhr} = \alpha + \beta_1 \text{migrant\_first} + \beta_2 \text{litscore1} + \beta_3 \text{erfaring} + \beta_4 \text{erfaring2} + \beta_5 \text{yrsqual}$$

$$(E): \text{learnhr} = \alpha + \beta_1 \text{migrant\_first} + \beta_2 \text{litscore1} + \beta_3 \text{erfaring} + \beta_4 \text{erfaring2} + \beta_5 \text{yrsqual} + \beta_6 \text{isco1c}$$

|                          | A                     | B                     | C                     | D                     | E                      |
|--------------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|
| Variabler                | learnhr               | learnhr               | learnhr               | learnhr               | learnhr                |
| Migrant_first            | -0,17600<br>(0,01728) | -0,10658<br>(0,01737) | -0,04208<br>(0,01748) | -0,07957<br>(0,01669) | - 0,05472<br>(0,01715) |
| Kontroll leseferdigheter | Nei                   | Ja                    | Ja                    | Ja                    | Ja                     |
| Kontroll Erfaring        | Nei                   | Nei                   | Ja                    | Ja                    | Ja                     |
| Kontroll Utdanning       | Nei                   | Nei                   | Nei                   | Ja                    | Ja                     |
| Kontroll Bransje         | Nei                   | Nei                   | Nei                   | Nei                   | Ja                     |
| Observasjoner            | 3020                  | 3020                  | 3018                  | 3017                  | 2629                   |
| R <sup>2</sup>           | 0,0332                | 0,0969                | 0,1546                | 0,2410                | 0,3199                 |

Tabell 3: Estimert lønnsgap basert på modellene (A)-(E)

**Kommentar:** Tabell 3 viser at lønnsgapet varierer fra 17,6 % til 4,2 %. Gapet er størst i modellen som ikke inkluderer kontrollvariabler og minst i modellen som kontrollerer for litscore1 og erfaring.

### Forklaring og resultat basert på appendiks

Ifølge SSBs undersøkelse fra 16. november 2017 er de viktigste faktorene for innvandrernes gjennomsnittslønn lese- og språkferdigheter, erfaring, bransje og utdanning. Vi har derfor

bestemt oss for å bruke disse kontrollvariablene i vår regresjonsanalyse; litscore1, erfaring, isco1c og utdanning.

## **Regresjonsanalyse**

### **Tester først uten kontrollvariabler**

(A):  $\text{Learnhr} = 5,39289 - 0,17600 \cdot \text{migrant first}$  ( $R^2: 0,0332$ )

Likning A viser lønnsforskjellen mellom innvandrere og nordmenn på logaritmisk form, målt i prosentvis forskjell. På logaritmisk form er koeffisienten lik 5,393. Likning A forteller at innvandrere tjener 17,6 % mindre i snitt enn nordmenn. Dette er svaret på hovedproblemstillingen; Tjener innvandrere mindre i gjennomsnitt enn majoritetsbefolkningen i Norge? Så ønsker vi å teste om regresjonen er signifikant. Er den det, kan vi med sikkerhet si at innvandrere tjener mindre enn majoritetsbefolkningen i Norge. Vi tester dette ved hjelp av en F-test.

$H_0: \beta_1 = 0$

$H_a: \beta_1 \neq 0$ .

Vi velger 5 % signifikansnivå og får kritisk verdi 3,85 med 1 i teller og 3018 i nevner.

Testobservatoren på 103,73 overstiger kritisk verdi i F-fordelingen hvor vi har frihetsgradene (1, 3018). Vi forkaster dermed klart  $H_0$  og aksepterer  $H_a$ . Vi kan derfor si at det å være innvandrer, har statistisk signifikant påvirkning på lønnen.

$R^2$  i modellen er 0,0332. Det betyr at 3,32 % av variasjonen i predikert timelønn kan forklares med det å være innvandrer.

### **Kan lønnsgapet forklares med leseferdigheter?**

(B):  $\text{Learnhr} = 5,37327 - 0,10658 \cdot \text{migrant\_first} + 0,08709 \cdot \text{litscore1}$  ( $R^2: 0,0969$ )

I likning B er kontrollvariabelen litscore1 inkludert. Vi får alfakoeffisienten 5,372 som er nesten lik den i likning A. Vi ser at regresjonskoeffisienten til litscore1 er 0,08709. Det vil si at et

standardavviks økning i leseferdigheter vil øke lønnen med 8,7 %. Vi tester om variabelen litscore1 er signifikant ved hjelp av en F-test i Stata.

H0:  $\beta_2 = 0$

HA:  $\beta_2 \neq 0$

Vi får F-verdien 212.78 med frihetsgradene 1 og 3017. Kritisk verdi i A3 med 1 i teller og 1000 i nevner er 3,85. Vi ser at F-verdien overskrider kritisk verdi, og kan forkaste nullhypotesen med god margin. Dette betyr at litscore1 ikke er ubetydelig når det gjelder innvandreres lønn.

Når kontrollvariabelen litscore1 inkluderes, ser vi at forklaringskraften til modellen øker fra 0,0332 til 0,0969.

### **Kan lønnsgapet forklares med erfaring?**

(C):  $\text{Learnhr} = 5,10887 - 0,04208 * \text{migrant\_first} + 0,10401 * \text{litscore1} + 0,02039 * \text{erfaring} - 0,00031 * \text{erfaring}^2$  ( $R^2: 0,1546$ )

I likning C er kontrollvariablene litscore1, erfaring og erfaring2 inkludert. Leddet som er ganget med erfaring, er positivt, mens leddet som er ganget med erfaring kvadrert, er negativt. Dette vil si at erfaring ikke er lineært stigende.

Likning C ser slik ut:  $\text{learnhr} = \alpha + \beta_1 \text{litscore1} + \beta_2 * \text{erfaring} + \beta_3 * \text{erfaring}^2$  (erfaring2). For å finne effekten på lønnen av ett års mer erfaring, må vi derivere likningen med hensyn på erfaring.

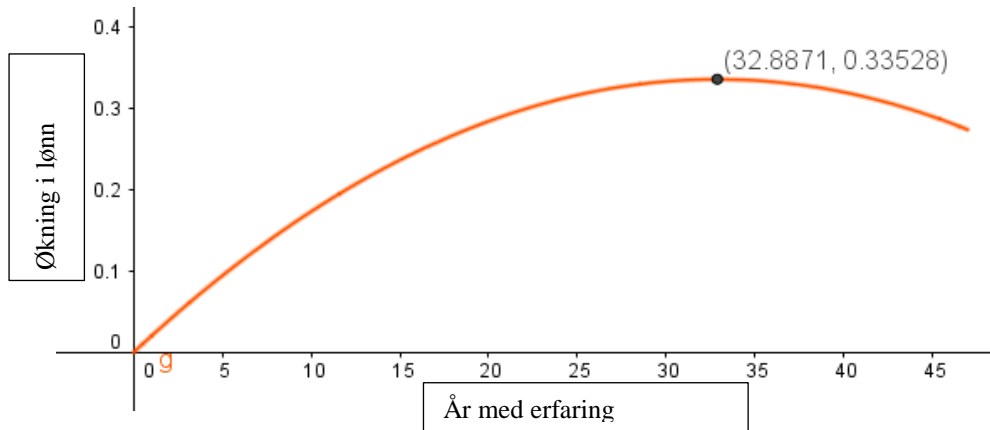
$$\frac{\delta \text{learnh}}{\delta \text{erfaring}} = \beta_2 - 2 * \beta_3 * \text{Erfaring}$$

$0,02039 - 2 * 0,00031 * x$  antall år erfaring = prosentvis endring i lønn med ett ekstra år erfaring.

Vi undersøker hvor margineffekten av erfaring er lik null ved å sette den deriverte lik null.

$0,02039 - 2 * 0,00031 * x = 0 \rightarrow 0,02039 - 0,00062 * x = 0 \rightarrow 0,02039 / 0,00062 = x \rightarrow x = 32,88 \approx 33$ .

Vi ser da, at lønnen øker for hvert år med erfaring inntil 33 år. Etter 33 år gir et ekstra år erfaring en negativ effekt på lønnen, som vist i figur 3.



Figur 3: Lønnsavkastning av erfaring

Vi tester med ti års erfaring:  $100*(0,02039 - 2*0,00031*10) = 1,419 \%$ . Med et utgangspunkt på ti års erfaring, tjener man 1,459 % mer ved en økning på ytterligere ett års erfaring.

Vi tester med 40 års erfaring:  $100*(0,02039 - 2*0,00031*40) = -0,441 \%$ . Med et utgangspunkt på 40 års erfaring, tjener man 0,441 % mindre ved en økning på ytterligere ett års erfaring.

Vi sjekker om erfaring og erfaring2 har signifikant effekt på lønnen ved hjelp av en F-test i Stata.

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

$H_a$ : minst  $\beta_3$  eller  $\beta_4$  er ikke lik 0

Vi får F-verdien 101,21 med 2 frihetsgrader i teller og 3013 i nevner. Kritisk verdi i tabell A3 med 2 frihetsgrader i teller og 1000 i nevner er 3. Vi forkaster dermed nullhypotesen klart, og kan si at erfaring er signifikant for lønnen.

Når erfaring og leseferdigheter inkluderes i regresjonen, ser vi at lønnseffekten av å være innvandrer går ned.  $R^2$  (0,1546) blir naturlig nok høyere når flere kontrollvariabler inkluderes.

### Kan lønnsgapet forklares av utdanning?

$$(D): \text{Learnhr} = 4,44350 - 0,07957*\text{migrant\_first} + 0,06040*\text{litscore1} + 0,02046*\text{erfaring} - 0,00030*\text{erfaring2} + 0,0444*\text{yrsqual} \quad (R^2: 0,2410)$$

I likning D er kontrollvariablene litscore1, erfaring, erfaring2 og yrsqual inkludert. Ett år ekstra utdanning gir ca. 4,44 % bedre lønn. Vi tar en F-test for å sjekke om utdanning har en signifikant effekt på timelønnen. Nullhypotese;  $H_0: \beta_5 = 0$ , mot alternativhypotesen  $H_A: \beta_5 \neq 0$ . Vi får F-

verdien 341,78 med 1 frihetsgrad i teller og 3017 i nevner. Kritisk verdi i tabell A3 med 1 frihetsgrad i teller og 1000 i nevner er 3,85. Dermed er det bekreftet at utdanning gir et signifikant utslag på timelønnen.

Vi ser at når vi har med enda en kontrollvariabel i regresjonen, reduseres effekten av å være innvandrere på lønnen. Forklaringskraften til modellen har blitt enda bedre, jfr.  $R^2$  som har gått opp fra 0,1546 til 0,2

### **Kan lønnsgapet forklares med bransje?**

(E):  $\text{Learnhr} = 4,87007 - 0,05472 * \text{migrant\_first} + 0,04400 * \text{litscore1} + 0,01885 * \text{erfaring} - 0,00028 * \text{erfaring2} + 0,02758 * \text{yrsqual} - 0,09587 * \text{isco1c3} - 0,08865 * \text{isco1c4} - 0,25282 * \text{isco1c5} - 0,31440 * \text{isco1c6} - 0,31440 * \text{isco1c7} - 0,13063 * \text{isco1c8} - 0,16434 * \text{isco1c9} - 0,32779 * \text{isco1c10}$   
( $R^2$ : 0,3199)

I likning E er kontrollvariablene litscore1, erfaring, erfaring2, yrsqual og isco1c inkludert. Vi ser at alle bransjetallene er negative i forhold til referansegruppen. Dette vil si at bransjegruppe 2 er den gruppen med høyest lønn. Vi kan for eksempel se på tallet til bransjegruppe 10, som er den bransjen med lavest gjennomsnittlig lønn. Tallet på -0,32779 forteller at en person i bransje 10 i gjennomsnitt tjener ca. 32,8 % mindre enn en i bransje 2. Disse tallene er ikke relevante i vår problemstilling, men det vi kan få ut av denne modellen, er forskjellen på koeffisienten til migrant\_first i likning D og E.

I likning D er regresjonskoeffisienten til migrant\_first -0,07957, mens i likning E er den på -0,05472. Dette betyr at bransje også er noe av forklaringen på hvorfor innvandrere tjener mindre enn majoritetsbefolkningen. Dette fordi relativt flere innvandrere enn nordmenn jobber i bransjer med et lavere lønnsnivå. Vi ser at i bransje 10 hvor det er lavest lønn, er det også høyest andel innvandrere, hele 46 %. Denne modellen gir den høyeste  $R^2$  (0,3199). av alle modellen fra A-E, og er dermed den med best forklaringskraft.

## Analyse av tilleggsspørsmål

### Har innvandrere lik avkastning på utdanning som majoritetsbefolkningen?

Vi lager her en ny modell med utdanning som avhengig variabel, istedenfor innvandring. Vi må lage en ny variabel, yrsqual\_migrant (yrsqual\*migrant\_first), som skal brukes i den nye regresjonen.

$$(F): \text{learnhr} = \alpha + \beta_1 \text{yrsqual} + \beta_2 \text{migrant\_first} + \beta_3 \text{yrsqual\_migrant}$$

$$(F): \text{learnhr} = 4,66526 + 0,04878 * \text{yrsqual} - 0,19143 * \text{migrant\_first} + 0,00047 \text{yrsqual\_migrant}$$

Majoritetsbefolkningens lønnsfunksjon kan skrives slik:  $\text{learnhr} = \alpha + \beta_1 \text{yrsqual}$ .

$\beta_1$  i regresjonen er på 0,04878. Den forteller oss at majoritetsbefolkningen tjener 4,878 % mer med ett års ekstra utdanning. Vi tester om koeffisienten er større enn 0 ved hjelp av en F-test.

$$H_0: \beta_1 = 0$$

$$H_a: \beta_1 \neq 0$$

Vi får F-verdien 372.30 med 1 frihetsgrad i teller og 3015 i nevner. Kritisk verdi i tabell A3 med 5 % signifikansnivå, 1 frihetsgrad i teller og 1000 i nevner gir kritisk verdi 3,8. Vi kan dermed forkaste nullhypotesen og bekrefte at majoritetsbefolkningen tjener mer ved ett års ekstra utdanning.

Innvandrerbefolkningens lønnsfunksjon kan skrives slik:  $\text{learnhr} = \alpha + \beta_2 + (\beta_1 + \beta_3) * \text{yrsqual}$ .

Her vil  $\alpha + \beta_2$  bestemme lønnen til en innvandrer, mens  $(\beta_1 + \beta_3) * \text{yrsqual}$  gir avkastning på ytterligere ett års utdanning for innvandrere. Vi finner altså en innvandrers alfakoeffisient ved å summere majoritetsbefolkningens alfakoeffisient med  $\beta_2$ .

Dette gir  $4,66526 + (-0,19143) = 4,47383$ , som er lønnen til en innvandrer uten noe utdanning, på logaritmisk form. Vi ser at denne verdien er lavere enn majoritetsbefolkningens lønn på logaritmisk form, som er på 4,66526. Regresjonsanalysen viser at betakoeffisienten til majoritetsbefolkningen er 0,04878. For å finne avkastningen på utdanning for en innvandrer, må vi summere denne med betakoeffisienten til yrsqual\_migrant. Dette gir  $0,04878 + 0,00047 = 0,04925$ .



Modellen vår viser altså at innvandrere har bedre avkastning på utdanning enn majoritetsbefolkningen. Innvandrere får i gjennomsnitt 4,925 % høyere lønn for hvert år ekstra utdanning de tar, mens majoritetsbefolkningen får 4,878 % mer. Vi tar deretter en hypotesetest for å sjekke om det er en signifikant forskjell på avkastningen.

$H_0: \beta_3 = 0$

$H_a: \beta_3 \neq 0$

Vi får F-verdien 0,01 med 1 frihetsgrad i teller og 3015 i nevner. Kritisk verdi med 1 frihetsgrad i teller og 1000 i nevner er 3,85. Vi beholder dermed nullhypotesen, og kan derfor ikke si at innvandrere og majoritetsbefolkning har forskjellig avkastning på utdanning.

## 6. Konklusjon

I denne oppgaven har vi ved hjelp av regresjonsanalyser og hypotesetesting forsøkt finne ut om lønnen til innvandrere er lavere enn majoritetsbefolkningens. Vi begynte med å gjengi tidligere forskning fra SSB som var relevant for vår problemstilling. Videre så vi på mulige forklaringsfaktorer gjennom teoretisk rammeverk. Gjennom metodedelen presenterte vi datamaterialet, definerte relevante variabler, utledet metoder, formulerte forutsetninger og valgte en passende funksjonsform. I kapittel fire «Deskriptiv statistikk» kom det frem at gjennomsnittlig timelønn for innvandrere er 85,37 % av timelønnen til majoritetsbefolkning. Dette stemte godt overens med tidligere forskning fra SSB, som viste at innvandrere i gjennomsnitt tjente 86 % av det majoritetsbefolkningen tjente.

Ved hjelp av en log-lineær funksjonsform har vi sett på hvordan lønnen påvirkes av forskjellige forklaringsfaktorer. De empiriske resultatene viste at innvandrere har en timelønn som er 17,6 % lavere enn majoritetsbefolkningens, når det ikke ble tatt hensyn til kontrollvariabler.

Det estimerte lønnsgapet varierte mellom 17,6 % og 4,2 % avhengig av hvilke kontrollvariabler det ble kontrollert for. Resultatene i vår undersøkelse viste at lønnsforskjellene kan forklares med leseferdigheter. Ett standardavviks økning i leseferdigheter vil øke lønnen med 8,7 %. Det kommer også frem at erfaring har en positiv effekt på lønnen i opptil ca. 33 år. Videre vil ett års ekstra utdanning øke lønnen med 4,44 %. Det ble vist at innvandrere oftere jobber i bransjer med

lavere lønn, noe som kan være med på å forklare lønnsdifferansen. Alle kontrollvariablene viste seg å være statistisk signifikante, og vi kan derfor konkludere med at innvandrere har lavere gjennomsnittlig timelønn enn majoritetsbefolkningen.

I tilleggsspørsmålet lurte vi på om innvandrere har lik avkastning på utdanning som majoritetsbefolkningen. Den deskriptive statistikken viste at innvandrere i gjennomsnitt har flere års utdanning enn majoritetsbefolkningen. Regresjonsanalysen viste at innvandrere i gjennomsnitt får 4,925 % høyere lønn for hvert år ekstra utdanning de tar, mens majoritetsbefolkningens lønn øker med 4,878 %. Dette viste seg imidlertid ikke å være statistisk signifikant, derfor kan vi ikke si at det er forskjellig avkastning på utdanning for innvandrere og majoritetsbefolkning.

Begrensninger i denne oppgaven, er først og fremst at vi hadde begrenset datamateriale. Det var for eksempel relativt få innvandrere i enkelte bransjer, noe som kan ha påvirket analysen. Videre kunne det vært interessant å se på diskriminering som en faktor for lønnsforskjell, men det er ikke inkludert i datasettet. Med et større og mer nyansert datasett med flere faktorer kunne resultatet blitt annerledes.

## 7. Appendiks

| Variabler                | lwage                 | lwage                 | lwage                 | lwage                 | lwage                 |
|--------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Migrant_first            | -0,17600<br>(0,01728) | -0,10658<br>(0,01737) | -0,04208<br>(0,01748) | -0,07957<br>(0,01669) | -0,05472<br>(0,01715) |
| Litscore1                |                       | 0,08709<br>(0,00597)  | 0,10401<br>(0,00606)  | 0,06040<br>(0,06212)  | 0,04400<br>(0,00654)  |
| erfaring                 |                       |                       | 0,02039<br>(0,00194)  | 0,02047<br>(0,00184)  | 0,01885<br>(0,00191)  |
| Erfaring 2               |                       |                       | -0,00031<br>(0,00004) | -0,00030<br>(0,00004) | -0,00028<br>(0,00004) |
| yrsqual                  |                       |                       |                       | 0,04435<br>(0,00240)  | 0,02758<br>(0,00287)  |
| Bransje 3                |                       |                       |                       |                       | -0,09587<br>(0,02031) |
| Bransje 4                |                       |                       |                       |                       | -0,08865<br>(0,02067) |
| Bransje 5                |                       |                       |                       |                       | -0,25282<br>(0,02715) |
| Bransje 6                |                       |                       |                       |                       | -0,30168<br>(0,02201) |
| Bransje 7                |                       |                       |                       |                       | -0,31440<br>(0,11996) |
| Bransje 8                |                       |                       |                       |                       | -0,13063<br>(0,02589) |
| Bransje 9                |                       |                       |                       |                       | -0,16434<br>(0,03108) |
| Bransje 10               |                       |                       |                       |                       | -0,32779<br>(0,03510) |
| Constant                 | 5,39289<br>(0,00606)  | 5,37327<br>(0,00601)  | 5,10887<br>(0,02083)  | 4,44530<br>(0,04095)  | 4,87007<br>(0,05240)  |
| F-test (1 og 3018 fr.gr) | 103,73                | 212,78                | 101,21                | 341,78                |                       |
| F-test (1 og 3017 fr.gr) |                       |                       |                       |                       |                       |
| F-test (2 og 3013 fr.gr) |                       |                       |                       |                       |                       |
| F-test (1 og 3017 fr.gr) |                       |                       |                       |                       |                       |
| Observasjoner            | 3020                  | 3020                  | 3018                  | 3017                  | 2629                  |
| R <sup>2</sup>           | 0,0332                | 0,0969                | 0,1546                | 0,2410                | 0,3199                |

## 8. Litteraturliste

Dzamarija, M,T. (2018) *Slik definerer SSB innvandrere*. Hentet fra:

<https://www.ssb.no/befolkning/artikler-og-publikasjoner/slik-definerer-ssb-innvandrere>

NHO (hentet 3 mars, 2019) *Innvandring og integrering*. Hentet fra:

<https://www.nho.no/tema/arbeidsliv/artikler/innvandring-og-integrering/>

Næsheim N,H. (2016) Flyktninger på arbeidsmarkedet. *Arbeid - mål og arena for integrering*.

Hentet fra

<https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/arbeid-mal-og-arena-for-integrering>

Selle, M. (2018). *Hypotesetesting*. Hentet fra

<https://wiki.math.ntnu.no/tma4245/tema/begreper/hypothesis>

SSB. (2017) *Lavere lønn blant innvandrere*. Hentet fra

<https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/lavere-lonn-blant-innvandrere>

SSB. (2019) *Fakta om innvandring*. Hentet fra

<https://www.ssb.no/innvandring-og-innvandrere/faktaside/innvandring>

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

Boundless-statitics. I Lumen learning. Hentet 4. Mai 2019 fra

<https://courses.lumenlearning.com/boundless-statistics/chapter/one-way-anova/>

