

Emilie Asplund Dyrseth  
Helene Melhus  
Jenny Vataker Seeberg

## Sammenhengen mellom utdanningsnivå og arbeidsledighet

En analyse av hvordan utdanningsnivå påvirker  
arbeidsledigheten i norske kommuner

Bacheloroppgave i Samfunnsøkonomi  
Veileder: Torberg Falch  
Mai 2023



Emilie Asplund Dyrseth  
Helene Melhus  
Jenny Vataker Seeberg

## **Sammenhengen mellom utdanningsnivå og arbeidsledighet**

En analyse av hvordan utdanningsnivå påvirker  
arbeidsledigheten i norske kommuner

Bacheloroppgave i Samfunnsøkonomi  
Veileder: Torberg Falch  
Mai 2023

Norges teknisk-naturvitenskapelige universitet  
Fakultet for økonomi  
Institutt for samfunnsøkonomi



**NTNU**

Kunnskap for en bedre verden



*Vi ønsker å rette en stor takk til vår veileder, Torberg Falch, for verdifulle innspill, nyttige tilbakemeldinger og rask respons.*

## SAMMENDRAG

I denne bacheloroppgaven har vi undersøkt sammenhengen mellom utdanningsnivå og arbeidsledighet. Analysen baserer seg på minste kvadraters metode (OLS) som økonometrisk metode. Data for den avhengige variabelen, arbeidsledige, er hentet fra NAV, mens variablene for utdanningsnivå er hentet fra SSB. Videre vil vi utvide modellen med flere kontrollvariabler, også hentet fra SSB, for å kunne besvare problemstillingen. Oppgavens konklusjon er at utdanning er en viktig faktor for å redusere arbeidsledigheten i de norske kommunene, spesielt høyere utdanning.

## ABSTRACT

In this bachelor's thesis, we have examined the relationship between level of education and unemployment. The analysis is based on ordinary least squares regression (OLS) as an econometric approach. Data for the dependent variable, unemployment, has been obtained from NAV, while the variables for educational attainment are from SSB. Furthermore, we will expand the model by including additional control variables, also obtained from SSB, to address the research question. The conclusion of the thesis is that education is an important factor in reducing unemployment in Norwegian municipalities, especially higher education.

# INNHALDSFORTEGNELSE

<b>1 INNLEDNING</b> .....	<b>1</b>
<b>2 ØKONOMETRISK METODE</b> .....	<b>3</b>
2.1 REGRESJONSANALYSE .....	3
2.2 MINSTE KVADRATERS METODE (OLS) .....	4
2.3 FORKLART VARIANS (R-SQUARED).....	5
2.4 HYPOTESETESTING .....	5
2.5 T-TEST OG P-VERDI.....	6
2.6 KONFIDENSINTERVALL .....	6
<b>3 BEGREPSAVKLARING OG TIDLIGERE LITTERATUR</b> .....	<b>7</b>
3.1 BEGREPSAVKLARING .....	7
3.2 TIDLIGERE LITTERATUR .....	7
<b>4 DATAMATERIALET</b> .....	<b>10</b>
4.1 VARIABLER OG HYPOTESER .....	11
<b>5 EMPIRISK ANALYSE</b> .....	<b>15</b>
5.1 DESKRIPTIV STATISTIKK .....	15
5.2 MINSTE KVADRATERS METODE (OLS) .....	17
5.3 HYPOTESETESTING .....	20
5.4 DRØFTING .....	23
5.5 ANTAGELSENE.....	28
5.6 MULIGE SVAKHETER VED OPPGAVEN .....	30
<b>6 KONKLUSJON</b> .....	<b>32</b>
<b>7 REFERANSELISTE</b> .....	<b>33</b>

## INNHALDSFORTEGNELSE FOR TABELLER OG FIGURER

<b>Tabell 1:</b> Deskriptiv statistikk for utdanningsnivå, avrundet til to desimaler, målt i prosent .....	15
<b>Tabell 2:</b> Deskriptiv statistikk for kontrollvariabler, avrundet til to desimaler, målt i prosent .....	16
<b>Tabell 3:</b> OLS i prosent for modell 1 og modell 2, standardfeil i parentes.....	18
<b>Figur 1:</b> Scatter plot modell 1 .....	29
<b>Figur 2:</b> Scatter plot modell 2 .....	29
<b>Figur 3:</b> Normalfordelingskurve modell 1 .....	29
<b>Figur 4:</b> Normalfordelingskurve modell 2 .....	29

# 1 INNLEDNING

Vi har i denne oppgaven valgt å se nærmere på hva som påvirker arbeidsledigheten i Norge. Arbeidsledighet defineres av Holden som: «de som er uten betalt arbeid, men som aktivt forsøker å få arbeid» (Holden, 2016, s. 88). Sammenliknet med andre land har Norge en svært lav arbeidsledighetsprosent og høy sysselsettingsprosent (Holden 2016). Arbeidsledigheten varierer likevel mye i de ulike delene av landet, hvor Sogn og Fjordane har 1,57 prosent arbeidsledighet, mens Rogaland har 3,90 prosent arbeidsledighet (NAV, 2018). Vi synes derfor det er interessant å gå nærmere inn på hvilke faktorer som kan påvirke arbeidsledigheten i ulike geografiske områder. Ettersom dette er et bredt og komplekst tema, har vi valgt å begrense vår oppgave til kun å omhandle arbeidsledighet på kommunenivå. Vi vil i det følgende ta for oss noen faktorer vi tror påvirker arbeidsledigheten, samt se på hvordan dette utspiller seg forskjellig fra kommune til kommune. Grad av utdanning er tett knyttet til sysselsetting og arbeidsledighet, og manglende utdanning blir ansett som den største risikofaktoren til å forklare hvorfor noen faller utenfor arbeidslivet (OECD, 2018). Høyere utdanning er likevel ingen garanti for deltakelse i arbeidslivet (Guldbrandsen, 2022). På bakgrunn av dette ønsker vi med denne oppgaven å identifisere om, og eventuelt i hvilken grad, utdanningsnivå har noe å si for arbeidsledigheten i de ulike kommunene. Vi har dermed valgt følgende problemstilling:

*Hvordan bidrar utdanningsnivå til å skape forskjeller i arbeidsledigheten i norske kommuner?*

For å besvare problemstillingen har vi valgt å kontrollere for andre faktorer som kan bidra til å forklare forskjellen i arbeidsledigheten i de ulike kommunene. Vi har valgt å begrense disse til innvandrere, norskfødte med innvandrerforeldre, alder, virksomheter og innbyggertall. I denne oppgaven har vi innhentet data om variablene for de ulike kommunene slik at vi kan foreta en regresjonsanalyse. For å gjennomføre analysene av datasettet har vi benyttet oss av programvaren Jupyter Notebook. På bakgrunn av resultatene diskuterer vi hvordan utdanningsnivå, og de resterende faktorene vi har ansett som relevante, påvirker arbeidsledigheten.



For å få et så korrekt bilde av den generelle arbeidsledigheten som mulig ønsket vi å innhente tall fra det man kan anse som et «normalår». Det vil si at år som i stor grad har vært påvirket av koronapandemien, 2020 og 2021, ikke har vært aktuell, da dette ga uvanlig høye arbeidsledighetstall sammenliknet med tidligere år. Siden vi ønsket å anvende så nye data som mulig var 2022 i utgangspunktet ideelt for oss. Dette lot seg imidlertid ikke gjennomføre, da tall og statistikk for 2022 ikke var tilgjengelig ved innhenting av data. 2018 og 2019 var ikke påvirket av pandemi, men årene fra 2018 og frem til 2020 var preget av kommunereformen, hvor Norge gikk fra å ha 426 til 356 kommuner. På bakgrunn av dette har vi valgt å benytte data fra 2017. Kun Sandefjord, Andebu og Stokke ble sammenslått til nye Sandefjord kommune dette året. Ved å velge 2017 har det gitt oss mulighet til å få et større antall observasjoner, og dermed et mer nøyaktig resultat ved våre regresjonsanalyser.

## 2 ØKONOMETRISK METODE

### 2.1 REGRESJONSANALYSE

Regresjonsanalyse er en metode innenfor statistikk som brukes til å forklare sammenhengen mellom én eller flere uavhengige variabler,  $x$ , og en avhengig variabel,  $y$ . I multippel regresjon inkluderer man flere uavhengige variabler i regresjonsmodellen. Matematisk ser den generelle likningen for multippel regresjon slik ut:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + u$$

Her er  $y$  den avhengige variabelen, mens  $\beta_0$  representerer konstantleddet. Det innebærer at når verdien på de uavhengige variablene er null får man verdien på den avhengige variabelen. I mange tilfeller er det normalt at konstantleddet bare har en teknisk fortolkning. I multippel regresjon sier koeffisientene,  $\beta_n$ , noe om den forventede verdien til den avhengige variabelen dersom den aktuelle uavhengige variabelen øker med én enhet, alt annet konstant. Den sier dermed noe om hvilken retning linjen har, om den er negativ eller positiv, samt hvor bratt linjen er. Dersom helningskoeffisientene er lik null, vil vi ikke ha en trend. Gitt en bestemt verdi på de uavhengige variablene kan regresjonslikningen brukes for å predikere en forventet verdi på den avhengige variabelen, og i denne sammenhengen kalles dette prediksjon.  $u$  er feilleddet, og representerer andre faktorer enn de uavhengige variablene som påvirker den avhengige variabelen. Med andre ord inneholder feilleddet andre variabler som kunne vært inkludert i regresjonsmodellen for å forklare variasjonen i den avhengige variabelen, og korrelerer med en eller flere av de uavhengige variablene (Christoffersen, Johannesen og Tufte, 2020).

I multippel regresjon er det seks antagelser som må være oppfylt for at vi skal kunne få den beste lineære forventningsestimatorene. Det er imidlertid vanskelig å oppfylle alle antagelsene, men det er viktig å være klar over svakhetene ved egen analyse. På den måten er man klar over at resultatene fra den multiple regresjonen ikke nødvendigvis gir korrekte konklusjoner.

**MLR.1:** Første antagelse for multippel regresjon er linearitet, som vil si at det er en lineær sammenheng mellom den avhengige variabelen og de uavhengige variablene.

**MLR.2:** Antagelse to innebærer at man har et tilfeldig utvalg som er representativ for populasjonen man ønsker å si noe om.

**MLR.3:** Den tredje antagelsen er at den lineære sammenhengen mellom to eller flere av de uavhengige variablene ikke er perfekt. Det er dette som kalles perfekt multikollinearitet.

**MLR.4:** Fjerde antagelse innebærer at restleddet har en forventet verdi lik null, gitt alle verdier på den uavhengige variabelen. Det er ofte utfordrende å kunne si noe om hvorvidt denne antagelsen er oppfylt eller ikke. Dette skyldes at man ikke kjenner til den forventede verdien på feilleddet, og det kan heller ikke testes.

**MLR.5:** Den femte antagelsen kalles homoskedastisitet, og innebærer at feilleddene rundt regresjonslinjen har konstant variasjon. Heteroskedastisitet betyr derimot at residualene har et slags mønster rundt regresjonslinjen, noe man ikke ønsker. For eksempel at desto høyere verdi det er på den uavhengige variabelen, desto større spredningen er det rundt regresjonslinjen.

**MLR.6:** Den sjette og siste forutsetningen handler om at feilleddet skal være normalfordelt med gjennomsnitt lik null, som vil være tilfelle dersom MLR.4 er oppfylt, og konstant varians, som vil si at MLR.5 er oppfylt. Normalfordeling er en funksjon som forklarer den tilfeldige fordelingen av en variabel, hvor de fleste verdiene vil samle seg rundt den sentrale delen av kurven, som derav kalles den sentrale tendensen. I tillegg skal restleddet,  $u$ , være uavhengig av forklaringsvariablene (Wooldridge, 2012).

## 2.2 MINSTE KVADRATERS METODE (OLS)

Avvikene fra regresjonslinjen kalles residualer, og er den predikerte verdien på den avhengige variabelen minus den observerte verdien. Minste kvadraters metode er en metode som brukes for å komme frem til en linje som totalt sett har de minste residualene (Christoffersen, Johannesen og Tufte, 2020). Man bruker estimering for å kunne få et anslag av den ukjente størrelsen, hvor dette kalles et estimat og betegnes med en «hatt». Metoden brukes dermed til å estimere verdien på konstantleddet og helningskoeffisientene. Matematisk skrives OLS regresjonslinjen for multipl regressjon med de predikerte verdiene slik (Wooldridge, 2012):

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_n x_n$$

### 2.3 FORKLART VARIANS (R-SQUARED)

Forklart varians (R-squared) sier noe om hvor mye av variasjonen i den avhengige variabelen som forklares av de uavhengige variablene. Denne verdien varierer mellom 0 og 1, hvor verdien 0 forteller at de uavhengige variablene ikke forklarer noe av variasjonen i den avhengige variabelen. Det vil si at det ikke eksisterer en lineær sammenheng mellom variablene. Dersom verdien er 1 forklarer de uavhengige variabelen all variasjon i den avhengige variabelen. Da er alle de observerte verdiene på regresjonslinjen, og det er dermed en perfekt lineær sammenheng. En lav forklart varians kan være en indikasjon på at det er andre forklaringsvariabler som burde vært inkludert i modellen. Hvis formålet er å se hvor bra regresjonsmodellen forklarer variasjonen i den avhengige variabelen er det hensiktsmessig å benytte den ujusterte R-squared. Den justerte R-squared bør derimot benyttes dersom vi ønsker å sammenlikne en regresjonsmodell med en annen regresjonsmodell hvor vi har tillagt flere variabler. Årsaken til dette er at den ujusterte R-squared øker dersom vi inkluderer flere uavhengige variabler. Den justerte R-squared derimot kan reduseres dersom vi inkluderer variabler i regresjonsmodellen som ikke er med på å forklare variasjonen i den avhengige variabelen (Christoffersen, Johannesen og Tufte, 2020).

### 2.4 HYPOTESETESTING

Ved hypotesetesting setter man gjerne opp to hypoteser, hvor man ønsker å teste hvorvidt regresjonsmodellen bidrar eller ikke bidrar til å forklare variasjonen i den avhengige variabelen i populasjonen. De to hypotesene kan formuleres slik:

- Nullhypotesen: Det er ingen sammenheng mellom den avhengige variabelen,  $y$ , og den uavhengige variabelen,  $x$ . Dette betyr at selv om koeffisienten viser en negativ sammenheng skyldes dette tilfeldig variasjon.
- Den alternative hypotesen: Det er en sammenheng mellom den avhengige variabelen,  $y$ , og den uavhengige variabelen,  $x$ . Med andre ord vil verdien vi forventer på den avhengige variabelen variere med verdien på den uavhengige variabelen.

Gjennom ulike statistiske tester vil man enten forkaste eller beholde nullhypotesen. Vi skal videre presentere to mye brukte metoder innenfor statistikk som vi skal bruke i vår oppgave for å sjekke om vi kan påstå våre alternative hypoteser (Christoffersen, Johannesen og Tufte, 2020).

## 2.5 T-TEST OG P-VERDI

Når man utfører en t-test regner man ut en t-verdi ved å ta den estimerte verdien til helningskoeffisienten delt på standardfeilen til helningskoeffisienten. Videre sammenlignes t-verdien med en kritisk verdi fra en t-tabell ut ifra signifikansnivå og antall frihetsgrader. Dersom nullhypotesen er sann viser signifikansnivået hvor stor sannsynlighet det er for at man forkaster en nullhypotese som egentlig stemmer. Når man velger å forkaste en nullhypotese skyldes det at forskjellen mellom nullhypotesen og den estimerte koeffisienten antas å være så stor at det er svært liten sannsynlighet for at det ikke er en sammenheng mellom den avhengige variabelen og den uavhengige variabelen. Både t- og p-verdi vil gi de samme konklusjonene ved hypotesetesting. Det er vanlig å bruke p-verdi større eller mindre enn 0,05 som signifikansnivå, men p-verdien vil også si noe om når estimatet ikke lenger er signifikant, samt når nullhypotesen kan forkastes (Christoffersen, Johannesen og Tufte, 2020).

## 2.6 KONFIDENSINTERVALL

Konfidensintervall indikerer et estimert intervall av sannsynlige verdier for en gitt populasjon. Det gir dermed en nedre og øvre grense for estimerte verdier. Det er mest vanlig å bruke 95 prosent konfidensintervall, men dette kan i utgangspunktet velges fritt ut ifra hensikten med analysen. Ved et slikt konfidensintervall vil observasjonene i 95 prosent av tilfellene ligge innenfor det estimerte konfidensintervallet. I vårt tilfelle velger vi å bruke et 95 prosent konfidensintervall. Et relativt stort konfidensintervall kan tyde på at estimatene er usikre, mens et relativt lite konfidensintervall kan derimot tyde på mer sikre estimater (Christoffersen, Johannesen og Tufte, 2020).

## 3 BEGREPSAVKLARING OG TIDLIGERE LITTERATUR

### 3.1 BEGREPSAVKLARING

Arbeidsstyrken i et land defineres som alle som tilbyr arbeidskraft i et arbeidsmarked. Det vil innebære summen av sysselsatte og arbeidsledige. Arbeidsledighet er summen av likevektsledighet og konjunkturledighet. Innen likevektsledighet skiller man mellom friksjonsledighet og strukturledighet. Friksjonsledighet vil si den ledighet som oppstår når man bytter jobb, mens strukturledighet handler om at kompetansen som etterspørres i markedet ikke samsvarer med arbeidstakernes kompetanse og utdanning. Konjunkturledighet skyldes derimot at det er for lav etterspørsel i økonomien, og vil være differansen mellom likevektsledigheten og den faktiske ledigheten (Holden, 2016).

### 3.2 TIDLIGERE LITTERATUR

Vi har brukt boken «Labor Economics» av George J. Borjas, en moderne introduksjon til arbeidsøkonomi, for å ta for oss teamet humankapital. Ifølge Borjas er humankapital definert som «hva hver arbeidstaker kan tilby i arbeidsmarkedet av forskjellige ferdigheter og egenskaper» (Borjas, 2008, s. 235). Gjennom skole og jobberfaringer tilegner vi oss mesteparten av humankapitalen. I tillegg øker humankapitalen ved skolegang og arbeid, ettersom sosiale ferdigheter utvikles. Det innebærer at kompetansen man opparbeider seg gjennom skolegang legger et solid grunnlag for den kunnskapen man bygger videre gjennom arbeid senere i livet. Gjennom hele livet vil man fortsette å investere i, samt utvikle egen humankapital helt til marginalavkastningen på humankapitalen er lik marginalkostnaden av å investere i mer humankapital. Det innebærer at som ung vil avkastningen på humankapital være større ettersom man har mangel på tidligere erfaring og kunnskap. Marginalavkastningen på humankapital vil derimot falle etter hvert som man blir eldre. Den humankapital man opparbeider seg som ung kan anvendes over en lang periode, mens humankapital innhentet som eldre bare vil kunne benyttes for en kortere periode. Det betyr at desto tidligere humankapitalen er innhentet, desto mer anvendelig er den (Borjas, 2008).

Basert på en undersøkelse gjennomført i USA fra 2005, er utdanning strekt korrelert med yrkesdeltakelse og sysselsettingsrate, og utdanningsnivå ses derfor på som svært betydningsfullt. Høyere utdanning vil ikke bare øke en persons fremtidige lønn, men også bidra med en positiv signaliserende effekt overfor arbeidsgiver. Man viser ved høyere utdanning at man er i stand til å tilegne seg kunnskap, være produktiv og dermed generelt være kvalifisert for å gjøre et godt arbeid. Dette gir vedkommende et positivt inntrykk overfor arbeidsgiver, og gjør det enklere å få seg jobb (Borjas, 2008).

Ifølge en artikkel av With publisert i 2017 kommer det frem at både utdanning og arbeid kan gi tilgang på en rekke goder, både for enkeltindivider og for samfunnet som helhet. Lav utdanning og høy arbeidsledighet vil derimot kunne føre til lav tilfredshet og mening med livet, og man ser tydelig at personer med lav formell kompetanse er overrepresentert blant dem som er arbeidsledige. Totalt sett er 7 av 10 personer over 16 år enten i utdanning eller i arbeid. Det kommer tydelig frem at vi er i en utdanningseksplansjon, hvor unge i dag har langt høyere utdanning enn eldre. Det kommer også frem i artikkelen at blant unge under 45 år er over 90 prosent enten i arbeid eller utdanning, mens tallet synker blant eldre mellom 45 og 66 år, hvor prosenten er på 78 (With, 2017).

I henhold til en artikkel av Østby fra 2017 om norskfødte med innvandrerforeldre, kommer det tydelig frem at norskfødte med innvandrerforeldre presterer langt over deres foreldregenerasjon. Forutsetningene for arbeid og inntekt skapes gjennom utdanningssystemet, og utdanning er et viktig mål på integrering. Artikkelen viser til at utbyttet av grunnutdanningen, med tanke på resultater, er dårligere for norskfødte med innvandrerforeldre enn den øvrige befolkningen. Man ser imidlertid at ved høyere utdanning vil en større andel av norskfødte med innvandrerforeldre ta høyere utdanning sammenliknet med andelen av norskfødte uten innvandrerbakgrunn. Her ligger de mellom 7 og 8 prosentpoeng over de med norskfødte foreldre. Avkastningen av utdanning er totalt sett lavest for innvandrerne. Ser vi på aldersgruppen 25-29 år er 65 prosent av innvandrerne, 73 prosent av norskfødte med innvandrerforeldre og 78 prosent av alle i Norge

sysselsatt. Norskfødte med innvandrereforeldre ligger altså nære den øvrige befolkningen både med tanke på utdanning og arbeid, mens innvandrere ligger et lite stykke under (Østby, 2017).

En artikkel skrevet av Vidal-Gil fra 2017 omhandlende utviklingen av arbeidsledigheten over tid, trekker frem at både endringer i økonomien, innvandring og utdanning har stor innvirkning på antall ledige over tid. Konjunkturer, strukturelle endringer i økonomien samt demografiske endringer påvirker hvor mange og hvem som blir ledige over tid. På grunn av EU-utvidelsen til Øst-Europa i 2004, økte innvandringen markant som medførte at personer med innvandrerbakgrunn utgjør en økende andel av de ledige. Fra 2002 til 2014 økte arbeidsledige innvandrere fra 17,5 prosent til 40 prosent av de ledige, men holdt seg stabilt frem til slutten av 2016 (Vidal-Gil, 2017).

En analyse publisert i 2017 av Andreev og Schou tar for seg sannsynligheten for at arbeidsledige flytter. Alder, utdanningsnivå og bosted trekkes frem som de viktigste faktorene når det kommer til flyttevillighet. Unge er mest tilbøyelige til å flytte og tilbøyeligheten til å flytte synker med alderen. Hovedgrunnene til dette er at unge er mindre etablert på arbeidsmarkedet og har dermed mindre grad av regions-spesifikk humankapital som vil gå tapt ved en eventuell utflytting. Forhold som familie, barn og egen bolig er med på å redusere flyttevilligheten, og det er en lavere andel unge som er påvirket av disse faktorene. I tillegg skriver Andreev og Schou at unge er mer ettertraktet i jobbmarkedet og har flere muligheter enn eldre. Det kan blant annet skyldes aldersdiskriminering. For eldre er flyttevilligheten lavere ettersom man ofte er mer etablert både i arbeidsmarkedet og i området man er bosatt. Aldersdiskriminering, kombinert med at det er vanskeligere å finne jobb med høyere inntekt, er faktorer som reduserer flyttevilligheten til personer i denne gruppen. Videre viser analysen at personer med høyere utdanning, både inntil og over 4 år, har høyere grad av flyttevillighet (Andreev og Schou, 2017).



## 4 DATAMATERIALET

Vi har i denne oppgaven benyttet oss av sekundærkilder for å finne nødvendig data. Alle tall er hentet fra NAV og Statistisk sentralbyrå (SSB) sine sider. SSB er en faglig uavhengig institusjon som har som oppgave å samle inn, produsere og publisere offentlig statistikk (SSB, u.å.). NAV har hovedansvar for å ivareta de norske velferdsordningene, samtidig som de samler inn og publiserer en rekke statistikk om blant annet arbeidsmarkedet i landet (NAV, 2023). Både SSB og NAV er dermed anerkjente og pålitelige kilder i Norge. Tallene for utdanningsnivå, innvandrere, alder, små og store virksomheter og innbyggertall er alle hentet fra SSB. Vi har valgt å bruke data fra SSB, da de har en fullstendig oversikt fra år 2017 på kommunenivå. I tillegg er data fra SSB sammenhengende, lett tilgjengelig, samt lett å anvende.

Tall på «helt ledige» har vi hentet fra NAV. Både NAV og SSB har like kriterier for arbeidsledige; «man må være helt uten inntektsgivende arbeid, ha forsøkt å få arbeid og være tilgjengelig for arbeid.» (SSB, 2020). Tallene fra SSB er basert på Arbeidskraftundersøkelsen (AKU) som gjennomføres hvert kvartal. I AKU er dataene basert på et utvalg av befolkningen hvor man innhenter informasjon gjennom telefonintervjuer. Dataen som samles inn brukes videre til å anslå, med så stor nøyaktighet som mulig, andelen arbeidsledige totalt sett i befolkningen. Det kan være noe usikkerhet knyttet til om AKU har et representativt utvalg eller ikke, og om det derfor kan være feil knyttet til tallene fra undersøkelsen. Ettersom AKU tar for seg alle i alderen 15 til 74 år ser vi på denne undersøkelsen som noe utvidet i forhold til det vi ser som relevant for å besvare vår oppgave. De fleste i alderen 15 til 19 år går fremdeles på skole, og pensjonsalderen i Norge er i all hovedsak satt til 67 år (SSB, 2020).

De som derimot er registrert som helt ledige hos NAV, kan man med sikkerhet si er arbeidsledige, da man evaluerer hvert individ som registrerer seg hos NAV i stedet for å gjøre et anslag. Personer som er arbeidsledige, men likevel ikke oppfyller kravene til dagpenger har imidlertid manglende insentiver til å melde seg som arbeidsledig hos NAV. Disse vil dermed ikke være med i statistikken over helt ledige, selv om de står uten arbeid. Selv om kriteriene for å

bli regnet som arbeidsledig er like, kan det altså være personer som regnes som arbeidsledige ifølge AKU, men som ikke har registrert seg hos NAV og motsatt. Dette kan for eksempel skyldes ulike referanseperioder. Statistikken fra NAV gir i motsetning til AKU detaljerte tall helt ned på kommunenivå. Ettersom vi ser på forskjeller i arbeidsledighet mellom kommunene, var det et opplagt valg for oss å benytte tall fra NAV.

#### 4.1 VARIABLER OG HYPOTESER

I oppgaven vår har vi valgt arbeidsledighet som den avhengige variabelen, som er et tall på helt ledige registrert hos NAV. De uavhengige variablene for utdanningsnivå er videregående skole og fagskole, høyere utdanning og ikke fullført eller ikke oppgitt, hvor grunnskole er referansevariabel. Som kontrollvariabler har vi valgt innvandrere, norskfødte med innvandrerforeldre, aldersgruppen 19-34 år, aldersgruppen 35-66 år, små virksomheter, samt innbyggertall i hver kommune.

SSB registrerer den høyest fullførte utdanningen som gjeldende for hvert individ over 16 år. Det vil si at dersom man har fullført to grader er det høyeste graden som er gjeldende, og for de under utdanning vil den tidligere fullførte graden være registrert. Grunnen til at vi har valgt utdanning som en uavhengig variabel er at vi tror det har en stor innvirkning på arbeidsledighet. I tillegg tror vi at forskjellene i utdanning mellom de ulike kommunene kan vise seg å være relativt stor. Vi har valgt å dele opp utdanning i fire variabler. Den første variabelen er grunnskole, som innebærer fullført 1. til 10. klasse. Deretter har vi valgt å sette sammen videregående skole og fagskole til én variabel, da fagskole i mange tilfeller bygger videre på videregående opplæring. Den tredje variabelen vi har valgt er høyere utdanning som er sammensatt av høyere utdanning under og over fire år, ettersom vi tenker at høyere utdanning totalt sett vil ha en liknende effekt på arbeidsledigheten. Den siste variabelen er ikke fullført eller ikke oppgitt, som vi videre velger å kalle ukjent. Dette er en liten variabel som fanger opp de som av ulike årsaker ikke har oppgitt eller ikke fullført utdanningen. Det er rimelig å anta at denne variabelen er liten fordi grunnskole er obligatorisk i Norge. For innvandrere med uklar bakgrunn har SSB gjort et statistisk estimat på hvilket utdanningsnivå man kan anta at hvert

enkelt individ har. SSB opplyser likevel om at en andel av innvandrere ikke har registrert utdanning i statistikken (SSB, 2022).

Vår hypotese er at for de med grunnskole som høyeste fullførte utdanning, samt ukjent, vil øke arbeidsledigheten fordi dette gir mindre muligheter som arbeidstaker på grunn av manglende kompetanse. I tillegg stiller man svakere i forhold til de med høyere utdanning. Videre tror vi at personer med videregående skole eller fagskole, samt høyere utdanning reduserer arbeidsledigheten fordi man anses som mer attraktiv på arbeidsmarkedet. Samtidig oppfyller man nødvendige krav arbeidsgiver kan ha til arbeidstakeren, som for eksempel legesertifisering eller fagbrev som snekker.

De første kontrollvariablene vi har valgt å inkludere er innvandrere og norskfødte med innvandrerforeldre. Grunnen til at vi har valgt disse variablene er at vi tror innvandring kan påvirke arbeidsledighet grunnet utfordringer knyttet til språk, kultur- og samfunnsforskjeller. I tillegg kan det i noen tilfeller, eller for noen yrkesgrupper, hende at utdanning fra utlandet er ugyldig i Norge, fordi det kan stilles andre krav til kompetanse (Meld. St. 7 (2020-2021)). Når det er sagt er det flere som kommer til Norge for å arbeide, noe som kan trekke arbeidsledigheten ned. Vi har valgt å skille mellom innvandrere og norskfødte med innvandrerforeldre fordi norskfødte med innvandrerforeldre ikke nødvendigvis har de samme utfordringene knyttet til språk, kultur- og samfunnsforskjeller. Samtidig følger de norsk utdanningsløp og stiller dermed likt med andre på arbeidsmarkedet. På bakgrunn av dette tror vi at innvandrere vil øke arbeidsledigheten i de norske kommunene. Når det gjelder norskfødte med innvandrerforeldre tror vi effekten på arbeidsledighet vil være mindre enn for variabelen innvandrere, da vi tror de stiller ganske likt som norskfødte uten innvandrerforeldre.

Videre har vi sett på hvordan alder fordeler seg i de ulike kommunene. Vi har valgt å ta utgangspunkt i det vi anser som den sentrale delen av arbeidsmarkedet, 19-66 år, da de fleste i dette aldersspennet vil være yrkesaktive. Vi skiller mellom to grupper, 19-34 år og 35-66 år. I aldersgruppen 19-34 år ser vi for oss at det er mange individer som er nyutdannet, noe som kan

være med på å øke arbeidsledigheten fordi de mangler erfaring fra arbeidslivet. Det kan derfor ta lang tid fra man er ferdig utdannet til man er i arbeid. Med bakgrunn i artikkelen av Andreev og Schou fra 2017, vil arbeidsledigheten sannsynligvis reduseres på sikt fordi unge uten arbeid vil være mer villig til å flytte for å skaffe seg relevant arbeid (Andreev og Schou, 2017). Det er imidlertid naturlig å se for seg at det vil oppstå friksjonsledighet fra man er arbeidsledig til man har flyttet til inntektsgivende arbeid. Vi tror dermed at alderen 19-34 år vil øke arbeidsledigheten.

For personer i alderen 35-66 år antar vi at de fleste er etablerte i arbeidslivet, har arbeidserfaring og er ferdig med eventuell utdanning, noe som virker inn ved å redusere arbeidsledigheten. Vi tror også at denne aldersgruppen har noe høyere terskel for å bytte jobb, noe som også underbygger at arbeidsledigheten reduseres. Likevel ser man tendenser til at individer som mister jobben i en relativt voksen alder, kan bli utsatt for aldersdiskriminering og dermed ha vanskeligheter med å få ny jobb (Andreev og Schou, 2017). Dette kan bidra til å øke arbeidsledigheten. Vi tror imidlertid at denne gruppen er såpass liten at arbeidsledigheten totalt sett vil reduseres.

Vi har valgt å inkludere antall store og små virksomheter innad i kommunene for å få et innblikk i kommunestrukturen, samt informasjon om type arbeidsledighet som kan oppstå. Ettersom vi ikke har funnet litteratur eller andre oppgaver som inkluderer denne variabelen, synes vi det var interessant å se effekten av den. I tillegg ønsker vi å se om det er de store hjørnesteinsbedriftene eller de mindre virksomhetene som er avgjørende for sysselsettingen i kommunene. Vi anser virksomheter som små dersom de har under 50 ansatte. Vi har valgt denne inndelingen fordi regnskapsloven gjør et skille mellom virksomheter over og under 50 ansatte jf. regnskapsloven § 1-6 første ledd (Regnskapsloven, 1998). I utgangspunktet er det naturlig å se for seg at kommuner med mange store virksomheter tilbyr flere arbeidsplasser, og dermed reduserer arbeidsledigheten. Samtidig tror vi at kommuner med høy andel store virksomheter er mer utsatt for friksjonsledighet. Små virksomheter kan dekke et behov i arbeidsmarkedet innad i en

kommune, og på den måten være en sterk bidragsyter til å øke sysselsettingen. Vi tror derfor at arbeidsledigheten vil redusere når andelene små virksomheter øker.

Den siste kontrollvariabelen er innbyggertall, som gir et innblikk i størrelsen på hver kommune. I tillegg ser man ved å inkludere denne variabelen ulike befolknings sammensetninger, utdannings- og kompetansenivåer, samt grad av konkurranse om jobbene som tilbys. Store kommuner har som regel også høyere levekostnader enn små kommuner, blant annet høyere boligpriser, som kan gjøre det vanskeligere å etablere seg. Dette tror vi kan medføre at arbeidsledigheten blir høyere i kommuner med mange innbyggere.

For å forenkle tolkningen av regresjonsanalysen, samt gjøre det lettere å sammenlikne estimeringsresultatene, er variablene omgjort til relative tall. NAV deler helt ledige på den totale arbeidsstyrken for å få arbeidsledige i prosent (NAV, 2009). Videre er tallene for alder og innvandrere delt på antall innbyggere i hver kommune. For utdanning er hver undervariabel, altså grunnskole, videregående skole eller fagskole, høyere utdanning, samt ukjent delt på dem som er 16 år og eldre ved utgangen av 2017. Nevneren tilsvarer altså den totale summen av alle utdanningsnivåene i hver kommune. Til slutt er store og små virksomheter delt på totalt antall virksomheter i hver kommune. Vi har valgt å omgjøre innbyggertall til en logaritme fordi dataene blir mer normalfordelt. I tillegg er det lettere å møte forventningene om konstant varians i regresjonsanalysen.

## 5 EMPIRISK ANALYSE

### 5.1 DESKRIPTIV STATISTIKK

*Tabell 1: Deskriptiv statistikk for utdanningsnivå, avrundet til to desimaler, målt i prosent*

	<b>Gjennomsnitt</b>	<b>Standardavvik</b>	<b>Median</b>	<b>Minimum</b>	<b>Maksimum</b>
<i>Helt ledige</i>	2,28	0,98	2,16	0	5,69
<i>Grunnskole</i>	29,98	5,47	29,53	16,32	52,58
<i>Videregående eller fagskole</i>	45,13	4,52	45,53	28,37	57,00
<i>Høyere utdanning</i>	24,34	5,94	23,24	13,72	52,21
<i>Ukjent</i>	0,55	0,53	0,53	0	1,87

Innledningsvis ønsker vi å se på noen sentrale egenskaper i datasettet, og bruker dermed deskriptiv statistikk. Datasettet inneholder 426 observasjoner, det vil si alle kommunene per 2017. Vi starter med å se på den avhengige variabelen i regresjonsanalysen, helt ledige. Gjennomsnittlig verdi er 2,28 prosent med et standardavvik på 0,98 prosent. Medianen ligger ikke så langt fra gjennomsnittet med en verdi på 2,16 prosent. Kommunene med minst andel arbeidsledige er Utsira, Røyrvik og Høylandet med 0 prosent. Vardø er derimot den kommunen med høyest andel arbeidsledige med 5,69 prosent.

Videre tar vi for oss de uavhengige variablene for utdanningsnivå i regresjonsanalysen; grunnskole, videregående skole eller fagskole, høyere utdanning, samt ukjent. For grunnskole er den gjennomsnittlige verdien 29,98 prosent med et standardavvik på 5,47 prosent, mens for videregående skole eller fagskole ligger gjennomsnittet på 45,13 prosent med et standardavvik på 4,52 prosent. Andelen med høyere utdanning har en gjennomsnittlig verdi på 24,34 prosent med et standardavvik på 5,94 prosent. Ukjent har derimot et lavt gjennomsnitt med 0,55 prosent, og et nesten tilsvarende standardavvik på 0,53 prosent. Det vil si at fordelt på alle kommunene er

det flest som har videregående skole eller fagskole som høyeste utdanning. Samtidig har denne variabelen lavest standardavvik sett bort fra variabelen ukjent.

Medianen for grunnskole er på 29,53 prosent, 45,53 prosent for videregående skole eller fagskole, 23,24 prosent for de med høyere utdanning og 0,53 prosent for ukjent. Medianene til variablene skiller seg dermed ikke noe spesielt ut fra de gjennomsnittlige verdiene. Kommunen med lavest andel som har grunnskole som høyeste fullførte utdanning er Bærum med 16,32 prosent. Oslo er kommunen med lavest andel videregående skole eller fagskole som høyeste fullførte utdanning med 28,37 prosent. Høyere utdanning har lavest verdi på 13,72 prosent gjeldende i Røst, og 0 prosent for ukjent i Granvin. Maksimumsverdiene for grunnskole og videregående skole eller fagskole finner vi i Værøy med 52,58 prosent og Granvin med 57,00 prosent. For høyere utdanning er den høyeste verdien 52,21 prosent i Bærum og 1,87 prosent for ukjent i Balestrand. Forskjellene mellom grunnskole, videregående skole eller fagskole og høyere utdanning er for de maksimale verdiene ikke så ulike hverandre, sammenliknet med gjennomsnittet og medianen. Dette kan være en indikasjon på at det er store variasjoner innad i kommunene.

*Tabell 2: Deskriptiv statistikk for kontrollvariabler, avrundet til to desimaler, målt i prosent*

	<b>Gjennomsnitt</b>	<b>Standardavvik</b>	<b>Median</b>	<b>Minimum</b>	<b>Maksimum</b>
<i><b>Innvandrere</b></i>	10,43	3,59	9,92	2,81	26,51
<i><b>Norskfødte med innvandrereforeldre</b></i>	1,40	1,06	1,14	0,00	7,86
<i><b>19-34 år</b></i>	18,56	2,32	18,38	12,78	29,10
<i><b>35-66 år</b></i>	41,49	1,78	41,37	36,69	47,00
<i><b>Små virksomheter</b></i>	98,89	0,64	99,01	96,40	100,00
<i><b>Innbyggertall (log)</b></i>	8,52	1,19	8,45	5,30	13,41

Fra tabellen ser vi gjennomsnittet, standardavviket, medianen, minimum- og maksimumsverdier for kontrollvariablene. Innvandrere har et gjennomsnitt på 10,43 prosent med et standardavvik på 3,59 prosent, mens norskfødte med innvandrereforeldre er vesentlig lavere med 1,40 prosent og

standardavvik lik 1,06 prosent. I aldersgruppen 19-34 år er gjennomsnittet 18,56 prosent med et standardavvik på 2,32 prosent. Aldersgruppen 35-66 år har gjennomsnitt lik 41,49 prosent og standardavvik 1,78. Det er naturlig at alderen 35-66 år har høyere gjennomsnitt enn alderen 19-34 år, da gruppen inkluderer omtrent dobbelt så mange aldersgrupper. Tallet for små virksomheter er relativt høyt. Det kan skyldes at kommuner med mange store virksomheter også har mange små virksomheter. Oslo har for eksempel 1712 store virksomheter, men likevel utgjør andelen små virksomheter 98 prosent av totalen. Etterfulgt av Oslo har Bergen 602 store virksomheter, men likevel 0,3 prosentpoeng mindre andel små virksomheter enn Oslo. Når vi ser på store virksomheter i antall er det store variasjoner, mens når vi ser på andelen er den relativt lik for de fleste kommunene. Dette gjør at det ikke er så stor forskjell mellom minimums- og maksimumsverdiene som det hadde vært dersom vi hadde sett på antall og ikke andel.

## 5.2 MINSTE KVADRATERS METODE (OLS)

Vi har valgt å bruke minste kvadraters metode i våre regresjonsanalyser. Den første regresjonsmodellen vår kaller vi modell 1, mens den andre regresjonsmodellen vi har gjennomført kaller vi modell 2. Begge modellene har helt ledige som avhengig variabel. Modell 1 er en multippel regresjonsmodell med videregående skole eller fagskole, høyere utdanning og ukjent som de uavhengige variablene. Vi har brukt grunnskole som referansevariabel for utdanningsnivåene. Grunnskole er den mest grunnleggende utdanningen, og dermed gunstig å sammenlikne med høyere utdanning. Ettersom ukjent er en såpass liten variabel ville det vært lite hensiktsmessig å bruke den som referansevariabel. Vi er nødt til å ha referansevariabler for å unngå perfekt multikollinearitet. I modell 2 har vi inkludert innvandrere, norskfødte med innvandrerforeldre, alder 19-34 år, alder 35-66 år, små virksomheter og innbyggertall som kontrollvariabler.



**Modell 1 kan skrives som:**

$$\widehat{\text{HeltLedige}} = \hat{\beta}_0 + \hat{\beta}_1 \cdot \text{VideregåendeFagskole} + \hat{\beta}_2 \cdot \text{HøyereUtdanning} + \hat{\beta}_3 \cdot \text{Ukjent}$$

**Modell 2 kan skrives som:**

$$\begin{aligned} \widehat{\text{HeltLedige}} = & \hat{\beta}_0 + \hat{\beta}_1 \cdot \text{VideregåendeFagskole} + \hat{\beta}_2 \cdot \text{HøyereUtdanning} + \hat{\beta}_3 \cdot \text{Ukjent} \\ & + \hat{\beta}_4 \cdot \text{Innvandrere} + \hat{\beta}_5 \cdot \text{Innvandrerforeldre} + \hat{\beta}_6 \cdot \text{19til34 år} + \hat{\beta}_7 \\ & \cdot \text{35til66år} + \hat{\beta}_8 \cdot \text{SmåVirksomheter} + \hat{\beta}_9 \cdot \text{logInnbyggertall} \end{aligned}$$

*Tabell 3: OLS i prosent for modell 1 og modell 2, standardfeil i parentes*

	<b>Modell 1</b>	<b>P-verdi</b>	<b>Modell 2</b>	<b>P-verdi</b>
<b>Konstantleddet</b>	4,8300	0,000	26,4828	0,001
	(0,750)		(7,593)	
<b>Videregående eller Fagskole</b>	-0,0562	0,000	-0,0297	0,007
	(0,012)		(0,011)	
<b>Høyere utdanning</b>	-0,0213	0,013	-0,0769	0,000
	(0,009)		(0,009)	
<b>Ukjent</b>	0,9183	0,000	0,5339	0,004
	(0,211)		(0,186)	
<b>Innvandrere</b>	-	-	0,1203	0,000
			(0,017)	
<b>Norskfødte med innvandrerforeldre</b>	-	-	-0,2516	0,000
			(0,062)	
<b>19 til 34 år</b>	-	-	0,0561	0,027
			(0,025)	
<b>35 til 66 år</b>	-	-	0,0261	0,342
			(0,027)	
<b>Små virksomheter</b>	-	-	-0,2740	0,000
			(0,075)	
<b>Innbyggertall (log)</b>	-	-	0,3265	0,000
			(0,057)	

Konstantleddet i modell 1 sier at dersom de uavhengige variablene er lik null, altså at alle har grunnskole, vil forventet arbeidsledighet i en kommune være 4,83 prosent. Når man har ett prosentpoeng økning i videregående- eller fagskoleutdanning, alt annet likt, vil arbeidsledigheten reduseres med -0,0562 prosent i forhold til grunnskoleutdanning. Dersom vi holder resten av variablene konstant, vil arbeidsledigheten reduseres med -0,0213 prosent ved ett prosentpoeng økning i høyere utdanning i forhold til grunnskoleutdanning. Effekten av høyere utdanning på arbeidsledighet er mindre enn ved ett prosentpoeng økning av videregående- eller fagskoleutdanning. For ukjent vil arbeidsledigheten ifølge modell 1 øke med 0,9183 prosent i forhold til grunnskoleutdanning når man har ett prosentpoeng økning i variabelen, alt annet konstant. Totalt sett betyr dette at arbeidsledigheten reduseres ved en økning i videregående- eller fagskoleutdanning og høyere utdanning, mens ved en økning i ukjent vil arbeidsledigheten øke i forhold til grunnskole.

Når vi inkluderer kontrollvariablene reduseres effekten av videregående skole eller fagskole på arbeidsledigheten til -0,0297 prosent. Effekten av høyere utdanning på arbeidsledigheten øker når vi inkluderer kontrollvariablene fra -0,0213 prosent til -0,0768 prosent. Når vi inkluderer kontrollvariablene reduseres effekten av variabelen ukjent på arbeidsledigheten til 0,5339 prosent. En større prosentandel som er høyere utdannet innad i en kommune bidrar dermed i størst grad til å redusere arbeidsledigheten når det kommer til utdanning ifølge modell 2.

Ett prosentpoeng økning av variabelen innvandrere i en kommune øker arbeidsledigheten med 0,1203 prosent. Når prosentandelen norskfødte med innvandrereforeldre i en kommune øker med ett poeng vil arbeidsledigheten reduseres med -0,2516 prosent. Innvandrere bidrar dermed til en høyere andel arbeidsledige, mens norskfødte med innvandrereforeldre bidrar til å redusere arbeidsledigheten innad i kommunene i forhold til norskfødte uten innvandrereforeldre. Arbeidsledigheten vil øke med 0,0561 prosent dersom det er ett prosentpoeng økning i aldersgruppen 19-34 år i forhold til referansegruppen, altså innbyggere under 19 år og over 66 år. Ved ett prosentpoeng økning i alderen 35-66 år i en kommune vil arbeidsledigheten øke med

0,0261 prosent i forhold til referansegruppen. Dette betyr at begge aldersgruppene øker arbeidsledigheten innad i en kommune, men den øker mindre i aldersgruppen 35-66 år.

Arbeidsledigheten vil reduseres med -0,2740 prosent ved ett prosentpoeng økning i antall små virksomheter i forhold til store virksomheter i en kommune. Det vil si at kommuner med et høyere andel små virksomheter enn store virksomheter vil ha lavere arbeidsledighet enn kommuner med stor prosentandel store virksomheter. Sett i forhold til flere av de andre variablene har små virksomheter en relativt stor effekt på arbeidsledigheten. For innbyggertall vil en økning på ti prosentpoeng gi en forventet økning på 0,0327 prosent i arbeidsledigheten. Dette innebærer at kommuner med mange innbyggere i gjennomsnitt vil ha en høyere andel arbeidsledige enn kommuner med få innbyggere, alt annet likt.

I modell 1 er R-squared 0,150 som betyr at 15 prosent av variasjonen i arbeidsledighet forklares av utdanningsnivå. Dette kan anses som et relativt lavt tall, noe som kan skyldes at vi har inkludert få variabler i modellen. Samtidig kan det tyde på at de inkluderte variablene forklarer mye av variasjonen. For modell 2 er R-squared 0,400, altså forklarer de uavhengige variablene 40 prosent av den avhengige variabelen. R-squared vil øke desto flere uavhengige variabler man inkluderer i modellen. Adjusted R-squared tar som nevnt tidligere hensyn til dette, og vi bruker dermed denne til å sammenlikne de to modellene. I modell 1 er adjusted R-squared 0,144, mens den er 0,387 i modell 2. Det betyr at modell 2 forklarer mer av variasjonen i den avhengige variabelen, helt ledige, enn modell 1.

### 5.3 HYPOTESETESTING

For å kunne se om våre estimater er troverdige utfører vi hypotesetesting. Den første variabelen vi ønsker å teste er videregående skole eller fagskole. Våre to hypoteser blir dermed som følger:

- Nullhypotesen: det er ingen sammenheng mellom videregående skole eller fagskole og arbeidsledighet.
- Alternative hypotesen: nullhypotesen er ikke sann, det vil si at det er en sammenheng mellom videregående skole eller fagskole og arbeidsledighet.

Vi starter med å utføre en t-test for både modell 1 og modell 2. Denne utføres med 10, 5 og 1 prosent signifikansnivå. Fra modell 1 har vi fått en t-verdi lik -4,508. Med tanke på at frihetsgradene er 422 vil kritisk verdi for 10 prosent signifikansnivå være -1,645. Ettersom t-verdien er mindre enn den kritiske verdien kan vi forkaste nullhypotesen, og dermed påstå at det er en signifikant sammenheng mellom videregående skole eller fagskole og arbeidsledighet ved 10 prosent signifikansnivå. Kritiske verdi for henholdsvis 5 prosent og 1 prosent signifikansnivå er -1,96 og -2,576. I begge tilfellene vil vi forkaste nullhypotesen, ettersom t-verdien er lavere enn disse verdiene. P-verdien forteller når estimatet ikke lenger er signifikant. For variabelen er p-verdien mindre enn 0,0005, som betyr at den ikke vil holde i 0,05 prosent av tilfellene. På bakgrunn av dette kan vi ifølge modell 1 påstå med stor sikkerhet at nullhypotesen ikke er sann. 95 prosent konfidensintervallet for den uavhengige variabelen er mellom -0,081 og -0,032. Det betyr at vi med 95 prosent sannsynlighet kan anta at videregående skole eller fagskole i kommunene har en negativ effekt på arbeidsledigheten, mellom -0,081 og -0,032 prosent i modell 1. Konfidensintervallet i modell 2 er -0,051 til -0,008, altså er effekten av den uavhengige variabelen mindre i forhold til modell 1. For modell 2 er t-verdi lik -2,692 og p-verdi lik 0,007. Det vil si at vi får tilnærmet lik konklusjon for t- og p-test, altså forkaster vi nullhypotesen frem til et signifikansnivå lik 0,007.

Neste variabel vi ønsker å test er høyere utdanning, hvor våre to hypoteser er følgende:

- Nullhypotesen: det er ingen sammenheng mellom høyere utdanning og arbeidsledighet.
- Alternative hypotesen: nullhypotesen er ikke sann, det vil si at det er en sammenheng mellom høyere utdanning og arbeidsledighet.

Med t-verdi lik -2,499 i modell 1 forkaster vi nullhypotesen med 10 og 5 prosent signifikansnivå. For 1 prosent signifikansnivå derimot, er t-verdien større enn kritisk verdi og vi kan dermed ikke påstå den alternative hypotesen i dette tilfellet. P-verdien er 0,013 som betyr at for et signifikansnivå mindre enn 0,013 beholder vi nullhypotesen. Til tross av at vi ikke kan forkaste nullhypotesen ved 1 prosent signifikansnivå, kan vi imidlertid påstå med en viss sikkerhet at det er en sammenheng mellom høyere utdanning og arbeidsledighet ifølge modell 1. Med hensyn på konfidensintervallet fra modell 1 kan vi med 95 prosent sannsynlighet påstå at det er en negativ sammenheng mellom høyere utdanning og arbeidsledighet, mellom -0,038 og -0,005 prosent. I

modell 2 får vi et konfidensintervall fra -0,095 til -0,059, som innebærer en større effekt på arbeidsledigheten enn i modell 1. T-verdien for modell 2 er lik -8,476, og p-verdien er lik 0,00. Det betyr at vi får en annen konklusjon ved 1 prosent signifikansnivå, hvor vi nå kan forkaste nullhypotesen. De ovennevnte konklusjonene består, og vi kan dermed med større sikkerhet påstå at det er en sammenheng mellom høyere utdanning og arbeidsledighet i denne modellen.

Siste variabel innen utdanning er ukjent, hvor vi har følgende hypoteser:

- Nullhypotesen: det er ingen sammenheng mellom ukjent og arbeidsledighet.
- Alternative hypotesen: nullhypotesen er ikke sann, det vil si at det er en sammenheng mellom ukjent og arbeidsledighet.

Med tanke på at t-verdien, 4,349, er større enn de kritiske verdiene kan vi med alle signifikansnivåer forkaste nullhypotesen. Også på bakgrunn av p-verdien kan vi forkaste nullhypotesen, da denne er 0,00. Ukjent har, med 95 prosent sannsynlighet, en positiv effekt på arbeidsledigheten mellom 0,503 og 1,333 prosent. Konfidensintervallet for modell 2 er 0,168 til 0,900. Det vil si at effekten fra variabelen, ukjent, på arbeidsledigheten er mindre sammenlignet med modell 1. T-verdien i modell 2 er lik 0,866, og p-verdien er 0,004. Resultatet vil dermed være det samme i modell 2 som i modell 1, vi påstår den alternative hypotesen i alle testene. Vi vil derimot beholde nullhypotesen ved et signifikansnivå lavere enn 0,004.

For de resterende uavhengige variablene, innvandrere, norskfødte med innvandrerforeldre, alder 19-34 år, alder 35-66 år, små virksomheter og innbyggertall, er p-verdiene mindre enn 0,05 for samtlige variabler, med unntak av alder 35-66 år. Denne variabelen har en p-verdi lik 0,342, og vi vil beholde nullhypotesen for 10, 5 og 1 prosent signifikansnivå. For alderen 19-34 år vil vi kun beholde nullhypotesen ved 1 prosent signifikansnivå, og forkaste den alternative hypotesen ved et signifikansnivå lavere enn 0,027. Resten av variablene har tilstrekkelig p- og t-verdier, som betyr at vi kan påstå den alternative hypotesen i alle tilfellene. Det innebærer at i modell 2 kan vi med stor sikkerhet påstå at det er en sammenheng mellom arbeidsledighet og de uavhengige variablene, innvandrere, norskfødte med innvandrerforeldre, alder 19-34 år, små virksomheter og innbyggertall. Sammenhengen mellom alderen 19-34 år og arbeidsledighet er

imidlertid svakere enn for de andre kontrollvariablene. For alderen 35-66 år kan effekten være lik null, og vi kan dermed ikke konkludere med hvilken effekt variabelen har.

#### 5.4 DRØFTING

Hvis den eller de uavhengige variablene enten forklarer hele eller ikke noe av variasjonen i den avhengige variabelen, vil ikke koeffisientene endre seg når vi inkluderer flere variabler. Dersom variablene derimot kun forklarer deler av variasjonen, vil koeffisientene endre verdier ettersom vi inkluderer eller fjerner kontrollvariabler som korrelerer med den avhengige variabelen.

Kontrollvariablene i modell 2 er en del av feilledet,  $u$ , i modell 1. Effekten videregående skole eller fagskole og ukjent har på arbeidsledigheten reduseres når kontrollvariablene trekkes ut av feilledet og inkluderes i modell 2. Det skyldes at feilledet i modell 1 inneholder variabler som bidrar til å forsterke effekten av de to uavhengige variablene på arbeidsledigheten. For høyere utdanning øker derimot effekten når vi inkluderer flere variabler. Årsaken til dette kan være at én eller flere variabler som inngår i feilledet i modell 1 trekker den uavhengige variabelen i motsatt retning. Når vi i modell 2 trekker disse ut av feilledet blir den negative effekten av høyere utdanning på arbeidsledighet forsterket. For modell 2 vet vi ikke hva feilledet inneholder, og det kan heller ikke testes. Dersom vi sammenlikner modellene vil modell 2 være mer troverdig enn modell 1, siden denne inneholder flere relevante variabler. I tillegg har modell 2 høyere adjusted R-squared. Vi velger med dette å gå videre med modell 2 i drøftingen.

Som nevnt i tidligere deler av oppgaven hadde vi forventninger til hvilke resultater vi ville få ved regresjonsanalysen. I modellen er det høyere utdanning som har størst negativ effekt på arbeidsledigheten, og vi får dermed et resultat av høyere utdanning slik vi hadde forventet oss. Gjennom høyere utdanning vil man tilegne seg sosial kompetanse, samt bedre evne til å lære, noe som sender positive signaler til arbeidsgiver. På denne måten vil høyere utdanning bidra til å øke enkeltindividets humankapital som igjen kan være med på å redusere arbeidsledigheten. At de med høyere utdanning har høyere flyttvillighet gjør at denne gruppen har mindre sannsynlighet for å være arbeidsledige, spesielt over lengre tid. Dersom man er arbeidsledig fordi ens kompetanse ikke etterspørres i kommunen man bor i er de med høyere utdanning, ifølge Andreev og Schou, mer villig til å flytte til en annen kommune hvor ens arbeidskraft

etterspørres. Dette kan bidra til å redusere strukturledigheten. Det er dermed ikke nødvendigvis kun utdanningen som bidrar til å redusere arbeidsledigheten, men også villigheten til å skaffe seg jobb som er relevant for utdanningen (Andreev og Schou, 2017).

Med videregående skole eller fagskole som høyest fullførte utdanning trodde vi arbeidsledigheten ville reduseres. I denne variabelen inngår yrkeskompetanse med fagbrev, studieforbereende og fagskole som yrkesrettet utdanning. Yrkeskompetanse med fagbrev er en yrkesrettet utdanning som gir spesifikk kompetanse og klar overgang til arbeidslivet. Fagskole vil gi mange av de samme mulighetene som yrkeskompetanse med fagbrev. Denne utdanningen forbereder en til å ta del i arbeidsmarkedet ettersom man tilegner seg kunnskap og ferdigheter som er særegen for yrket man søker jobb i. Det kan også bidra til å redusere strukturledighet, da mange kan havne i en situasjon der kravene innen deres bransje endres. Fagskole er dermed et godt alternativ for å imøtekomme arbeidsmarkedet. Vi tror dette kan være noen av årsakene til at variabelen videregående skole eller fagskole som høyest fullførte utdanning, er med på å påvirke arbeidsledigheten negativt. Studieforbereende utdanningsprogram gir bred og lite spesifikk kompetanse, og innebærer et løp hvor det legges til rette for videre utdanning. Ettersom det er tiltenkt at denne gruppen skal ta høyere utdanning tror vi at individer som ikke fortsetter utdanningen etter videregående skole vil øke arbeidsledigheten. Dette kan være årsaken til at videregående skole eller fagskole har mindre effekt på arbeidsledigheten enn høyere utdanning. Det er likevel vanskelig å avgjøre, spesielt ettersom variabelen inkluderer flere undergrupper. De som fortsetter med høyere utdanning vil være registrert med videregående skole som høyeste utdanningsnivå frem til de er ferdig utdannet. Mange av disse vil ikke inngå i statistikken hos NAV, da mange ikke oppfyller kravene til dagpenger og dermed mangler insentiv til å registrere seg som arbeidsledig. Vi tror disse faktorene er grunnen til at variabelen, videregående skole eller fagskole, totalt sett reduserer arbeidsledigheten.

Den siste variabelen innen utdanning er ukjent, hvor vi forventet at variabelen ukjent i forhold til grunnskoleutdanning ville øke arbeidsledigheten. Dette kan skyldes manglende kompetanse og erfaring, noe som gjør at de generelt stiller svakere i arbeidsmarkedet, sammenliknet med andre

utdanningsnivå. Disse personene har ikke nødvendigvis opparbeidet seg samme type humankapital. I tillegg inkluderer denne variabelen uoppgitt nivå av utdanning, og det er dermed vanskelig å si hvilken påvirkning det vil ha for arbeidsledigheten. Samtidig vet vi at mange av de med ukjent utdanning er innvandrere, og vi ser fra regresjonsanalysen at denne gruppen øker arbeidsledigheten. Vi er dermed ikke overrasket over at denne variabelen gir høyere arbeidsledighet.

Høyere utdanning utpeker seg som viktig for å redusere arbeidsledigheten i våre analyser, men vil samtidig ikke være noen garanti mot arbeidsledighet i alle situasjoner. Ifølge With kommer det frem at vi er i en utdanningsekspansjon, hvor unge har langt høyere utdanning enn eldre (With, 2017). Det er verdt å merke seg at høyere utdanning for enkeltindivider kan være positivt, da det kan redusere arbeidsløshet og fattigdom. Samtidig vil marginalnyttens for samfunnet som helhet minke etter hvert som flere tar høyere utdanning. Eksempelvis, dersom en bachelorgrad ikke lenger er noe unikt vil ikke utdanningen ha samme verdi som tidligere. Kravene som stilles til arbeid vil bli høyere og det vil være flere kvalifiserte arbeidstakere. Dermed vil ikke nødvendigvis høyere utdanning i alle situasjoner være eneste avgjørende faktor for å redusere arbeidsledigheten. Desto flere høyt utdannede, desto større vil gapet mellom de med lav og høy utdanning bli. Dette underbygger påstanden om at flere høyt utdannede ikke nødvendigvis vil redusere den totale arbeidsledigheten i de norske kommunene. Ut ifra regresjonsanalysen ser vi likevel at i 2017 var utdanningsnivå en svært viktig faktor for å redusere arbeidsledigheten.

Som nevnt så vi for oss at variabelen innvandrere kom til å øke arbeidsledigheten, noe regresjonsanalysen bekrefter. Vi tror faktorene vi presenterte tidligere i oppgaven omhandlende språk-, kultur- og samfunnsforskjeller, er hovedårsakene til dette. Selv om arbeidsinnvandrere er en del av variabelen ser det ut til at disse ikke har hatt en stor nok påvirkning på variabelen. Det kan imidlertid hende at den positive effekten på arbeidsledigheten ville vært større dersom arbeidsinnvandrere ikke inngikk i variabelen. Også regionspesifikk humankapital kan være en årsak til dette resultatet; ved å være ny i et samfunn eller miljø vil man være nødt til å opparbeide seg ny humankapital gjennom nye relasjoner og kompetanse. Fra artikkelen av Østby kommer



det frem at innvandrere ligger et godt stykke under norskfødte uten innvandrerforeldre når det gjelder utdanning og arbeid (Østby, 2017). Vidal-Gil skriver også at innvandrere utgjør 40 prosent av de ledige (Vidal-Gil, 2017). Dette bidrar dermed til å underbygge resultatene fra vår regresjonsanalyse.

Fra vår hypotese trodde vi norskfødte med innvandrerforeldre ville ha en svakere effekt på arbeidsledigheten enn innvandrere, og stille relativt likt som den øvrige befolkningen. Dette skyldes at mange av de faktorene som kan være en barriere for innvandrere ikke trenger å ha betydning for deres barn, da de gjerne er bedre integrert i samfunnet og utdanningssystemet. Norskfødte med innvandrerforeldre har også gjerne opparbeidet seg mer regionspesifikk humankapital, noe som kan være en årsak til at de stiller sterkere som arbeidstakere enn sine foreldre. Resultatet fra regresjonsanalysen viser at variabelen har en relativt sterk negativ effekt på arbeidsledigheten. Det kan skyldes at en større andel av de norskfødte med innvandrerforeldre tar høyere utdanning sammenliknet med andelen av norskfødte uten innvandrerforeldre (Østby, 2017).

Vår antakelse var at variabelen alder 19-34 år vil øke arbeidsledigheten. Som beskrevet tidligere i oppgaven antar vi at det kan skyldes mangel på erfaring og humankapital. På bakgrunn av dette anser vi det som noe vanskeligere for denne aldersgruppen å skaffe seg arbeid i forhold til aldersgruppen 35-66 år. Samtidig vil denne variabelen inkludere mange unge som er flyttevillige, noe som kan redusere arbeidsledigheten. Likevel vil mange i denne aldersgruppen søke seg til store byer der konkurransen gjerne er høy, som igjen kan øke arbeidsledigheten. Ettersom flere unge enda ikke er så godt etablert på arbeidsmarkedet, og gjerne ønsker å utforske nye muligheter kan det oppstå friksjonsledighet, da man oftere bytter jobb i denne aldersgruppen (Andreev og Schou, 2017). Totalt sett ser vi for oss at dette er årsakene til at en større andel av personer i alderen 19-34 år øker arbeidsledigheten.

Videre har vi antatt at aldersgruppen 35-66 år vil bidra til å redusere arbeidsledigheten. Dette på bakgrunn av at man i denne fasen av livet gjerne er etablert i arbeidsmarkedet, samt har

opparbeidet seg høy grad av humankapital i form av erfaringer og regionspesifikk humankapital. Det er verdt å merke seg at denne variabelen ikke er signifikant, og at vi dermed ikke vet effekten av den.

Vi trodde at med en større andel av små virksomheter i forhold til store virksomheter i en kommune ville arbeidsledigheten reduseres. Dette viser seg å stemme i henhold til regresjonsanalysen, men effekten var enda større enn forventet. Hypotesen vår var basert på at kommuner med høy andel store virksomheter er mer utsatt for friksjonsledighet, samt at små virksomheter dekker behov i arbeidsmarkedet. Disse faktorene kan være noen av årsakene, men det kan samtidig være mange andre mulige årsaker som spiller inn. Det kan blant annet være at små virksomheter bidrar til å diversifisere og skape verdier, samt innovasjon fordi de raskere kan tilpasse seg endringer i markedet. Selv om små virksomheter er mer utsatt ved nedgangstider på grunn av færre midler enn store virksomheter, vil ikke dette ha like stor betydning på den totale arbeidsledigheten, i motsetning til nedbemanning av store virksomheter. Vi klarer ikke å se en tydelig trend fra datasettet, noe som tidligere nevnt kan skyldes at vi ser på andeler istedenfor antall. Samtidig kommer det frem at andelen små virksomheter er avgjørende for å redusere arbeidsledigheten i de norske kommunene. Det betyr at de store hjørnesteinsbedriftene ikke har like stor effekt som de små virksomhetene, ifølge vår analyse.

Fra hypotesen trodde vi at kommuner med høyere innbyggertall ville øke arbeidsledigheten, noe analysen bekrefter. Dette tror vi kan ha sammenheng med større konkurranse om arbeidsplassene, samt høyere krav til kvalifikasjoner i større kommuner, noe som kan skyldes at jobbene er mer attraktive. Vi har fra analysen sett en tendens til at andelen med høyere utdanning er større i kommunene med store byer der befolkningen er høy. Eksempelvis ser vi at andelen med høyere utdanning i Oslo er på 50 prosent, noe som er over 25 prosentpoeng mer enn gjennomsnittet på 24,34 prosent. Det vil si at de som er lavt utdannet vil stille svakere på arbeidsmarkedet fordi gapet opp til de høyt utdannede er større, og kan være en årsak til at arbeidsledigheten øker når innbyggertallet i en kommune er større. Altså er effekten av innbyggertall interessant når den sees i sammenheng med de andre uavhengige variablene.

## 5.5 ANTAGELSENE

Som tidligere nevnt er det viktig å sjekke for antagelsene i multippel lineær regresjon for å se om resultatene fra regresjonsanalysen er holdbare og dermed til å stole på. Dersom de fire første antagelsene, MLR.1 til MLR.4, holder, kan vi konkludere med at estimatene fra OLS-regresjonen er forventningsrett, altså:

$$E(\hat{\beta}_j) = \beta_j$$

Det vil si at den forventede effekten fra koeffisientene gjennomsnittlig er lik effekten av populasjonsparametret. Hvis alle antakelsene, MLR.1 til MLR.6, holder vil vi ha beste lineære forventningsestimator (BLUE).

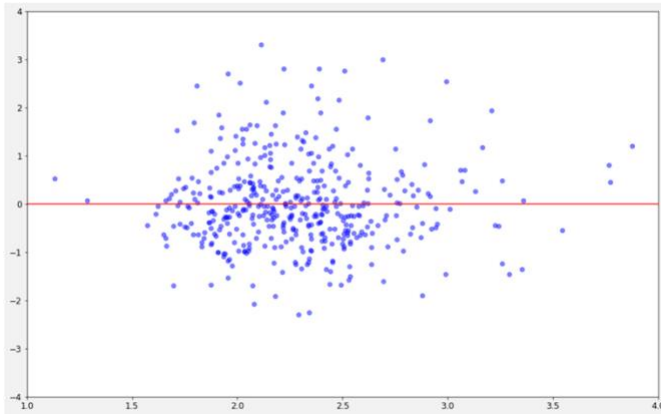
**MLR.1:** Vi kan ikke si noe om lineære parametere i modellen vår, og antar dermed linearitet i våre parametere.

**MLR.2:** Hensikten med analysen er å kunne si noe om arbeidsledigheten i de norske kommunene. Med tanke på at vi har inkludert alle kommunene vil vi påstå at modellen vår er representativ for populasjonen vi ønsker å si noe om.

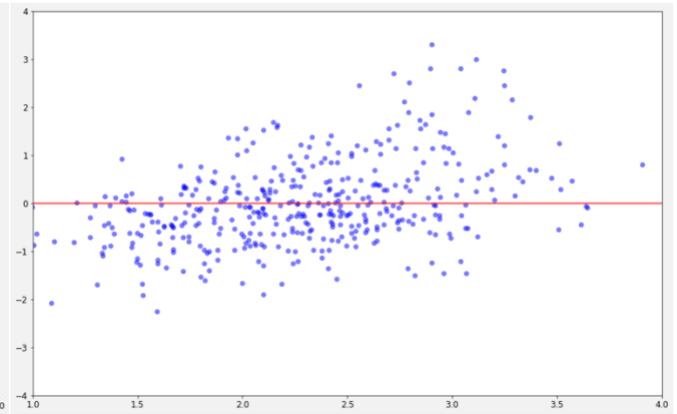
**MLR.3:** Selv om vi har valgt å fjerne noen variabler for å unngå perfekt multikollinearitet kan vi likevel ikke utelukke at det er en viss korrelasjon mellom de uavhengige variablene. Samtidig er variablene signifikante, og vi antar derfor at denne antagelsen holder.

**MLR.4:** Antagelsen om at restleddet har en forventet verdi lik null er nødvendig for å tolke estimatene. I modell 1 har vi bevist at ved å inkludere kontrollvariabler i modell 2 hadde ikke restleddet en forventet verdi lik null. Det er imidlertid stor sannsynlighet for at det er andre enn de vi valgte å bruke som uavhengige variabler i modell 2 som også kunne vært inkludert i modellen. Årsaken er at det sannsynligvis er mange andre faktorer som forklarer forskjellene i arbeidsledigheten innad i de norske kommunene som korrelerer med de uavhengige variablene. Dette er ikke mulig å teste, og antagelsen om at restleddet har en forventet verdi lik null er det grunn til å være i tvil om.

**MLR.5:** For å sjekke om modellene oppfyller kravene til homoskedastisitet bruker vi modellene under med de predikerte verdiene på x-aksen og residualene på y-aksen.



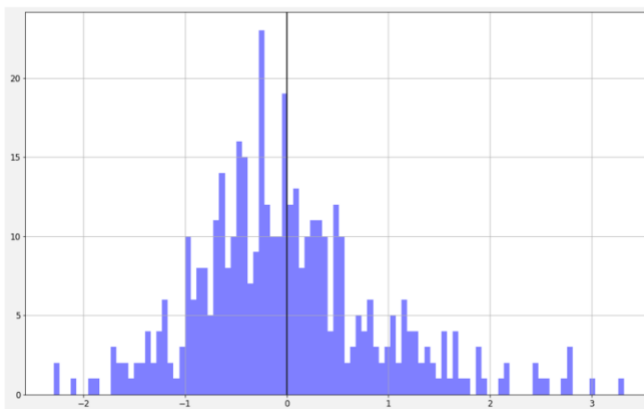
**Figur 1:** Scatter plot modell 1



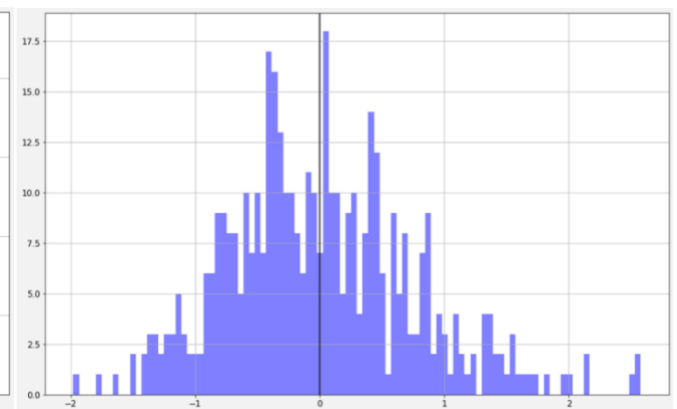
**Figur 2:** Scatter plot modell 2

Den første modellen viser ingen tegn til mønster. Det er relativ stor spredning med en tendens til sentralisering mellom 2,0 og 2,5 på de predikerte verdiene. Modellen ser dermed ut til å ha en relativ konstant variasjon, og vi kan dermed anta at kravet om homoskedastisitet er oppfylt. For modell 2 ser man en relativt stor spredning, men en svak positiv lineær sammenheng hvor lave predikerte verdier henger sammen med lave residualer og motsatt. Samtidig kan det se ut som det er litt større variasjon i feilleddene ved høye predikerte verdier, men denne tendensen er relativt lav. Det kan tyde på at det er en svak tendens til heteroskedastisitet.

**MLR.6:** For å sjekke om residualene er normalfordelte med gjennomsnitt lik 0 lager vi et histogram som viser normalfordelingen, samt regner ut gjennomsnittet av residualene.



**Figur 3:** Normalfordelingskurve modell 1



**Figur 4:** Normalfordelingskurve modell 2

Den svarte linjen representerer gjennomsnittet, og vi ser at den er lik 0 i begge modellene. Utregningen forteller oss at gjennomsnittet er lik 0 for både modell 1 og modell 2. Begge modellene er tilnærmet normalfordelt med kun en svak tendens til å være høyreskjev. Modell 1 er litt mer høyreskjev enn modell 2. Vi antar dermed at denne antagelsen er oppfylt.

## 5.6 MULIGE SVAKHETER VED OPPGAVEN

Ved utarbeidelsen av problemstilling og resterende deler av denne oppgaven har vi sett oss nødt til å begrense tematikken. Dette har gjort at vi kun kan forklare deler av arbeidsledigheten i kommunene, men målet med oppgaven var i all hovedsak å kunne si noe om effekten av utdanning på arbeidsledigheten. Det at vi benytter tall på arbeidsledige fra NAV, som ikke fanger opp alle AKU anser som arbeidsledige, kan gjøre at det er mange arbeidsledige som faller utenfor statistikken. Samtidig vet vi ikke hvordan de arbeidsledige som ikke er inkludert i statistikken fordeler seg i de ulike kommunene, noe som kan gi en skjevfordeling.

Som nevnt innledningsvis har vi sett oss nødt til å begrense denne oppgaven betraktelig for å besvare den så godt som mulig, gitt de rammene vi har. På bakgrunn av dette er det mange interessante variabler som kan forklare arbeidsledigheten i de ulike kommunene som dessverre faller bort. En relevant, uavhengig variabel som kunne vært forklarende er hvilke næringer de ulike kommunene har. Man skiller gjerne mellom primær-, sekundær-, og tertiærnæringer. Spesielt potensiale til primærnæringer som fiske, er mer tilgjengelig for kommunene med kystlinje. I tillegg vil jordsmonn og mulighet til skog- og/ eller jordbruk være mer egnet og utbredt i noen kommuner. Likeledes vil tertiærnæringsens størrelse i de ulike kommunene avhenge av innbyggertall eller i hvilken grad kommunene har turisme som kan bidra til at tertiærnæringsen vokser. De ulike bransjene er også påvirket av etterspørselen i arbeidsmarkedet, dermed kan det oppstå høy arbeidsledighet i yrker hvor etterspørselen er lav.

Ulike familiestatuser i kommunene kunne også vært interessant å trekke inn i vår analyse. Det vil si sivilstatus, antall barn og husholdning. Dette vil kunne påvirke valg av bosted og flyttevillighet. Også forskjellige satser på arbeidsgiveravgift, eiendomsskatt og inntektsskatt vil

kunne si noe om arbeidsmarkedet, boligmarkedet og dermed arbeidsledigheten innad i kommunene. Ved å trekke inn disse faktorene som uavhengige variabler ville man fått et mer helhetlig bilde av hva som påvirker arbeidsledigheten og eventuelt øke forklaringsevnen til analysen. Det er dog vanskelig å si sikkert om dette ville forbedret vår regresjonsanalyse.

Vi har gjort en grundig jobb i å forsøke å finne tidligere litteratur omkring dette emnet og denne problemstillingen. Det har blitt publisert mange relevante artikler vi har benyttet oss av som viser til hvordan flere av de uavhengige variablene er tilknyttet arbeidsledighet. Likevel har vi møtt på utfordringer med å finne tidligere litteratur som tar for seg arbeidsledighet som avhengig variabel. Dette har vist seg å være svært begrenset, noe som har gjort vårt arbeid noe vanskeligere fordi vi har mistet muligheten til å innhente informasjon fra, samt sammenlikne våre resultater med annen litteratur. Vi mener dette kunne vært til hjelp, og relevant for å løfte oppgaven. Dette gjør imidlertid oppgaven enda mer spennende, da vi får mulighet til å se på noe som ikke er forsket så mye på tidligere.

Vi har fått et godt bilde på hvordan utdanningsnivå påvirker arbeidsledigheten i 2017. Selv om vi regner med at våre resultater og konklusjoner også hadde holdt i dag, kan det likevel være forskjeller med tanke på at vi bruker tall som er 6 år gamle. I årene etter 2017 har vi sett store forskjeller i arbeidsledigheten, blant annet på grunn av koronapandemien. Det gjør at vår analyse ikke nødvendigvis er like relevant i dag. I tillegg kan det være hendelser vi ikke er klar over innad i kommunene som påvirker arbeidsledigheten i 2017. Dette kan blant annet ha ført til unormalt høye tall på arbeidsledigheten for noen kommuner. For å unngå disse utfordringene kunne vi brukt tidsserier eller panel data istedenfor tverrsnitt data, da dette kunne gitt et mer helhetlig bilde over tid. Dette hadde imidlertid blitt såpass omfattende at vi har valgt å bruke tverrsnitt data.

## 6 KONKLUSJON

I denne oppgaven har vi undersøkt sammenhengen mellom utdanningsnivå og arbeidsledighet i norske kommuner. For å besvare problemstillingen har vi brukt minste kvadraters metode. I tillegg til utdanningsnivå har vi inkludert andre uavhengige variabler som er med på å forbedre, utvide og gi en dypere forståelse av analysen. De bidrar blant annet til å forklare hvorfor arbeidsledigheten varierer på tvers av kommunene. Kontrollvariablene som hadde størst negativ effekt på arbeidsledigheten ifølge vår analyse, var andelen små virksomheter og norskfødte med innvandrerforeldre. Dette var overraskende resultater, da vi i utgangspunktet trodde at høyere utdanning ville ha størst effekt på arbeidsledigheten. I forhold til grunnskole er høyere utdanning den utdanningsvariabelen som bidrar til å redusere arbeidsledigheten mest etterfulgt av videregående skole eller fagskole. Variabelen ukjent øker arbeidsledigheten i størst grad. At høyere utdanning er viktig for å redusere arbeidsledigheten, samt at de med lavt utdanningsnivå utgjør en stor andel av de arbeidsledige, støttes opp av tidligere litteratur. Vi kan på bakgrunn av oppgaven konkludere med at utdanningsnivå er en viktig faktor for å forklare forskjellene i arbeidsledigheten i norske kommuner. Ifølge våre analyser innebærer dette at desto høyere utdanningsnivået i kommunene er, desto lavere blir arbeidsledigheten.

## 7 REFERANSELISTE

Andreev, L., Schou, L. (2017) *Mobilitet og flyttevillighet blant arbeidsledige*. Tilgjengelig fra: [https://arbeidogvelferd.nav.no/asset/2017/3/Arbeid\\_og\\_velferd-2017-03\\_art-2.pdf](https://arbeidogvelferd.nav.no/asset/2017/3/Arbeid_og_velferd-2017-03_art-2.pdf) (Hentet: 02. mars 2023).

Borjas, G. J. (2008) *Labor economics*. 4. utg. New York: McGraw-Hill Companies.

Christoffersen, L., Johannesen, A., og Tuft, P. A. (2020). *Forskningsmetode for økonomisk-administrative fag*. 4. utg. Oslo: Abstrakt forlag.

Gulbrandsen, C. (2022) *Utdanning er ikke en garanti mot arbeidsløshet og fattigdom*. Tilgjengelig fra: <https://www.oslomet.no/forskning/forskningsnyheter/utdanning-ingen-garanti> (Hentet: 03. mars 2023).

Holden, S. (2016). *Makroøkonomi*. 1.utg. Oslo: Cappelen Damm

Meld. St. 7 (2020-2021) (2021) *En verden av muligheter – Internasjonal studentmobilitet i høyere utdanning*. Oslo: Kunnskapsdepartementet. Tilgjengelig fra: <https://www.regjeringen.no/no/dokumenter/meld.-st.-7-20202021/id2779627/?ch=10> (Hentet: 28. mars 2023).

NAV (2009) *4. Begreper, kjennemerker og grupperinger*. Tilgjengelig fra: [https://www.nav.no/no/nav-og-samfunn/statistikk/arbeidssokere-og-stillinger-statistikk/relatert-informasjon/om-statistikken-arbeidssokere/4.begreper-kjennemerker-og-grupperinger\\_kap](https://www.nav.no/no/nav-og-samfunn/statistikk/arbeidssokere-og-stillinger-statistikk/relatert-informasjon/om-statistikken-arbeidssokere/4.begreper-kjennemerker-og-grupperinger_kap). (Hentet: 14. mars 2023).

NAV (2018) *Helt ledige. Fylke og kommune. Tidsserie måned*. Tilgjengelig fra: <https://www.nav.no/attachment/download/92e77709-aad3-4998-90f1-f9c04b06c180:bfc0d53b35dbb3b7f12bb44ef2e0f75c2b981127/201712-hl060-helt-ledige.fylke->



[og-kommune.tidsserie-maaned.pdf](#) (Hentet: 7. mars).

NAV (2019) *Helt ledige*. Tilgjengelig fra: <https://www.nav.no/no/nav-og-samfunn/statistikk/arbeidssokere-og-stillinger-statistikk/helt-ledige> (Hentet: 7. mars 2023).

NAV (2023) *Hva er NAV?* Tilgjengelig fra: <https://www.nav.no/hva-er-nav#samfunnsoppdrag> (Hentet: 14. mars 2023).

OECD (2018) *Investing in Youth: Norway*. Tilgjengelig fra: <https://doi.org/10.1787/9789264283671-en> (Hentet: 03. mars 2023).

Regnskapsloven (1998). *Lov om årsregnskap m.v.* Tilgjengelig fra: <https://lovdata.no/dokument/NL/lov/1998-07-17-56> (Hentet: 08. mars 2023).

SSB (u.å.) *SSBs virksomhet*. Tilgjengelig fra: <https://www.ssb.no/omssb/ssbs-virksomhet>. (Hentet: 14. mars 2023).

SSB (2020) *Hvorfor ulike arbeidsledighetstall?* Tilgjengelig fra: <https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/hvorfor-ulike-arbeidsledighetstall--410344>. (Hentet: 8. mars 2023).

SSB (2022) *Befolkningens utdanningsnivå*. Tilgjengelig fra: <https://www.ssb.no/utdanning/utdanningsniva/statistikk/befolkningens-utdanningsniva> (Hentet: 7. mars 2023).

Vidal-Gil, E. (2017) *Endringer i arbeidsledigheten over tid*. Tilgjengelig fra: <https://www.nav.no/no/nav-og-samfunn/kunnskap/analyser-fra-nav/arbeid-og-velferd/arbeid-og-velferd/endringer-i-arbeidsledigheten-over-tid> (Hentet: 21. mars 2023).

With, M. L. (2017) *Arbeid og utdanning*. Tilgjengelig fra: <https://www.ssb.no/sosiale-forhold-og-kriminalitet/artikler-og-publikasjoner/arbeid-og-utdanning> (Hentet: 07. mars 2023).

Wooldridge, J. M. (2012) *Introductory Econometrics a modern approach*. 5. utg. Mason: Cengage Learning, Inc.

Østby, L. (2017). *Bedre integrert enn mor og far*. Tilgjengelig fra:

<https://www.idunn.no/doi/full/10.18261/ISSN1504-3045-2017-01-05> (Hentet: 02. mars 2023).

