

Brita Engen Finne
Sunniva Wergeland

Lærermangel og arbeidsledighet

En paneldatanalyse av effekten av
arbeidsledighet på lærermangel i grunnskolen

Masteroppgave i Samfunnsøkonomi
Veileder: Torberg Falch
Juni 2023

Brita Engen Finne
Sunniva Wergeland

Lærermangel og arbeidsledighet

En paneldatanalyse av effekten av arbeidsledighet
på lærermangel i grunnskolen

Masteroppgave i Samfunnsøkonomi
Veileder: Torberg Falch
Juni 2023

Norges teknisk-naturvitenskapelige universitet
Fakultet for økonomi
Institutt for samfunnsøkonomi



Kunnskap for en bedre verden

Sammendrag

Lærere med kompetanse er av betydning for å sikre kvaliteten på undervisningen i grunnskolen. Imidlertid har vi en utfordring på skoler i norske kommuner med mangel på formelt kvalifiserte lærere. Denne masteroppgaven undersøker om arbeidsmarkeds-situasjonen påvirker læreres valg mellom å bli i læreryrket eller finne andre jobber utenfor skolesystemet. Hypotesen er at når arbeidsledighetsraten reduseres fører det til økt lærermangel. I analysen defineres lærermangel som andelen årsverk utført av lærere uten formell kompetanse. Oppgaven baserer seg på en tidligere analyse gjort av Falch, Johansen og Strøm i 2009. I deres analyse fant de en negativ signifikant effekt av regional arbeidsledighet på lærermangel for tidsperioden 1981-2002.

Utgangspunktet for oppgaven er et paneldatasett som inneholder data for norske kommuner og skoler i perioden 1995-2019. Datasettet omfatter informasjon fra alle skoler i Norge. I tillegg har vi arbeidsledighetsraten på kommunenivå og kommunestatistikk.

I perioden 1995-2007 finner vi en negativ signifikant effekt av arbeidsledighet på lærermangel. Dette indikerer at en reduksjon i arbeidsledighetsraten er assosiert med økt lærermangel i grunnskolen. Imidlertid finner vi at denne effekten forsvinner i resten av perioden, 2008-2019, noe som tyder på en endring i løpet av tidsperioden.

Abstract

Qualified teachers are of importance to ensure the quality of teaching in primary school. However, we have a challenge in schools in Norwegian municipalities with a shortage of formally qualified teachers. This master's thesis examines whether the condition in the labor market affects teachers' choices between staying in the teaching profession or finding other jobs outside the school system. The hypothesis is that a reduction in the unemployment rate leads to increased teacher shortages. In the analysis, teacher shortage is defined as the proportion of full-time positions carried out by teachers without formal qualifications. The thesis is based on a previous analysis conducted by Falch, Johansen, and Strøm in 2009. In their analysis, they found a negative significant effect of regional unemployment on teacher shortages for the period 1981-2002.

The starting point for the thesis is a panel dataset that contains data for Norwegian municipalities and schools from 1995 to 2019. The dataset includes information from all schools in Norway. In addition, we have the unemployment rate at the municipal level and municipal statistics.

In the period 1995-2007, we find a negative significant effect of unemployment on teacher shortages. This indicates that a reduction in the unemployment rate is associated with increased teacher shortages in primary schools. However, we find that this effect disappears in the rest of the period, 2008-2019, suggesting a change over time.

Forord

Denne masteroppgaven markerer slutten på vår studietid ved NTNU i Trondheim. Oppgaven er et felles arbeid gjort av Brita Engen Finne og Sunniva Wergeland. Samarbeidet har vært spennende, lærerikt og kjekt. Vi ønsker å rette en takk til vår veileder, Torberg Falch. Vi setter stor pris på diskusjonene, tilbakemeldingene og rådene vi har fått underveis i arbeidsprosessen. Videre ønsker vi å takke samtlige medstudenter som har sørget for god stemning!

Brita Engen Finne & Sunniva Wergeland
Trondheim, 1.juni 2023

Innhold

1	Introduksjon	1
2	Teori og litteratur	3
2.1	Betydningen av lærere	3
2.2	Lærererarbeidsmarkedet	4
2.3	Betydningen av arbeidsledighet på lærermangel	6
2.4	Utbredelse av lærermangel	8
3	Institusjonelt rammeverk	9
3.1	Skolenes rammebetingelser	9
3.2	Lønnsdannelse	10
3.3	Arbeidsledighet	11
4	Datamaterialet	12
4.1	Datakilder	12
4.2	Lærermangel	14
4.3	Data på arbeidsledighet	16
4.4	Kontrollvariabler	20
5	Metode	24
5.1	Modellspesifikasjoner	24
5.1.1	Modellspesifikasjoner på skolenivå	24
5.1.2	Modellspesifikasjoner på kommunenivå	25
5.2	Statistiske paneldatamodeller	26
5.2.1	Sammenslått minste kvadraters metode (POLS)	28
5.2.2	Faste effekter (FE)	29
5.2.3	Tilfeldige effekter (RE)	29
5.3	Tester	30
5.3.1	F-test	31
5.3.2	Hausman-test	32
6	Resultat	33
6.1	Hovedmodell: Skolenivå	33
6.1.1	Valg mellom estimeringsmetoder	35
6.1.2	Skolenivå med kontrollvariabler	36
6.2	Hovedmodell: Kommunenivå	39
6.2.1	Kommunenivå med kontrollvariabler	40
6.3	Robusthet	42
6.3.1	Analyse med den andre variabelen på arbeidsledighet	43
6.3.2	Annet mål på lærermangel	45

6.3.3	Arbeidsmarkedsregion	47
6.4	Heterogenitet	50
6.4.1	Dele i to tidsperioder	50
6.4.2	To perioder med arbeidsmarkedsregional ledighet	53
6.4.3	Skille mellom skole typer	54
7	Avslutning	56
7.1	Endring fra tidligere funn	57
7.1.1	Endringer i perioden	58
7.2	Utfordringer med oppgaven	59
7.3	Konklusjon	60
	Referanser	61
A	Appendiks	63
A.1	Deskriptiv statistikk	63
A.2	Tabeller	64

1 Introduksjon

I følge Utdanningsforbundet er det omlag 40 000 lærerutdannede som jobber i andre bransjer (Utdanningsforbundet, 2022). Høsten 2022 var det en langvarig streik blant lærere etter forhandlinger med kommunen som arbeidsgiver. Streiken ble stoppet av tvungen lønnsnemnd fordi konflikten mellom partene syntes å være helt fastlåst (Regjeringen, 2022). At lærerne er misfornøyd med lønnsutviklingen til yrket er med på å legge press på statusen til læreryrket. Lærerutdanning sikrer at lærere besitter viktig og spesialisert fagkompetanse og didaktikk. Studier på betydningen av kvalifiserte lærere i klasserommet viser at kompetanse gir utslag på elevenes prestasjoner.

Problemstillingen om lærermangel i skolen kan betraktes i forbindelse med bærekraftsmålsetting om god utdanning. Utdanning er en grunnleggende menneskerettighet, i tillegg til at det er en viktig forutsetning for å kunne oppnå et bærekraftig samfunn. Å løse lærermangelproblemet kan dermed bidra til å oppfylle andre målsettinger. UNESCO stadfester at målsettingen må sikre at undervisning er undervist av godt kvalifiserte, utdannende, tilstrekkelig avlønnede og motiverte lærere (UNESCO, 2015, s.30).

Denne masteroppgaven har som formål å undersøke hvilken effekt arbeidsledighet har på graden av lærermangel i grunnskolen. Oppgaven er et forsøk på å oppdatere en tidligere analyse på lærermangel og konjunktursvingninger av Falch, Strøm og Johansen i 2009. Vi benytter nyere data av bedre kvalitet enn dem. Vi definerer vår variabel på lærermangel på lik måte, ved at vi ser på andelen uten formell kompetanse i skolen. Hypotesen er at lavere arbeidsledighet gir høyere lærermangel. I perioder med lav arbeidsledighet vil lærere med kompetanse ha større muligheter til å finne andre jobber utenfor skolesystemet. Det betyr at under høykonjunktur forventer vi at flere lærere vil søke seg vekk fra yrket, slik at disse til en viss grad må erstattes av ukvalifiserte arbeidstakere.

Lærermangel vil kunne identifiseres gjennom endringer i kvalitetskaraktistikker til lærerstaben. Kvalitetskaraktistikker refererer til de egenskapene og kvalifikasjonene som lærerne besitter. I denne sammenhengen handler det spesifikt om andelen lærere som ikke oppfyller kravene til formell kompetanse. Selv om det vil være tilstrekkelig antall lærere tilstede på skolene og i klasserommet, er det lærerenes kompetanse som er av interesse når graden av lærermangel vurderes. En økning i andelen lærere uten tilstrekkelig kompetanse indikerer dermed økt lærermangel, selv om klasserommet alltid er bemannet.

Vi benytter paneldata fra 1995-2019 som omfatter informasjon om norske skoler og kommuner. Når vi deler perioden i to, 1995-2007 og 2008-2019, finner vi at arbeidsledigheten i kommunen har en negativ signifikant effekt på lærermangel i den første perioden. I den andre perioden finner vi imidlertid ikke lenger den signifikante effekten av arbeidsledighet. Vi gjennomfører analysen både på skole- og kommunenivå. I robusthetskapittelet benytter vi ledighet for arbeidsmarkedetsregionen istedenfor kommunal arbeidsledighet, og finner en signifikant negativ effekt på lærermangel. Dette indikerer at arbeidsmarkedet i en noe større region enn kommunen har betydning.

Opgaven er strukturert slik at vi først presenterer tidligere litteratur om betydningen av lærere og sammenhengen mellom lærerarbeidsmarkedet og arbeidsledighet. Kapittel 3 orienterer om rammeverket til skolene, lønnsdannelsen for lærere og arbeidsledighet. Datamaterialet som er brukt i analysen blir presentert i kapittel 4, hvor variabelen for lærermangel blir først lagt frem. Deretter presenteres arbeidsledighetsvariablene og kontrollvariablene. I kapittel 5 presenteres modellspesifikasjonene og estimeringsmetodene som brukes i analysen. Resultatene blir lagt frem i kapittel 6, hvor vi begynner med hovedanalysen på skolenivå, etterfulgt av en tilsvarende analyse gjort på kommunenivå. Deretter følger en robusthetsanalyse hvor vi tester målene på arbeidsledighet og lærermangel. Videre tester vi for heterogenitet, hvor vi deler utvalget i to perioder og skiller mellom skoletyper. Avslutningsvis oppsummerer vi oppgaven og diskuterer funnene opp mot tidligere analyser.

2 Teori og litteratur

I dette kapitlet blir tidligere litteratur og teori presentert. Først blir betydningen av lærere og kompetanse trukket frem. Analysen fra Falch, Strøm og Johansen presenteres, og videre knyttes lærermangel til mekanismene i lærerarbeidsmarkedet. I tillegg presenteres andre analyser på sammenhengen mellom lærermangel og tilstanden i arbeidsmarkedet. Til slutt kommenteres det på utbredelsen av lærermangel.

2.1 Betydningen av lærere

Andelen lærere med kvalifisert kompetanse ved skoler er av interesse da den har betydning for kvaliteten på grunnskoletilbudet. Det er funnet sterke bevis for at lærerkvalitet er en av de viktigste faktorene for skoleelevers prestasjoner og oppnåelser (Rivkin, Hanushek, Kain, 2000). Imidlertid er det utfordrende å måle hvor stort bidraget av lærerne er. Det har heller ikke vært mulig å identifisere noen spesifikke egenskaper ved lærere som er pålitelig relatert til elevresultater (Hanushek, 2011). Læring er utfordrende å måle da det er et resultat av et komplisert samspill mellom flere faktorer rundt elevene. Disse faktorene er blant annet lærere, foresatte og tilgang på relevante læringsressurser. I den forbindelse diskuteres betydningen av formell kompetanse i forbindelse med kvalitet på undervisningstilbudet blant forskere.

Kvaliteten på undervisningen kan ses i sammenheng med bestemmelsen av antall elever per lærer som settes, og med det hvor tett oppfølging hver enkelt elev får. En fremgangsmåte for å sette et mål på lærerkvalitet, er å basere seg på elevenes faglige prestasjoner. Kompetanse og kvalitet hos læreren kan knyttes til om en oppfyller krav til formell kompetanse gjennom å fullføre lærerutdanning. En svensk analyse har funnet at lærere som ikke er formelt kvalifisert fører til svakere elevprestasjoner (Andersson et al., 2008). Elevprestasjonene ble i analysen målt utifra "testscoren" til elevene, gitt ved hvorvidt lærerne var kvalifisert eller ukvalifisert.

Sosiale ferdigheter og personlige egenskaper kan også argumenteres for at påvirker kvaliteten på undervisningen. Lærerutdanning skal derimot sikre at lærere har kompetanse innenfor fag, didaktikk og pedagogikk, i tillegg til å kunne bygge relasjoner. Bruken av kvalifikasjonskrav ved ansettelse kan derfor fungere som en nedre grense for å sikre en viss kvalitet. En formell kvalifikasjon kan effektivisere søknadsprosessen ved å selektere ut mindre aktuelle kandidater til en stilling.

2.2 Lærerarbeidsmarkedet

Denne oppgaven tar utgangspunkt i analysen gjort av Falch, Strøm og Johansen fra 2009 i "Teacher shortage and the business cycle". I analysen undersøkes tilgjengeligheten av lærere med visse karakteristikk og hvordan tilgjengeligheten varierer med arbeidsledighetsraten (Falch et al., 2009). For å forklare sammenhengen mellom lærermangel og arbeidsledighetsraten bruker de en modell for arbeidsmarkedet til lærere. I denne modellen antas det at lærermangel og relativ lønn for lærere bestemmes simultant. Boardman, Darling-Hammond og Mullin (1982) formulerer en modell for lærerarbeidsmarkedet gitt ved tilbud og etterspørsel, hvor individet må ta stilling til om en skal arbeide som lærer eller velge en alternativ jobb. Denne modellen fungerer som et rammeverk for å forstå mekanismen i lærerarbeidsmarkedet.

Etterspørselen etter lærere avhenger av elevtallet og det myndighetene setter som minimumskrav for antall elever per lærer. Grunnskolen finansieres gjennom kommunen, og deres budsjettbetingelse som gir etterspørselsfunksjonen er

$$Y = W \cdot L + K$$

Her er Y kommunenes disponible inntekt. W er lærerlønn og L er antall lærere, slik at $W \cdot L$ er de totale utgiftene til lærerlønn. K viser til kostnader som går til andre kommunale tjenester. Budsjettbestemmelsene i kommunene antas å være eksogene i denne modellen. Kommunene ønsker å maksimere tjenestetilbudet gitt budsjettet de setter. Med hensyn til budsjettbetingelsen vil kommunens etterspørselen etter lærere være gitt ved

$$L = L(W, K, Y, X)$$

Endringer i variablene for lærerlønn, kostnader til andre kommunale tjenester og kommunenes inntekter vil føre til bevegelse langs etterspørselskurven. Vektoren X representerer andre karakteristikk som antas å påvirke tjenesteproduksjonen i skolen. Endringer i disse faktorene vil medføre skift i etterspørselskurven.

Etterspørselen er fallende, da økt lærerlønn vil redusere tilgjengelige midler i budsjettet for å ansette flere. Helningen på etterspørselskurven avhenger av elasticiteten på etterspørselen etter lærere. En elastisk etterspørsel vil respondere sterkere på endringer i lønnen enn en uelastisk etterspørsel. Grafisk kommer dette til uttrykk ved at en elastisk etterspørselskurve er slakere. Et skift i tilbudskurven vil medføre en stor endring i sysselsetting, men en liten endring i lønn.

Tilbudet av lærere gis ved antallet kvalifiserte personer som er villige til å jobbe i læreryrket med de gitte lønnsbetingelsene. Lærerne står ovenfor individuelle valg knyttet til å velge det alternative som gir høyest nytteverdi. En nyttefunksjon blir brukt til å måle individets preferanser gitt et sett med muligheter, og blir antatt å være en funksjon av personens inntekt i tillegg til en vektor som reflekterer individens preferanser (Boardman et al., 1982). I denne sammenhengen vil lærere kunne stå ovenfor avveininger mellom arbeid, fritid og lønn. For å beregne verdien av disse avveiningene over tid er det ønskelig å diskontere fremtidige verdier til dagens verdi. En arbeidstaker vil slutte i en jobb når nåverdien i den nåværende jobben blir lavere enn nåverdien assosiert med alternative jobbmuligheter. Arbeidsmarkedssituasjonen vil være relevant da den sier noe om sannsynligheten for å finne en annen jobb. Tilbudsfunksjonen kan formuleres som en funksjon av inntekt og arbeidsmarkedssituasjon, gitt ved

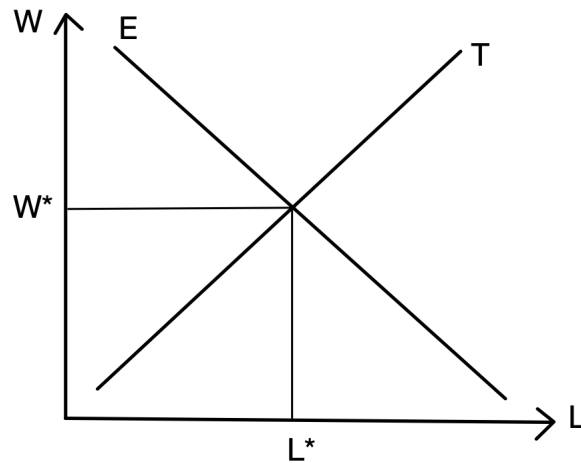
$$T = T(W, WR, U)$$

Her er W lærerlønn, WR er alternativlønn og U representerer arbeidsledighet. Alternativlønn viser til den potensielle lønnen en lærer kan opptjene ved å arbeide utenfor læreryrket. Endringer i lærerlønn gir bevegelse langs tilbudskruven, mens endringer i alternativlønn og arbeidsledighet gir skift i tilbudskurven.

Tilbudskurven er stigende, ved at økt lærerlønn vil gjøre det mer attraktivt å arbeide som lærer. Helningen på tilbudskruven avhenger av substitusjonselastisiteten mellom læreryrket og alternative yrker. Substitusjonselastisiteten vil si hvor enkelt det er å bytte fra læreryrket og over til et alternativt yrke. Et elastisk tilbud betyr at en liten endring i lærerlønn vil føre til en stor endring i lærertilbudet.

Når vi setter sammen etterspørsels- og tilbudskurven finner vi markedstilpasningen i skjæringspunktet mellom kurvene. Med en antagelse om fri konkurranse vil aktørene ta pris og lønn for gitt. Figur 1 illustrerer en initiell likevekt mellom tilbud og etterspørsel. X-aksen representerer antallet lærere med kompetanse, Y-aksen viser lærerlønn. Tilbudskurven, som er stigende, representerer tilgjengeligheten av lærere, mens etterspørselskurven er fallende og reflekterer arbeidsgivernes behov for lærere.

Figur 1: Lærerarbeidsmarkedet

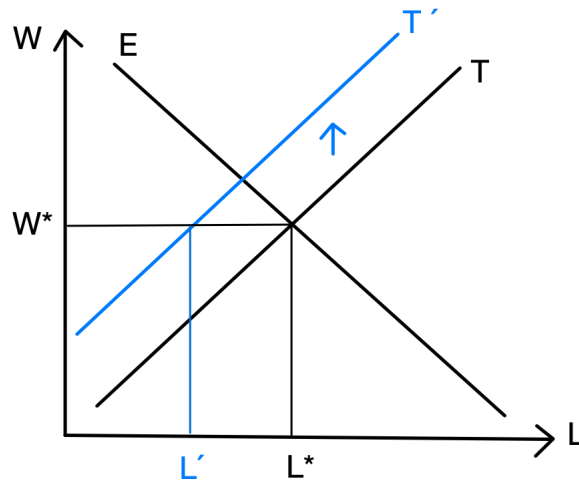


Skift i en av kurvene vil gi ulikvekt dersom lønnsbetingelsene er rigide og ikke justerer seg. For å gjenopprette likevekt er det nødvendig med lønnsjustering. Ubalansen i markedet vil derimot ikke ende i en situasjon hvor det er klasserom uten en lærer. Under forhold med overskuddsetterspørsmål etter lærere og lavere lønn enn markedslikevektslønnen, vil skolene reagere ved å ansette lærere uten formell kompetanse. På den måten vil mangelen på tilbudet av lærere ikke kunne ses i kvantum, men heller som en endring i kvalitetskarakteristikker i lærerstaben.

2.3 Betydningen av arbeidsledighet på lærermangel

I forbindelse med lærermangel, er arbeidsledigheten av interesse da den kan gi en pekepinn på om lærere ønsker å fortsette i læreryrket til tross for muligheten for å finne alternative jobber. En reduksjon i arbeidsledighetsraten vil gi et skift oppover i tilbudskurven, illustrert i figur 2. Hypotesen er at ved gitt lærerlønn vil det oppstå en mangel i tilbudet av lærere når arbeidsledigheten reduseres. Lavere arbeidsledighetsrate tilsier at det er færre tilgjengelige kandidater for stillingsutlystninger, noe som øker sannsynligheten for at en lærer kan finne en jobb utenfor skolesystemet. Som en konsekvens må arbeidsgivere øke lønningene for å tiltrekke seg attraktive arbeidstakere. Lærerlønnen må økes tilstrekkelig, slik at lærere med kompetanse vil oppleve læreryrket som like attraktivt som før reduksjonen i arbeidsledighetsraten. Økt lærerlønn fører til redusert budsjett, som igjen fører til at muligheten for å ansette flere lærere reduseres. Tilpasningen i markedet, som følger av skiftet i tilbudskurven, resulterer i en reduksjon av antall lærere representert ved L' . Avstanden mellom L^* og L' i figur 2 illustrerer lærermangelen.

Figur 2: Skift gitt redusert arbeidsledighet



Paneldatanaanlysen til Falch et al. (2009) brukte datamateriale på heltidsstillinger. De antar at lærermangel er synkende i arbeidsledighetsraten. I deres modell vil høyere arbeidsledighetsrate øke tilbudet av lærere ettersom mulighetene for å få en annen jobb enn lærer synker for gitt relativ lønn. Den forventede utbetalingen man får som lærer vil derfor øke i forhold til alternative stillinger. Det ble funnet et negativt forhold mellom lærermangel og regional arbeidsledighet. Effekten av arbeidsledighet var på omlag $-0,9$. Når det tolkes ved prosentpoengs endring i arbeidsledigheten divideres koeffisienten på ledighetsnivået. Gjennomsnittlig arbeidsledighetsrate for deres tidsperiode (1981-2002) var på $2,6\%$. Det gir at økning i arbeidsledigheten på ett prosentpoeng fører til en reduksjon i lærermangel på nesten $0,4$ prosentpoeng. Lærermangel i analysen er gitt som andelen heltidsansatte lærere uten formell kvalifikasjon, (Falch et al., 2009).

Fraenkel har gjort tilsvarende studier på det amerikanske arbeidsmarkedet. Analysen undersøkte hvordan tilstanden i arbeidsmarkedet ved ansettelsestidspunktet har påvirkning på tilfredsheten med yrket. Hun finner at lærere som ansettes når arbeidsmarkedet er svakt har høyere sannsynlighet for å uttrykke utilfredshet med yrket, men at deres sannsynlighet for å forbli i yrket ikke er lavere (Fraenkel, 2022). Virkningen av tilstanden på arbeidsmarkedet i forhold til å forlate læreryrket ble også studert av Falch og Strøm i 2005. De fant at regional arbeidsledighet reduserer sannsynligheten for å forlate læreryrket (Falch & Strøm, 2005). I denne analysen ble det i tillegg undersøkt hvilke ikke-økonomiske forhold ved skolene som kunne påvirke andelen lærere som sluttet i yrket. Det ble funnet at lærere tenderer til å forlate skoler som har en høy andel elever med minoritetsbakgrunn og spesielle behov. I tillegg har lærersammensetningen og skolestørrelsen innvirkning på sannsynligheten for å slutte.

Goldhaber og Theobald undersøkte, i sin artikkel fra 2022, sammenhengen mellom lærerslitasje og mobilitet på offentlige skoler i Washington med arbeidsledighet, over en periode på 35 år. De skilte mellom i) lærere som forlot læreryrket, ii) lærere som gikk over i administrative stillinger innenfor det offentlige skoleverket og iii) lærere som forlot en lærerstilling til fordel for en tilsvarende lærerstilling på en annen offentlig skole. Lærerslitasje omfatter det to første kategoriene, mens mobiliteten viser til lærere som forblir i lærerstillingen. De fant at det var relativt lav differanse i de to ratene opp mot arbeidsledighet. Videre fant de at lærerslitasje er motsyklisk, ved at det er mindre sannsynlig for lærere å forlate en stilling som lærer når arbeidsmarkedet er stramt (Goldhaber & Theobald, 2022).

2.4 Utbredelse av lærermangel

Mangel på kvalifiserte lærere vil ikke ende i en situasjon hvor et klasserom står uten en lærer. Imidlertid er spørsmålet hvorvidt læreren i klasserommet besitter den nødvendige kompetansen eller ikke. Det kan være regionale forskjeller i graden av lærermangel, og faktorer som kommunestørrelse og geografisk plassering kan påvirke tilgangen til kvalifiserte lærere. Større kommuner er oftere omringet av mindre, nærliggende kommuner, slik at det reelle arbeidsmarkedet omfatter flere kommuner. Tilgangen på arbeidskraft vil være høyere i kommuner med større befolkningstall. Samtidig vil konkurransen om arbeidskraften fra andre sektorer være større, enn i en mindre kommuner med færre alternative jobbmuligheter.

Hvordan problemstillingen med lærermangel innenfor læreryrket utvikler seg kan også sees i sammenheng med antall lærerstudenter. Et lavt antall lærerstudenter indikerer en reduksjon i antall kvalifiserte lærere som kommer ut i lærerarbeidsmarkedet. Lønnsvilkårene innenfor læreryrket kan også påvirke beslutningen om å gå inn i, og forbli i yrket. I en undersøkelse av utdanningsvalg i Storbritannia, ble det funnet at lønnen i læreryrket sammenlignet med andre yrker har en signifikant påvirkning på sannsynligheten for at studenter velger å ta lærerutdanning (Chevalier et. al., 2007). Lønnen vil da, etter endt studie, ha en mindre avgjørende rolle i beslutningen om å fortsette i læreryrket eller ikke i følge Chevalier.

Dagens tendens til lærermangel kan knyttes til kvinners økte deltagelse i yrkeslivet. Læreryrket har tradisjonelt vært det primære yrket for utdannede kvinner. Bacolod undersøkte effekten av økte alternativer i arbeidsmarkedet for kvinner opp mot reduksjonen av lærerkvalitet. Han fant at flere muligheter resulterte i at færre høyt kvalifiserte kvinner valgte læreryrket (Bacolod, 2007).

3 Institusjonelt rammeverk

Dette kapittelet orienterer om skolens rammebetingelser og opplæringslovens bestemmelser for rekruttering. Videre orienteres det om lønnsdannelsen for lærere og om arbeidsledighet.

3.1 Skolenes rammebetingelser

Grunnskoleordningen i Norge er en obligatorisk 10-årig skolegang for alle barn fra og med 6 år. I 1997 ble 6-års reformen innført, som reduserte alder for skolestart fra 7 år til 6 år. Grunnskolene reguleres av bestemmelsene i opplæringsloven. De offentlige grunnskolene er eid av kommunen der de er geografisk lokalisert. Kommunene er pliktet til å tilby alle barn og unge grunnskoleopplæring. Finansiering av skolene skjer i hovedsak gjennom den generelle rammefinansieringen av kommunesektoren, hvor kommunene selv får prioritere ressursbruken. Kun 5-6 prosent av utgiftene til grunnskolen finansieres ved øremerkede tilskudd (St.meld. nr 33 (2002-2003), s.9).

Myndighetene fastsetter et minimumskrav for antall elever per ansatt i klasserommet. Dermed vil skolens størrelse i stor grad bestemmes av antallet barn som er tilknyttet skolen. Skolens etterspørsel av ansatte kan ses på som uavhengig av tilstanden i arbeidsmarkedet. Samtidig er etterspørselen begrenset av budsjettet i rammefinansieringen. Opplæringsloven setter bestemmelser til rekruttering ved at det er kvalifiserte kandidater som skal ansatte.

Opplæringsloven §10-1: Krav om kompetanse ved tilsetjing av undervisningspersonell.

Den som skal tilsetjast i undervisningsstilling i grunnskolen og i den videregående skolen, skal ha relevant fagleg og pedagogisk kompetanse. Departementet gir nærmare forskrifter om krav til utdanning og praksis for den som skal tilsetjast i undervisningsstillinger på ulike årstrinn og i ulike skoleslag (Opplæringsloven, 1998, §10-1).

Ukvalifiserte kandidater kan kun ansettes dersom det ikke er kvalifiserte søkere, og kun i midlertidig stilling. Fagforeningen for lærere er involvert i intervjuprosessen, og vil bli informert dersom det ikke finnes noen kvalifiserte søkere til stillingsutlystningen.

Opplæringsloven §10-6 : Mellombels tilsetjing.

Dersom det ikkje er søkjarar som fyller kompetansekrava for tilsetjing i §10-1, kan andre tilsetjast mellombels (Opplæringsloven, 1998, §10-6).

Lærermangel i vår oppgave defineres som andel lærere i undervisningsstillinger som ikke har godkjent kompetanse. Ansettelse av ukvalifiserte kandidater er da et resultat av at ingen kvalifiserte kandidater har søkt på utlysningen.

3.2 Lønnsdannelse

Lønnsvekst i offentlig sektor reguleres av frontfagsmodellen som setter gjennomsnitts lønnsvekst i kommunene. Modellen går ut på at det er den konkurranseutsatte sektoren som først forhandler lønn. Lønnsdannelse for ansatte ved offentlige grunnskoler gjøres både ved sentrale og lokale lønnsforhandlinger. Det er tre fagforbund i Norge for lærere. Norsk Lektorlag dekker lektorutdannede, Skolens Landsforbund er hovedsaklig for yrkesfaglærere, mens Utdanningsforbundet dekker pedagogisk personale med godkjent utdanning i utdanningssektoren over alle utdanningsnivå. De ulike forbundene blir i lønnsforhandlingene representert av hovedsammenslutningene ved henholdsvis Akademikerne, LO og Unio.

Kommunesektorens organisasjon (KS) er en sammenslutning av kommuner og bedrifter med kommunalt eierskap. Under de sentrale lønnsforhandlingene opptrer KS som arbeidsgiver og forhandler med hvert forbund i forhandlingssammenslutningene Akademikerne Kommune, LO Kommune og Union.

Den endelige lønnen bestemmes ved de lokale forhandlingene, gitt av den sentralt bestemte "potten". Denne potten gir kommunene en fleksibilitet til å ta lokale avveininger på bruk av ressursene. En kommune som sliter med å rekruttere lærere til skolene har dermed mulighet til å tilby høyere lønn for å kompensere for mangelen. Lokal lønnsrespons kan føre til mindre variasjon i lærermangel. Nytilsatte lærere kan bli tilbudt høyere lønn enn lønnen i tariffavtaler. I vår oppgave vil vi bruke tidsfaste effekter for å kontrollere for nasjonale endringer i lønn.

KS tok over ansvaret for lærere i 2004. Før den tid var lønnsforhandlingene sentralisert, og uten lokalt handlingsrom. Lønnsfastsettingen baserte seg utelukkende på utdanningsnivå og ansiennitet.

Minimums lønnsats som gjelder for de ulike utdanningsnivåene og erfaringene innenfor yrket er fastsatt i den sentrale tariffavtalen. Fra og med høsten 2017 ble grunnskolelærerutdanningen omgjort til en femårig utdanning. Dette gav en forlengelse med ett år på grunnskoleutdanningen, og i kompetanselønnssystemet gir det uttelling som lektor (Utdanningsforbundet, 2021). Det første studentkullet med femårig grunnskolelærerutdanning, med normert studieprogresjon, var da ferdig utdannet til sommeren 2022.

Kompetansen til lærere vi ser på i vår oppgave er begrenset til perioden før det var nyutdannede med den nye utdanningslengden.

3.3 Arbeidsledighet

Arbeidsledighet er definert som andelen av arbeidsstyrken som er tilgjengelig for arbeidsmarkedet. Dette utelukker blant annet de som er uføre og dermed ikke yrkesaktive. Det er forskjellige måter å måle arbeidsledigheten på, og i Norge brukes blant annet arbeidskraftsundersøkelsen (AKU) til Statistisk sentralbyrå. En person blir definert som AKU-ledig dersom de både er uten arbeid, oppgir at de aktivt har søkt arbeid i løpet av de siste fire ukene og kan begynne i arbeid i løpet av de neste to ukene (Stokke et al., 2023). Et annet mål på arbeidsledigheten i Norge er basert på de som registrerer seg som arbeidssøkere hos NAV.

Videre skilles det mellom ulike typer arbeidsledighet. Friksjonsarbeidsledighet oppstår når det tar tid å gå fra en jobb til en annen. Sesongarbeidsledighet er knyttet til arbeid som kun foregår i enkelte årstider. Arbeidsledighet som påvirkes av tilstanden på økonomien kalles konjunkturarbeidsledighet. Strukturarbeidsledighet oppstår dersom arbeidsstyrken ikke er tilpasset for nye forhold eller konkurranse.

Arbeidsledighetstall brukes som en indikator på økonomiens tilstand. Den er en sentral variabel i makroøkonomiske analyser og som en konjunkturindikator (Sparrman, 2012). Lav arbeidsledighet knyttes til at det er høy aktivitet i økonomien, mens høy arbeidsledighet knyttes til lav aktivitet. Dette skyldes at det da er relativt mange som står utenfor arbeidsmarkedet og ikke bidrar til verdiskapning.

4 Datamaterialet

Dette kapittelet beskriver datamaterialet som er brukt i analysen. Datakildene blir presentert og vi definerer variabelen for målet på lærermangel. Videre beskrives fokusvariablene for arbeidsledighet. Til slutt presenteres kontrollvariablene som inngår i analysen.

4.1 Datakilder

Datamaterialet med informasjon om grunnskolene er hentet fra Grunnskolens Informasjonssystem (GSI) som administreres av Utdanningsdirektoratet. Statistikken rapporteres av skolene selv gjennom rektorene per 1 oktober for gjeldende skoleår. Innsamlingen av data startet i 1992, og alle registrerte år ligger offentlig tilgjengelig. Blant opplysningene som registreres er antall årsverk, ansatte og elever. Det rapporteres også informasjon om hvor stor andel av undervisningen som går til ordinær undervisning og til spesialundervisning.

Opplysningene som registreres i GSI er knyttet til den planlagte og forventede ressursbruken for det kommende året. Dette innebærer at vikarer på timesbasis ikke inkluderes. Opplæringsloven viser til at en kvalifisert lærer må ha relevant faglig og pedagogisk kompetanse. Rektorer kan med denne definisjonen påvirke hva ”relevant” betyr. Senter for økonomisk forskning har gjennomført kvalitative undersøkelser av rapporteringen i GSI. Disse undersøkelsene viser at vikarer som rapporteres i GSI hovedsaklig er vikarer for langtidsfravær som allerede er kjent ved registreringstidspunktet (Senter for økonomisk forskning, 2021, s.9). Videre vises det at disse vikarene i større grad har godkjent kompetanse sammenlignet med lærere på korttidskontrakter. Konsekvensen av dette kan være at dataene fra GSI gir en underestimert av den faktiske lærermangel i norske skoler.

Det har vært flere utfordringer knyttet til å benytte GSI som datakilde. For det første har innhenting av opplysninger vært krevende, da vi har måttet manuelt laste ned relevante variabler for hvert år og deretter kombinere dem til ett datasett. For det andre har GSI tilbakeført endringer som er gjort på et senere tidspunkt til å også gjelde før endringen inntraff. Dette kommer til uttrykk ved at kommuner og skoler som ble påvirket av kommunesammenslåingen i 2020 har blitt registrert med det nye kommunenummeret også for årene før sammenslåingen. Utfordringen ved dette har da vært at kommunenummerene ikke har fulgt ett og samme system. For å håndtere dette har vi justert kommunenummerene tilbake til det opprinnelige nummeret før

sammenslåingen fant sted. En tredje utfordring er knyttet til at datamaterialet har bestått av flere duplikater. Det betyr at samme enhet har blitt registrert flere ganger. Problemet med duplikater har i stor grad vært relevant for kombinerte skoler, hvor informasjonen om barneskole- og ungdomsskoletrinn har blitt registrert hver for seg, men med samme navn på skoleenheten. Slike duplikater medfører et problem under estimering, og vi har derfor valgt å summere de to enhetene med samme navn til én skole. En siste utfordring involverer manglende observasjoner i datasettet. Når observasjoner ikke er rapportert i GSI har disse verdiene blitt satt til 0. Eksempelvis begynte ikke rapporteringen av antall lærere i skolen før 2004. I GSI er det derfor rapportert 0 lærere i norske skoler før 2004, noe som ikke er representativt for virkeligheten. Vi har derfor endret observasjonene til ”manglende” i vårt datasett for de aktuelle tilfellene, for å unngå skjevhet mot null i estimatene.

Vi har konstruert et paneldatasett som består av 3 938 skoler, med elever på både barne- og ungdomstrinn, fordelt på 436 kommuner over hele fastlands-Norge. Tidsperioden strekker seg fra og med 1992 til og med 2019. Paneldatasettet er ubalansert, som vil si at det ikke er like mange tidsobservasjoner for alle enheter. Det er fordi noen skoler har endret seg gjennom årene, enten ved å bli opprettet, lagt ned, splittet eller slått sammen.

Kommunedata og arbeidsledighet på kommunenivå er hentet fra Fiva, Halse og Natvik sitt ”Local Government Dataset” (Fiva et.al., 2020). Refererer videre til ”FIVA” for å henvise til dette datasettet gjennom hele oppgaven. Materialet er samlet fra Kommunedatabasen (NSD) og fra Statistikkbanken til Statistisk sentralbyrå (SSB). Paneldatasettet spenner seg fra 1972 til 2019 og inneholder informasjon om kommunale strukturer, demografier, politiske valg og kommunal finanspolitikk. Vi vil benytte oss av de kommunale arbeidsledighetsratene i tillegg til data på befolkningstall og aldersgrupper. Datasettet inneholder også en variabel som indikerer hvilke kommuner som til sammen utgjør en større arbeidsmarkedsregion. Arbeidsmarkedsregioner er et mellomliggende nivå mellom kommuner og fylker, og det er ingen arbeidsmarkedsregioner som krysser fylkesgrenser. Denne inndelingen ble laget i 2000, hvor kommunene er delt inn i 90 ulike regioner. Hensikten bak arbeidsmarkedsregionen er å ha et enhetsnivå som tar hensyn til mobiliteten i arbeidsmarkedet og økonomiske samarbeid.

I tillegg til arbeidsledighetsraten fra FIVA har vi også hentet kommunal arbeidsledighetsrate fra NAV. Raten er oppgitt for hver måned, og starter i 1995 og varer frem til 2021. Datasettet inkluderer også en variabel som registrerer antallet arbeidsledige i kommunene hver måned.

Vi benytter oss av paneldata, som betyr at vi følger de samme enhetene over flere år. Vårt paneldatasett som kombinerer GSI, FIVA og NAV spenner seg over tids-

perioden 1995 - 2019. Vi har valgt å avgrense datamaterialet på bakgrunn av flere hensyn. Den nedre grensen er satt til 1995 fordi arbeidsledighetsdata fra NAV kun er tilgjengelig fra dette året, mens den øvre er av hensyn til sammenslåinger av kommuner. Kommunereformen som trådte i kraft i 2020 førte til at 119 kommuner ble til 47 nye kommuner. Vi har derfor valgt å kun la datasettet gå opp til 2019 for å forenkle analysen. Det er imidlertid flere tilfeller av kommuner som har slått seg sammen i løpet av tidsperioden 1995-2019. Eksempelvis ble Nord- og Sør-Trøndelag i 2018 slått sammen til et felles Trøndelag fylke. Sammenslåingen førte til nye kommunenummer, og for å kunne matche arbeidsledigheten på kommunenivå korrekt har vi brukt samsvarende kommunenavn for de tre ulike datasettene når vi har kombinert dem.

4.2 Lærermangel

Lærermangel på skolenivå definerer vi som andelen årsverk uten kompetanse i forhold totalt antall årsverk ved skolen. Variabelen fanger da opp hvor stor andel av årsverkene ved skolen som ikke har tilstrekkelig kompetanse etter opplæringsloven. Årsverk viser til arbeidet som utføres i løpet av ett år for en arbeidstaker som er i en heltidsstilling. Det inkluderer arbeidstakere i mindre stillingsprosenter. Denne variabelen tolkes som lærermangel da ansettelse av ukvalifiserte er et resultat av at kvalifiserte kandidater ikke har søkt på utlysningen. Vi ønsker å ha det på prosent og multipliserer derfor med 100. Fotskrift s viser til hver unike skole, mens t viser til årene innenfor tidsperioden 1995 - 2019.

$$Lærermangel_{st} = \frac{\text{årsverk uten kompetanse}_{st}}{\text{årsverk med kompetanse}_{st} + \text{årsverk uten kompetanse}_{st}} \cdot 100$$

For å finne lærermangel på kommunenivå bruker vi gjennomsnittlig lærermangel over alle skolene innenfor én kommune. Fotskrift k viser til kommunen som den enkelte skole ligger i.

$$\begin{aligned} Lærermangel_{kt} &= \frac{\sum \text{årsverk uten kompetanse}_{skt}}{\sum (\text{årsverk med kompetanse}_{skt} + \text{årsverk uten kompetanse}_{skt})} \\ &= \frac{\text{ukvalifiserte}_{kt}}{\text{totalt}_{kt}} \cdot 100 \end{aligned}$$

I datasettet har vi 77 484 observasjoner fra 3 938 skoler i 436 kommuner. På kommunenivå har vi 10 760 observasjoner. Vi skiller i denne oppgaven mellom variablene for lærermangel på skole- og kommunenivå ved å skrive henholdsvis $Lærermangel_s$ og $Lærermangel_k$. Tabell 1 viser at gjennomsnittlig lærermangel over tidsperioden

på både skole- og kommunenivå ligger på omtrent 4%. Det er skoler og kommuner som i løpet av tidsperioden registrerer at ingen årsverk gjennomføres av ukvalifiserte lærere. Maksimalt har det blitt registrert at halvparten av årsverkene ved en skole har vært av ukvalifiserte lærere. På kommunenivå er den største graden av lærermangel på 41%.

Overordnet standardavvik refererer til den samlede spredningen av variabelen på tvers av alle enheter og tidsperioden. Den kan videre dekomponeres i to komponenter, nærmere bestemt variasjonen mellom og variasjonen innenfor. Variasjonen mellom viser til variasjonen mellom enhetene som er konstant over tid, mens innenfor variasjon fanger opp endringer over tid innenfor hver enhet.

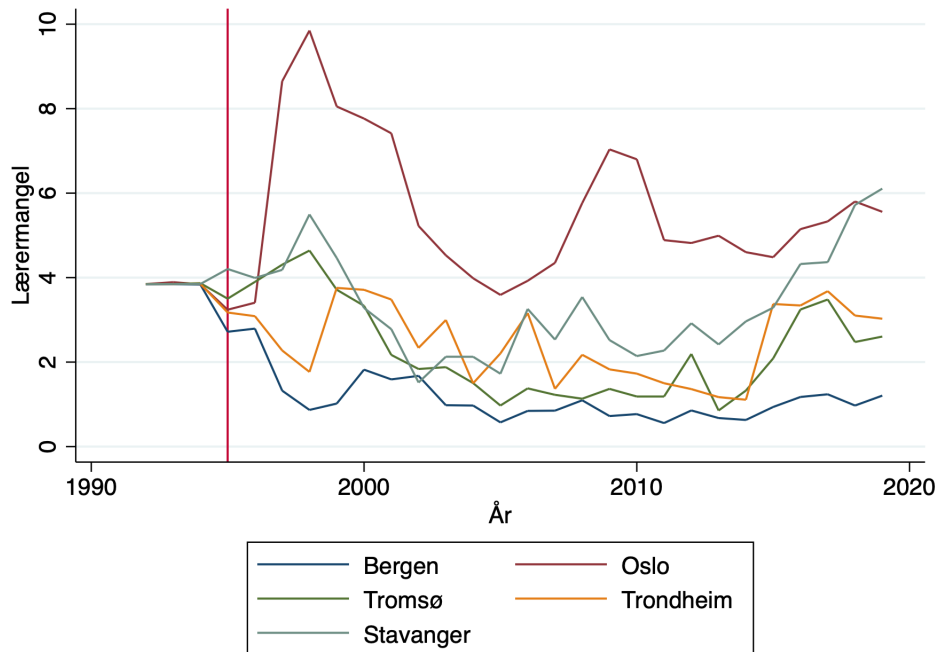
Tabell 1 viser at det er større variasjon i lærermangel innenfor hver skoleenhet enn det er i mellom dem. Dette gir en indikasjon på at det ikke er en symmetrisk fordeling på lærermangel. Tilsvarende på kommunenivå er det større variasjon innenfor kommunene enn mellom.

Tabell 1: Deskriptiv statistikk på lærermangel

Variabelnavn	Variasjon	Gjesnitt	Standardavvik	Min	Maks	Observasjoner
Lærermangel.s	Overordnet	4,044	6,415	0	50	N = 77 484
	Mellom		4,575	0	50	n = 3 938
	Innenfor		5,139	-35,51	51,66	T-bar = 19,68
Lærermangel.k	Overordnet	4,346	4,335	0	41,33	N = 10 760
	Mellom		2,732	0,577	16,15	n = 436
	Innenfor		3,370	-11,80	31,96	T-bar = 24,68

Hvordan utviklingen av lærermangel har vært i utvalgte større byer i Norge fremgår i figur 3, hvor den røde linjen markerer året 1995. Det er variasjon gjennom hele perioden, både innenfor hver by og mellom byene. Vi ser at Oslo skiller seg ut med klart høyest grad av lærermangel, mens Bergen ligger på et lavere nivå over hele perioden. Lærermangelraten er brattere ved slutten av tidsperioden for Trondheim, Stavanger og Tromsø. Avslutningsvis ligger Stavanger høyere enn Oslo.

Figur 3: Trenden i lærermangel i utvalgte byer



Det eksisterer en bekymring angående påliteligheten av rapporteringen i GSI for de første årene. Figur 3 illustrerer dette ved at lærermangel i årene før 1995 ligger på et jevnt og uendret nivå, og at alle byene viser tilsvarende grad av lærermangel, som en samlet strek. I 1995 ser vi at Oslo tar et klart hopp fra under 4% lærermangel opp til over 8% i 1997. En mulig årsak kan være at samme tall har blitt registrert over flere år. Vi utelater de første upålitelige årene ved at paneldataanalysen begynner i 1995.

Falch et al. brukte i sin analyse fra 2009 antall heltidsstillinger for å måle lærermangel. Vi er interessert i lærermangel fra elevperspektiv og da er årsverk et riktigere mål. Ved å benytte årsverk får vi inkludert deltidsansatte lærere uten formell kompetanse, noe som gir et mer realistisk bilde av den faktiske lærermangelen eleven møter i klasserommet. Derfor har vi et bedre mål enn Falch et al. som bruker heltidsstillinger.

4.3 Data på arbeidsledighet

Vi benytter to variabler som begge måler arbeidsledighet på kommunenivå. NAV sitt mål på arbeidsledighet baserer seg på andelen uten arbeid fra arbeidsstyrken. NAV rapporterer månedlig arbeidsledighetsraten og vi har i analysen valgt raten som er registrert i januar for inneværende år. Dette gir oss muligheten til å undersøke hvordan

arbeidsledigheten påvirker lærermangelen som blir registrert i oktober samme år.

$$\text{Arbeidsledighet_NAV}_{kt} = \frac{\text{arbeidsledige}_{kt}}{\text{arbeidsstyrken}_{kt}} \cdot 100$$

Variabelen på arbeidsledighet fra FIVA tar utgangspunkt i hele befolkningen i arbeidsdyktig alder, altså mellom 16 og 65 år, og hvor mange av disse som ikke er i arbeid. Nevneren er større og vil derfor konsekvent rapportere en noe lavere arbeidsledighetsrate enn NAV. I datasettet er arbeidsledighetsraten gitt som et årsgjennomsnitt.

$$\text{Arbeidsledighet_FIVA}_{kt} = \frac{\sum \text{arbeidsledige}_{kt}}{\sum \text{arbeidsdyktig alder}_{kt}} \cdot 100$$

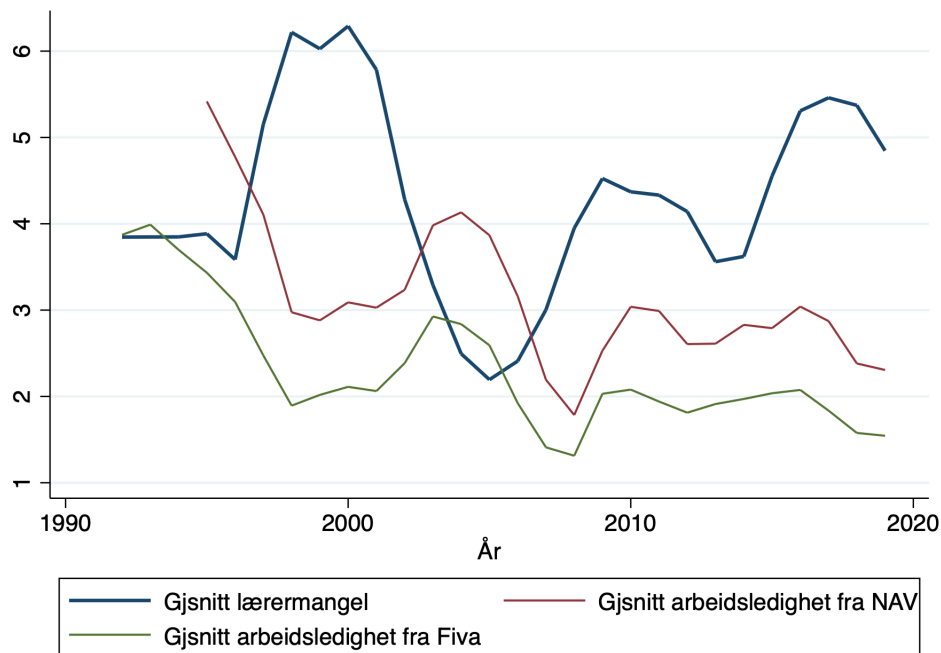
Tabell 2 viser at gjennomsnittlig arbeidsledighetsrate i januar måned fra NAV ligger på 3,15%. Årsgjennomsnittlig arbeidsledighetsrate i kommunene i FIVA ligger på 2,23%. NAV registrerer en høyere maksimal arbeidsledighetsverdi på 24,3%, mens maksimal årsgjennomsnittlig arbeidsledighet i FIVA er på nesten 13%. Denne forskjellen er forventet på grunnlag av hvordan variablene er definert. Standardavviket er også høyere for NAV, noe som kan forklares av at variasjonen i arbeidsledigheten innenfor en kommune er høyere for NAV enn FIVA.

Tabell 2: Deskriptiv statistikk på arbeidsledighet

Variabelnavn	Variasjon	Gjennomsnitt	Standardavvik	Min	Maks	Observasjoner
Arbeidsledighet_NAV	Overordnet	3,154	1,685	0	24,3	N = 10 671
	Mellom		1,125	1,080	8,604	n = 436
	Innenfor		1,261	-1,502	19,25	T-bar = 24,47
Arbeidsledighet_FIVA	Overordnet	2,214	1,054	0	12,94	N = 10 722
	Mellom		0,796	0,872	6,075	n = 436
	Innenfor		0,703	-1,287	10,19	T-bar = 24,59

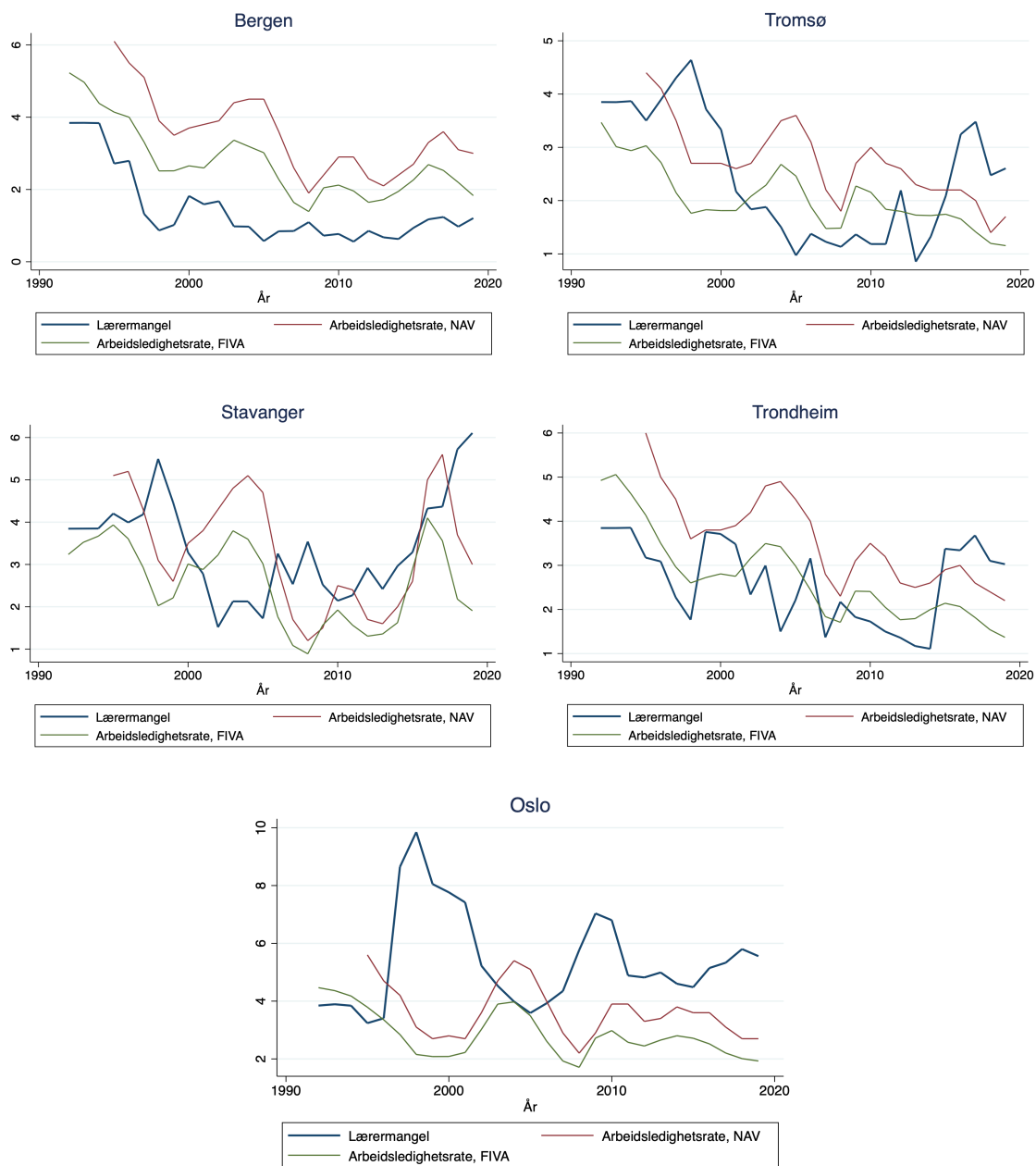
På aggregert nivå viser figur 4 tendenser til et negativt forhold mellom arbeidsledighet og lærermangel, gitt ved den tykkere linjen. Når arbeidsledigheten synker så stiger lærermangel. Denne tendensen ses tydeligst for årene frem til 2010. Figuren illustrerer også at linjene for arbeidsledighet fra de to variablene har like bevegelser gjennomgående, men på ulike nivå. Arbeidsledigheten fra FIVA ligger på et lavere nivå enn raten fra NAV, men begge er relativt stabile fra rundt 2010. Gjennomsnittlig aggregert lærermangel er på sitt laveste nivå rundt år 2005. Fra rundt 2010 er bevegelsene mer samsvarende mellom lærermangelraten og arbeidsledighetsratene, ved at lærermangel stiger når arbeidsledighetsraten stiger.

Figur 4: Trenden i de aggregerte arbeidsledighetsratene og aggregert lærermangel



Figur 5 viser lærermangelen i de større norske byene isolert opp mot arbeidsledighetsratene i kommunen. I Bergen og i Trondheim er tendensen at bevegelsene mellom lærermangel og arbeidsledighet samsvarer. Stavanger har større variasjoner både i arbeidsledighet og lærermangel. De tydelige stigningene i arbeidsledigheten i Stavanger rundt 2008 og 2014 kan være relatert til utviklingen og sjokk i petroleumsvirksomhet. Vi ser videre at når arbeidsledigheten i Stavanger er høy den første gangen, så synker lærermangelen. Derimot ser vi at både arbeidsledigheten og lærermangelen er stigende ved slutten av perioden. Tendensen i Tromsø og Oslo er at arbeidsledigheten og lærermangelen har et negativt forhold.

Figur 5: Trenden i lærermangel i utvalgte byer opp mot arbeidsledighetsratene



4.4 Kontrollvariabler

I tillegg til arbeidsledighet inkluderer vi en rekke kontrollvariabler i analysen. Dette er for å kontrollere for ulike forhold som også kan tenkes å ha påvirkning på graden av lærermangel. Tabell 3 viser variabelnavn og forklaring av dem. Kontrollvariablene er knyttet til forhold ved skolene og i kommunen som kan være relevant for lærermangel.

Tabell 3: Variabelnavn og forklaringer

Variabelnavn	Forklaring
Elever0110_sum	Antall elever ved grunnskolene fra 1.klasse til 10.klasse
Lærere_sum	Antall lærere ved skolen
Lærerandel_kvinner	Andelen av lærerne som er kvinner
Spesandelu	Andelen av årsverk til undervisning som går til spesialundervisning ved grunnskolene fra 1.klasse til 10.klasse
Befolkningstall	Befolkningstall i kommunen
Arbeidsalder	Andelen av befolkningen som er i alderen mellom 15 og 66
Barn	Andelen av befolkningen som er yngre enn arbeidsrelevant alder
Eldre	Andelen av befolkningen som er eldre enn arbeidsrelevant alder
Barneskole	Dummyvariabel for om skolen er en barneskole
Ungdomsskole	Dummyvariabel for om skolen er en ungdomsskole
Kombinert	Dummyvariabel for om skolen er en barne- og ungdomsskole
Barneskole_andel	Andelen barneskoler i kommunen
Ungdomsskole_andel	Andelen ungdomsskoler i kommunen
Kombinertskole_andel	Andelen skoler i kommunen som er barne- og ungdomsskoler

Antall elever og lærere gir en indikasjon på størrelsen på skolen, og brukes til å kontrollere for om skolestørrelsen påvirker graden av lærermangel. Tabell 4 viser at i gjennomsnitt er det 194 elever på skolene. Variasjonen mellom minste og største elevtall er relativt stor, da minste verdi er på én elev mens største skole har 1 049 elever. Det er færre observasjoner på antall lærere og kjønnsfordelingen mellom lærere

fordi det ikke ble registrert i GSI før i 2004. Lærerstaben ligger i gjennomsnitt på omlag 21 lærere, hvor andelen kvinnelige lærere er på 74%. Vi inkluderer en variabel med andelen av totale årsverk som går til spesialundervisning, ettersom dette kan gi en indikasjon på hvor belastende eller utfordrende en arbeidshverdag som lærer kan være på en skole. Andelen årsverk som går til spesialundervisning på skolene ligger i gjennomsnitt på 20%. Vi har et høyt standardavvik på 13,97, noe som viser til at det er stor variasjon i graden av spesialundervisning mellom skolene.

Kommunekaraktistikker er gitt ved befolkningstall og aldersfordelingen. Tabell 4 viser at vi har 10 760 antall observasjoner. Gjennomsnittlig befolkningstall i kommunene er 57 463 personer. Minste befolkningstall i en kommune er på 196 personer, mens største er på 681 071. Det viser at det er stor variasjon i befolkningstall i de ulike kommunene. Den største aldersgruppen i befolkningen er de i arbeidsrelevant alder, satt til 16-65 år, som utgjør 62,8%. Andelen i arbeidsrelevant alder sier noe om tilgangen på potensiell arbeidskraft. Variabelen for barn viser til andelen i befolkningen som er yngre enn arbeidsrelevant alder. I gjennomsnitt utgjør barn 20% av befolkningen. Eldre, som viser til andelen i befolkningen som er eldre enn arbeidsrelevant alder, utgjør 17% av befolkningen i gjennomsnitt. Kommunene er ansvarlige for å tilby tjenester for både unge og eldre. En stor andel eldre i kommunen gjør det nødvendig med prioritering av ressurser mellom ulike kommunale tjenester i budsjettet. Det er stor spredning i størrelsen på befolkningen i de norske kommunene, noe vi kontrollerer for ved å inkludere en variabel som tar for seg befolkningsantall.

Tabell 4: Deskriptiv statistikk på kontrollvariabler

Variabelnavn	Gjennomsnitt	Standardavvik	Min	Maks	Observasjoner
Elever0110_sum	194	150,2	1	1 049	77 529
Lærere_sum	21,89	13,67	1	171	47 640
Lærerandel_kvinner	74,07	14,32	0	100	47 640
Spesandelu	20,08	13,97	0	100	77 369
Befolkningstall	11 080	33 121	196	681 071	10 760
Barn	20,15	2,568	11,16	29,88	10 760
Arbeidsalder	62,85	2,489	52,99	70,92	10 760
Eldre	17,00	3,738	6,810	31,61	10 760

Dummyvariabler er numeriske variabler som enten tar verdien 0 eller 1 for å skille mellom forskjellige grupper i datasettet. Verdien 1 viser til at en hendelse inntreffer i den aktuelle variabelen, og vil ellers ta verdien 0 om hendelsen ikke inntreffer. For å kontrollere for om det er forskjell på graden av lærermangel på de ulike skoletypene lager vi dummyvariabler. Variablene for barne- og ungdomsskole genereres henholdsvis ut i fra om det kun er registrert elever i GSI på årstrinnene 1.-7. eller 8.-10.. Kombinertskoler viser til skoler som både har barne- og ungdomsskoletrinnene på skolen. Dummyvariabelen for kombinertskoler opptrer som referansekategori i analysen ved at den er utelatt fra modellen.

Tabell 5 viser at 59,1% av skolene i datasettet er barneskoler, mens 16% er ungdomsskoler. De resterende skolene er kombinerte skoler, og utgjør med det 24,9%.

Tabell 5: Deskriptiv statistikk på dummyvariabler for skoler

Variabelnavn	Gjennomsnitt	Standardavvik
Barneskole	0,591	0,492
Ungdomsskole	0,160	0,367
Kombinertskoler	0,249	0,432
Antall observasjoner		77 529

På kommunenivå genererer vi variabler for andelen av de ulike skoletypene i kommunen. I gjennomsnitt utgjør barneskoler 51% av skolene i kommunene. Ungdomsskoler utgjør 12% mens kombinerte skoler utgjør 38%. I noen kommuner er det kun barneskoler eller kombinerte skoler, mens ungdomsskoler utgjør maksimalt halvparten av skolene i kommunene.

Tabell 6: Deskriptiv statistikk på andelen av skoletyper i kommunene

Variabelnavn	Gjennomsnitt	Standardavvik	Min	Maks
Barneskole_andel	50,68	27,04	0	100*
Ungdomsskole_andel	11,71	12,31	0	50
Kombinertskole_andel	37,61	34,43	0	100
Antall observasjoner	10 760			

*Kommuner registrert med kun barneskoler er Rømskog, Hægebostad og Mosvik.

5 Metode

Analysen blir gjort på et paneldatasett som varierer både over tverrsnitt og over tid. Siden vi har data nede på skolenivå, kan vi ved å benytte paneldata lettere kontrollere for andre faktorer enn arbeidsledighet som påvirker lærermangel. Paneldata gjør det også mulig å kontrollere for effekter av makroøkonomiske forhold som for eksempel lærerlønn og uobserverte faktorer som kan påvirke den avhengige variabelen.

I dette kapitlet legger vi frem modellspesifikasjoner og estimeringsmetoder. Først presenteres modellspesifikasjonene som benyttes i analysen. Deretter presenteres ulike estimeringsmetoder for paneldata. Til slutt presenteres hvilke tester som er hensiktsmessige å benytte for å vurdere hvilken estimeringsmetode vi skal bruke. Kapitlet tar utgangspunkt i læreboken Wooldrige (2019), kapittel 13 og 14, og Verbeek (2017), kapittel 10.

5.1 Modellspesifikasjoner

I analysen benyttes ulike modellspesifikasjoner. Den avhengige variabelen er, som tidligere spesifisert, andelen årsverk som ikke oppfyller kravet om formell kompetanse. Fokusvariablene for arbeidsledighet er i likhet med Falch et al. log-transformert og blir illustrert med ln i variabelnavnet. Vi tolker koeffisienten som β_1 prosentpoeng endring i lærermangel gitt én prosent endring i arbeidsledighetsraten. Arbeidsledigheten er målt på kommunenivå, så hver skole innenfor samme kommune vil i det samme året ha lik arbeidsledighetsrate. Vi skiller mellom lærermangel på skolenivå og kommunenivå for å ta hensyn til faktorer som kan påvirke situasjonen på de ulike nivåene.

5.1.1 Modellspesifikasjoner på skolenivå

Hovedmodellen på skolenivå spesifiseres som

$$L\ddot{a}rermangel_{st} = \beta_0 + \beta_1 lnArbeidsledighet_{kt} + \eta_s + \tau_t + \varepsilon_{skt} \quad (1)$$

Hvor β_0 er konstantledd, η_s fanger opp uobserverte faste skoleeffekter som påvirker lærermangelen på skolenivå, og τ_t fanger opp tidsfaste effekter. ε_{skt} er det idiosynkratiske restleddet og fanger opp både tids- og enhetsvarierende faktorer. Denne enkle modellspesifikasjonen benyttes blant annet for å teste hvilken estimeringsmetode vi vil anvende i analysen når vi inkluderer kontrollvariabler.

Videre ønsker vi å utvide modellen for å kontrollere for forhold ved skolene og i kommunene som kan påvirke graden av lærermangel. Inkluderer i første omgang variabler som omhandler forhold ved skolene, som skolestørrelse og skoletype. Kontrollvariablene er representert ved vektoren \mathbf{X}_{skt} . Vi får modellspesifikasjonen

$$Lærermangel_{st} = \beta_0 + \beta_1 \ln Arbeidsledighet_{kt} + \mathbf{X}_{skt} \beta + \eta_s + \tau_t + \varepsilon_{skt} \quad (2)$$

I tillegg til forhold ved skolene, inkluderer vi ulike demografiske forhold i kommunen, som befolkningstall og andel i de ulike aldersgruppene inn i \mathbf{X}_{skt} i modellspesifikasjon (2).

Størrelsesvariabler som befolkningstall har ofte en skjev fordeling med stor spredning. For å redusere den relative forskjellen i størrelsen på slike variabler kan vi ta logaritmen. Dette vil resultere i en mer symmetrisk og jevn fordeling ved at vi unngår de ekstreme verdiene. I analysen har vi derfor valgt å log-transformere variabelen for antallet elever i skolen og befolkningsantallet i kommunen og vi tolker koeffisienten som en endring i lærermangel på β prosentpoeng ved en dobling av antallet.

Det er rimelig å anta at valget om å jobbe som lærer og endringer i arbeidsledigheten ikke skjer samtidig. Vi ønsker derfor å ytterligere utvide modellen for å undersøke de dynamiske effektene av arbeidsledighet. Dette gjøres ved å forsinke arbeidsledigheten med én periode i modellen. På denne måten kan vi undersøke om det er en treghet i effekten, ved at arbeidsledighetsraten i perioden før påvirker valget om å gå inn i læreryrket eller velge en annen yrkesretning i den nåværende perioden. Denne utvidelsen er også inkludert i \mathbf{X}_{skt} i modellspesifikasjon (2).

5.1.2 Modellspesifikasjoner på kommunenivå

For å oppnå en mer helhetlig forståelse av lærermangelsituasjonen, benytter vi en modell hvor vi anvender lærermangel på kommunenivå som den avhengige variabelen. Hovedmotivasjonen er at dette enhetsnivået samsvarer med nivået arbeidsledigheten måles på. I estimeringen av modellspesifikasjonene som følger, rapporterer vi klostrede standardavvik på kommunenivå. Å klostre på kommunenivå er hensiktsmessig fordi det gir oss bredere informasjon om lærermangelen i en gitt kommune, selv om det ikke nødvendigvis fanger opp hvordan arbeidsledigheten påvirker lærermangelen på de enkelte skolene. Skolene har ulik størrelse, men blir vektet likt når det klostres på skolenivå. Ved klostring på kommunenivå blir de ulike skolestørrelsene tatt hensyn til i noen grad.

Den enkle modellen på kommunenivå er spesifisert som

$$Læremangel_{kt} = \beta_0 + \beta_1 \ln \text{Arbeidsledighet}_{kt} + \eta_k + \tau_t + \varepsilon_{kt} \quad (3)$$

Koeffisienten, β_1 tolkes på samme måte som tidligere. I stedet for skoleeffekter representerer nå komponenten η_k de uobserverte faste kommuneeffektene. Det idiosynkratiske restleddet, ε_{kt} , fanger opp variasjoner mellom tid og kommuner.

Analogt til skolenivå, utvider vi modellens spesifisering (3) på kommunenivå ved å først spesifisere en modell hvor skolekarakteristika inkluderes. Disse inkluderes som \mathbf{X}_{kt} i modell (4). På kommunenivå er kontrollvariablene for skoletypene oppgitt som andel av totalt antall skoler i kommunen. Vi tolker koeffisientene slik at en økning på ett prosentpoeng i andelene barne- eller ungdomsskole assosieres med en endring i lærermangel med β prosentpoeng. Øvrige koeffisienter og variabler tolkes og er identisk som på skolenivå.

$$Læremangel_{kt} = \beta_0 + \beta_1 \ln \text{Arbeidsledighet}_{kt} + \mathbf{X}_{kt} \beta + \eta_k + \tau_t + \varepsilon_{kt} \quad (4)$$

Videre utvider vi modellspesifisering (4) ved å inkludere karakteristika for kommunen. I tillegg legger vi til en forsinket arbeidsledighetsvariabel for å ta hensyn til dynamiske effekter. Variablene blir inkludert i \mathbf{X}_{kt} i (4).

5.2 Statistiske paneldatamodeller

I dette delkapittelet gjennomgås de ulike estimeringsmetodene som kan benyttes for å estimere modeller basert på paneldata. Begynner med den generelle modellen for paneldata som kan gis ved

$$y_{it} = \beta_0 + \mathbf{X}_{it} \beta + u_{it}$$

Hvor restleddet deles opp i to komponenter

$$u_{it} = \eta_i + \varepsilon_{it}$$

Slik at den generelle modellen da blir

$$y_{it} = \beta_0 + \mathbf{X}_{it} \beta + \eta_i + \varepsilon_{it} \quad (5)$$

Tverrsnittskomponenten gitt ved η_i fanger opp alle uobserverte tidkonstante effekter

som påvirker den avhengige variabelen, y_{it} . Den inneholder ikke fotskriften t og antas å være konstant over tid og homoskedastisk på tvers av enheter. Vi har også en komponent som varierer over tid, den idiosynkratiske komponenten ε_{it} . Denne representerer alle de uobserverbare effektene som varierer over tid og påvirker y_{it} .

Videre har vi noen forutsetninger om restleddet ε_{it}

$$E(\varepsilon_{it}|x_{i1}, x_{i2}, \dots, x_{iT}) = 0 \quad (\text{I})$$

$$E(\varepsilon_{it}, \varepsilon_{js}|x_{it}) = \begin{cases} \sigma_\varepsilon^2 & \text{for } i = j \text{ og } t = s \\ 0 & \text{ellers} \end{cases} \quad (\text{II})$$

$$E(\eta_i|x_{i1}, x_{i2}, \dots, x_{iT}) = 0 \quad (\text{III})$$

$$E(\eta_i, \eta_j|x_{i1}, x_{i2}, \dots, x_{iT}) = \begin{cases} \sigma_\eta^2 & \text{for } i = j \\ 0 & \text{ellers} \end{cases} \quad (\text{IV})$$

$$E(\varepsilon_{it}\eta_j|x_{i1}, x_{i2}, \dots, x_{iT}) = 0 \text{ for alle } i, j \text{ og } t \quad (\text{V})$$

Tolkning av forutsetningene:

Forutsetning (I) innebærer at forklaringsvariablene, x , er sterkt eksogene i forhold til det idiosynkratiske restleddet, ε_{it} . Forutsetning (II) forteller at det idiosynkratiske restleddet ikke skal være korrelert innenfor eller mellom enheter, altså er det ingen systematisk sammenheng mellom s og t og i og j . η_i er ikke seriekorrelert og har konstant varians, altså er den homoskedastisk. Forutsetning (III) handler om at x er sterkt eksogen relativt til det enhetsspesifikke restleddet, η_i . Dette er en veldig streng antagelse som innebærer at mulig utelatte variable som kun varierer mellom enhetene er ukorrelert med de inkluderte forklaringsvariablene. Forutsetning (IV) innebærer at det enhetsspesifikke restleddet ikke korrelerer mellom enheter og er homoskedastisk. Forutsetning (V) forteller at de to restleddene, det idiosynkratiske og det enhetsspesifikke, er ukorrelert med hverandre.

Til sammen utgjør disse forutsetningene standardegenskapene. Disse må være oppfylt for at vi skal få konsistente og forventningsrette estimater når vi foretar en regresjonsanalyse. At estimatoren er konsistent betyr at når observasjonene øker, $n \rightarrow \infty$, så konvergerer estimatoren, β , mot den sanne verdien. Med andre ord vil ikke estimatoren i gjennomsnitt avvike systematisk fra den sanne verdien av parameteren når datasettet er stort. Med forventningsrett estimator menes det at den i gjennomsnitt gir det sanne parameterestimatet når den anvendes på et datasett som er uendelig stort.

5.2.1 Sammenslått minste kvadraters metode (POLS)

Sammenslått minste kvadraters metode (Pooled Ordinary Least Squares (POLS)) er en vanlig metode for å estimere lineære regresjonsmodeller med paneldata.

Når vi benytter POLS har vi at estimatoren for koeffisienten β er gitt ved:

$$\hat{\beta}^{POLS} = \frac{\sum^N \sum^T (y_{it} - \bar{y})(x_{it} - \bar{x})}{\sum^N \sum^T (x_{it} - \bar{x})^2} \quad (6)$$

Her illustrerer vi at vi tar gjennomsnittet til variabelen. Med POLS får vi utnyttet all variasjon, både mellom enheter og over tid, ved at summeringene $\sum^N \sum^T$ går over alle observasjonene i datasettet. Derimot tar denne estimeringsmetoden ikke hensyn til tidseffekter, i den forstand at den behandler hele datasettet som et tverrsnitt og ikke ser på de samme enhetene over tid. En stor fordel ved å benytte paneldata er å kunne kontrollere for tidsfaste effekter i restleddet, noe vi ikke får utnyttet med denne estimeringsmetoden.

For å benytte denne metoden er vi avhengig av at forutsetning (III), om at det enhets-spesifikke restleddet er ukorrelert med de inkluderte forklaringsvariablene i modellen, gjelder. Dersom denne ikke er oppfylt vil ikke POLS gi forventningsrette eller konsistente estimater.

Estimeringsmetoden kan altså gi et skjevt estimat på estimatoren $\hat{\beta}_1$, ved at den ikke konvergerer mot den samme parameterverdien β_1 når utvalget går mot uendelig. Dette er konstante effekter i restleddet.

$$plim \hat{\beta}_1 \neq \beta_1 \quad (7)$$

Det er realistisk å anta at det finnes enhetsspesifikke variabler som egentlig skulle vært inkludert i modellen, men som er utelatte og dermed inngår i den enhetsspesifikke komponenten η_i . Årsaken til at de ikke er inkludert i modellen kan være at vi ikke har tilgang på dem, eller at de av andre grunner er uobserverbare. Vi får antagelig et problem med seriekorrelasjon i og med at paneldatamodellen vår ser på de samme kommunene og skolene over en periode på 25 år. Det er derfor lite sannsynlig at restleddene ikke korrelerer med hverandre.

5.2.2 Faste effekter (FE)

Paneldata gjør det mulig å kontrollere for konstante effekter i restleddet. Fra delkapittel 5.2 har vi at en viktig antagelse når vi benytter paneldata er forutsetningen om at det enhetsspesifikke restleddet er ukorrelert med kontrollvariablene. Denne antagelsen er veldig streng, og for å unngå å bryte denne forutsetningen kan vi benytte faste effekter (Fixed Effects (FE)). Regresjonsanalyse med FE brukes for å se på avviket mellom tverrsnittsvariansen og det globale gjennomsnittet. Det gjør at vi eliminerer det enhetsspesifikke restleddet, og at denne antagelsen dermed ikke trenger å være oppfylt, i motsetning til under POLS-estimering. Det er rimelig å anta at kovariansen mellom det enhetsspesifikke restleddet og variablene som er inkludert i modellen er forskjellig fra null.

Vi tar utgangspunkt i den generelle modellen (5), hvor restleddet er dekomponert og får ved hjelp av FE en estimerbar relasjon gitt ved

$$y_{it} - \bar{y}_i = \beta(x_{it} - \bar{x}_i) + (\varepsilon_{it} - \bar{\varepsilon}_i) \quad (8)$$

Ser fra (8) at det enhetsspesifikke restleddet, η_i , forsvinner fordi det ikke varierer over tid. Videre får vi estimatoren for FE

$$\hat{\beta}^{FE} = \frac{\sum^N \sum^T (y_{it} - \bar{y}_i)(x_{it} - \bar{x}_i)}{\sum^N \sum^T (x_{it} - \bar{x}_i)^2} \quad (9)$$

Ulempen med denne metoden er at den ikke kan estimere variabler som kun varierer over tverrsnitt, og ved lite tverrsnittsvariasjon vil vi derfor få upålitelige estimater. FE fjerner alt som ikke varierer over tid innenfor enheter. Lite innenfor-variasjon kan også føre til at estimatene ikke blir forventningsrette.

5.2.3 Tilfeldige effekter (RE)

Estimeringsmetoden tilfeldige effekter (Random Effects (RE)) tar hensyn til både tidskonstante og tidsspesifikke faktorer ved å anta at individuelle effekter er tilfeldige og korrelert med de uavhengige variablene i modellen. Dersom denne estimatoren skal gi konsistente estimater forutsettes det, i likhet med POLS, at antagelse (III) om at forklaringsvariablene er ukorrelert med det enhetsspesifikke restleddet gjelder. Dersom dette er tilfelle er RE-estimatoren også effisient, noe som betyr at den er estimatoren med den laveste variansen. Det er altså et mål på hvor presis estimatoren

er. Likevel betyr det ikke at den er ideell, da en estimatoren kan være effisient, men likevel ikke være forventningsrett.

Vi kan skrive den estimerbare RE modellen gitt ved:

$$y_{it}^* = \beta_0^* + x_{it}^* \beta + u_{it}^* \quad (10)$$

Her er hvert enkelt ledd bestemt av:

$$y_{it}^* = y_{it} - (1 - \theta)\bar{y}_i$$

$$x_{it}^* = x_{it} - (1 - \theta)\bar{x}_i$$

$$\beta_0^* = \beta_0 - (1 - \theta)\beta_0$$

Til sammen blir dette en regresjonsligning med Gauss-Markov egenskaper, som betyr at den oppfyller visse forventninger. Altså er ligningen en lineær spesifisering uten perfekt multikollinearitet. I tillegg er den forventningsrett, har homoskedastiske feilledd og ingen autokorrelasjon. Vi antar også at det ikke er noen endogenitet i feilleddet. θ er videre gitt som

$$\theta = \sqrt{\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_\eta^2}} = \sqrt{\frac{1}{1 + T\frac{\sigma_\eta^2}{\sigma_\varepsilon^2}}}$$

FE og POLS kan betraktes som spesialtilfeller av RE-estimatoren når $\theta = 0$, ettersom RE i dette tilfellet blir ekvivalent med FE. Dette betyr at variansen mellom enhetene, σ_η^2 , dominerer den idiosynkratiske variansen, σ_ε^2 , og tidsperioden er kort, altså t er liten. I motsatt tilfelle, dersom $\theta = 1$, vil den idiosynkratiske variansen dominere tverrsnittsvariansen. Dette skjer dersom $t \rightarrow \infty$ og $\frac{\sigma_\eta^2}{\sigma_\varepsilon^2} \rightarrow \infty$.

5.3 Tester

Hvilken estimeringsmetode som gir de mest konsistente resultatene avhenger av hvilke antagelser som er ivaretatt. Vi kan bruke ulike tester i vurderingen om hvilken metode som er hensiktsmessig å benytte videre i analysen. Skal i dette delkapittelet presentere F-test og Hausman-test.

5.3.1 F-test

En F-test kan benyttes ved å teste signifikansen av å bruke (FE) i paneldatasettet i forhold til (POLS). Imidlertid er det viktig å merke seg at F-testen ikke er avgjørende i vurderingen av hvilken metode som er mest hensiktsmessig å bruke.

I praksis tester vi en generell modell opp mot en restriktiv. Den generelle modellen har enhetsspesifikke konstantledd, mens den restriktive kun blir representert med konstantledd og forklaringsvariabler. Nullhypotesen vår er at alle de enhetsspesifikke konstantleddene, representert ved dummier, er lik null. Dette vil si at det ikke er systematiske forskjeller i kryssningene på tvers av enhetene. Alternativhypotesen er at minst én ikke er lik null, som betyr at det er tilstedeværelse av enhetsspesifikke effekter.

Mer formelt har vi nullhypotesen

$$H_0 : \delta_1 = \delta_2 = \dots = \delta_{N-1} = 0$$

Videre får vi F-testen gitt ved F-statistikken:

$$F = \frac{R_u^2 - R_r^2}{1 - R_u^2} \cdot \frac{n - k - 1}{q}$$

Hvor R_u^2 er R-kvadrert fra den generelle modellen med alle parametre, mens R_r^2 er R-kvadrert fra den restriktive modellen. q er antall restriksjoner, n er antall observasjoner og k er antall parameter i den generelle modellen.

F-statistikken ses opp mot en kritisk verdi som er gitt ved to typer frihetsgrader. Frihetsgradene i telleren er antallet restriksjoner i den restriktive modellen. Frihetsgradene i nevneren er gitt ved $n - k - 1$. Dersom F-statistikken er større enn den kritiske verdien forkaster vi nullhypotesen. Forkasting innebærer at vi foretrekker faste effekter fremfor sammenslått minste kvadraters metode.

$$F > F_{q, n-k-1}$$

Nullhypotesen kan forkastes på ulike signifikansnivå. Signifikansnivået er vanligvis angitt som en prosentandel og representerer terskelen eller grensen for å avgjøre om en statistisk test gir signifikante resultater eller ikke. Med andre ord kan den forklares som hvor stor sannsynligheten er for å feilaktig forkaste en sann nullhypotese.

5.3.2 Hausman-test

Hausman-testing er en vanlig tilnærming for å teste for endogenitet eller for å velge riktig modellspesifikasjon når paneldata benyttes. Vi anvender derfor denne testen i vurderingen om vi skal bruke FE eller RE ved estimering av våre paneldatamodeller.

Denne testen sammenligner koeffisientene i en modell med tilfeldige med en modell med faste effekter og tester om koeffisientene med tilfeldige effekter er forventningsrette eller ikke, og om de er effisiente. Hausman-testen tester om forskjellen mellom koeffisientene er statistisk signifikant. RE er mer effisient enn FE, så dersom de ikke er signifikant forskjellige kan RE benyttes. RE tar også i bruk mer av variasjonen i datamaterialet i motsetning til FE.

Hausman-testen er basert på antagelsen om at de tilfeldige effektene er ukorrelert med forklaringsvariablene, som betyr at de tilfeldige effektene ikke er korrelert med noen utelatte variabler som kan påvirke utfallet. Dette kan dermed føre til at feil modell blir valgt, dersom antagelsen ikke holder.

Når Hausman-testen skal utføres estimeres først begge estimeringsmetodene. Deretter beregnes forskjellen mellom koeffisientene på de to modellene og dividerer på standardavviket for forskjellen. Nullhypotesen vår er at koeffisientene til modellen med tilfeldige effekter er konsistente og effisiente.

Mer formelt formulerer vi nullhypotesen

$$H_0 : E(\eta_i|x_{it}) = 0$$

$$H_A : E(\eta_i|x_{it}) \neq 0$$

Intuisjonen er at under nullhypotesen vil FE-estimatoren og RE-estimatoren ikke være systematisk forskjellige fordi begge er konsistente, som vil si at

$$plim(\hat{\beta}_{FE} - \hat{\beta}_{RE}) = 0$$

Vi bruker χ^2 -fordeling med k frihetsgrader, som er gitt ved

$$\chi_k^2 = (\hat{\beta}_{FE} - \hat{\beta}_{RE})[\hat{V}(\hat{\beta}_{FE} - \hat{V}(\beta_{RE}))]^{-1}(\hat{\beta}_{FE} - \hat{\beta}_{RE})$$

Forkaster nullhypotesen dersom testobservatoren er større enn kritisk verdi, som innebærer at vi foretrekker FE-estimatoren.

6 Resultat

I dette kapittelet presenteres resultatene fra estimeringer gjort på modellspesifikasjonene, lagt frem i forrige kapittel. Først, i delkapittel 6.1, gjøres analysen på skolenivå ved å starte med den enkle modellen. Deretter blir det gjennomført tester for å vurdere de ulike estimeringsmetodene. Videre blir resultatene med kontrollvariabler presentert. I delkapittel 6.2 gjennomføres analysen på kommunenivå med lik fremgangsmåte som ved analysen på skolenivå. Delkapittel 6.3 undersøker robustheten av resultatene ved å foreta endringer i hovedanalysen, ved å bytte ut variabelen for arbeidsledighetsraten og benytte et annet mål på lærermangel. Vi utvider analysen ytterligere ved å se på ledighet i arbeidsmarkedsregionene i stedet for i kommunene. I delkapittel 6.4 undersøkes det for heterogenitet, hvor vi først deler utvalget i to tidsperioder. Deretter skiller vi mellom de ulike skoletypene.

6.1 Hovedmodell: Skolenivå

Vi begynner analysen med å se på lærermangel på skolenivå, noe som gjør det mulig å kontrollere for faktorer som også påvirker innad i en kommune. Det er rimelig å anta at skolene har ulik attraktivitet basert på hvor de er lokalisert, spesielt i de store kommunene.

Alle modellene som er estimert i tabell 7 samsvarer med modellspesifikasjon (1) og er vår enkleste modell. Modell (1) benytter sammenslåtte minste kvadraters metode (POLS) og vi inkluderer klustrede standardavvik (kommandoen "cluster" i STATA) for å korrigere for heteroskedastisitet og seriekorrelasjon på skolenivå. I modell (2) inkluderer vi tidsfaste effekter, men korrigerer ikke for heteroskedastisitet og seriekorrelasjon. Det er ønskelig å fange opp tidseffektene som påvirker lærermangel likt. Ved å inkludere tidsdummier for hvert av årene kan vi kontrollere for den nasjonale variasjonen, som for eksempel lønnsvariasjoner.

I modellene (3) - (5) benytter vi faste effekter (FE) som estimeringsmetode, og får dermed kontrollert for uobserverte effekter på skolenivå som ikke varierer over tid. I (3) og (4) har vi valgt å se på effekten av arbeidsledighet uten å korrigere for seriekorrelasjon og heteroskedastisitet når vi benytter FE-metoden. I modell (3) har vi også utelatt å inkludere tidsdummier, mens i (5) har vi både tidsfaste effekter og klustrede standardavvik.

Modell (6) benytter estimeringsmetoden tilfeldige effekter (RE). Her kontrollerer vi for tidsfaste effekter, og korrigerer for seriekorrelasjon og heteroskedastisitet.

Resultatene presentert i tabell 7 viser at den estimerte effekten av arbeidsledighet på lærermangel varierer mellom de forskjellige metodene. Modell (1) gir en signifikant negativ effekt av arbeidsledighet, mens i modell (2) får vi at koeffisienten endrer fortegn. Vi finner en svak positiv effekt av arbeidsledighet på lærermangel, men denne er ikke-signifikant.

R-kvadrert uttrykker hvor stor andel av den totale variasjonen i datamaterialet som regresjonsmodellen forklarer. Koeffisienten kan ta verdier mellom 0 og 1, hvor en verdi mot 1 tilsier at det som estimeres ligger nærme dataene, altså hvor godt modellen passer til de faktiske observasjonene i datasettet. Vi finner at R-kvadrert øker fra 0,000 til 0,026 når de tidsfaste effektene inkluderes, som tyder på at modell (2) forklarer mer av variasjonen og gir et mer presist bilde av sammenhengen mellom arbeidsledighet og lærermangel.

Modell (3) estimert med FE gir en signifikant negativ effekt på 0,516 når vi ikke inkluderer tidsfaste effekter. Vi tolker koeffisienten som at en økning i arbeidsledighetsraten med én prosent fører til en reduksjon i lærermangel på 0,516 prosentpoeng. At effekten ikke lenger er signifikant når tidsdummiene inkluderes antyder at de signifikante resultatene i (1) og (3) drives av makrovariasjon og ikke av ulik utvikling i arbeidsledighet mellom kommunene. Vi mangler imidlertid kunnskap om arten av denne makrovariasjonen, da det ikke undersøkes i vår analyse. Når vi tar hensyn til heterogenitetsfaktorer som er konstante over tid innenfor hvert panel i modell (4), reduseres koeffisienten til 0,086, og vi har ikke lenger en signifikant effekt av arbeidsledighet.

Videre, når vi i tillegg korrigerer for seriekorrelasjon og heteroskedastisitet på skolenivå ved å inkludere klustrede standardavvik i (5) får vi samme effekt av arbeidsledighet som i (4), men høyere standardavvik. En mulig årsak til dette kan være at det er mer variasjon mellom enhetene som tilhører samme kluster enn det som antas i FE-metoden. Resultatet ved bruk av RE-metoden i modell (6) gir en svak negativ effekt av arbeidsledighet på lærermangel som ikke er signifikant.

Tabell 7: Lærermangel på skolenivå uten kontrollvariabler

SKOLENIVÅ						
Metode	(1)	(2)	(3)	(4)	(5)	(6)
Variabler	POLS	POLS	FE	FE	FE	RE
	Lærermangel _s	Lærermangel _s	Lærermangel _s	Lærermangel _s	Lærermangel _s	laerermangel _s
lnArbeidsledighet_NAV	-0,212** (0,107)	0,042 (0,061)	-0,516*** (0,057)	-0,086 (0,081)	-0,086 (0,117)	-0,050 (0,106)
Tidsfaste effekter	Nei	Ja	Nei	Ja	Ja	Ja
Skolefaste effekter	Nei	Nei	Ja	Ja	Ja	-
Kluster	Ja	Nei	Nei	Nei	Ja	Ja
Observasjoner	77 358	77 358	77 358	77 358	77 358	77 358
R-kvadrert	0,000	0,026	0,001	0,039	0,039	-
Antall skoler	-	-	3 938	3 938	3 938	3 938

Standardavvik i parentes
 *** p<0.01, ** p<0.05, * p<0.1

6.1.1 Valg mellom estimeringsmetoder

Noen estimeringsmetoder er mer hensiktsmessige enn andre, og som tidligere diskutert så avhenger valget av metodene av hvilke forutsetninger som er ivaretatt. I tillegg må metodene ses i sammenheng med formålet med analysen og det spesifikke forskningsspørsmålet som ønskes besvart.

I delkapittel 5.2.1 ble bruken av POLS som estimeringsmetode presentert og diskutert. Basert på denne diskusjonen betraktes POLS som en metode med visse svakheter når det gjelder bruk på paneldatasett. Vi ønsker å undersøke med en F-test hvorvidt POLS er upassende som estimeringsmetode i forhold til FE i vårt tilfelle. Fra delkapittel 5.3.1 er nullhypotesen formulert slik at de enhetsspesifikke restleddene, representert ved dummyvariabler, er lik null. Under alternativhypotesen er minst én av koeffisientene foran dummyvariablene ikke lik null, noe som innebærer at vi har tilstedeværelse av enhetsspesifikke restledd og dermed foretrekker FE som estimeringsmetode.

Vi utfører testen og får at F-statistikken er $F = 82,76$. Den kritiske verdien på et 1% signifikansnivå er $F_{1,3937} = 6,63$. Dette innebærer at F-statistikken er tilstrekkelig høy til at vi kan forkaste nullhypotesen på et 1% signifikansnivå. Videre vil vi likevel inkludere POLS i vår analyse når vi utvider hovedmodellen. Grunnen til at vi ikke forkaster estimeringsmetoden på bakgrunn av resultatet i F-testen er at det er interessant å sammenligne med øvrige estimeringsmetoder gjennom analysen da vi ved FE mister en del av variasjonen. I tillegg ønsker vi å gjøre analysen på kommunenivå identisk til skolenivå, og inkluderer derfor POLS som estimeringsmetode.

Videre undersøker vi hvilken estimeringsmetode av FE og RE som er mest hensiktsmessig for analysen ved å benytte en Hausman-test. Under nullhypotesen er både

FE-metoden og RE-metoden konsistent, mens kun RE er effisient. Vi vil i en slik situasjon derfor foretrekke RE fremfor FE. Under alternativhypotesen er kun FE-metoden konsistent, og derfor å foretrekke som estimeringsmetode i denne analysen. Når vi gjennomfører en hausman-test kan vi ikke bruke klustrede standardavvik i estimeringen. Vi har en χ^2 -fordelt testobservator med 25 frihetsgrader. Kritisk verdi er $\chi_{0,10}^2(25) = 34,38$ og testobservatoren er $\chi_{0,10}^2(25) = 3,43$, noe som betyr at vi kan forkaste på et 10% signifikansnivå. I denne testen har vi antatt at de to estimatorene har samme varians, som i noen tilfeller kan være en upresis antagelse. Dersom vi tar hensyn til at variansen i de to estimatorene er ulike får vi testobservatoren $\chi_{0,05}^2(25) = 73,41$. Med kritisk verdi $\chi_{0,05}^2(25) = 37,65$ kan vi klart avvise RE som estimator på 5% signifikansnivå, og fortsetter derfor med faste effekter som hovedmetode videre i analysen.

6.1.2 Skolenivå med kontrollvariabler

Fortsetter hovedanalysen med å undersøke hvordan graden av lærermangel påvirkes av arbeidsledighet når kontrollvariabler inkluderes. Resultatene presenteres i tabell 8. Modell (7) inkluderer skolekarakteristikker, mens modell (8) er ytterligere utvidet med kommunekarakteristikker. I modell (9) er arbeidsledigheten forsinket med én periode. Modell (7) - (9) er alle estimert med FE. I modell (10) benytter vi POLS som estimeringsmetode og inkluderer alle kontrollvariablene forutenom den forsinkede arbeidsledigheten.

Begynner med å kommentere på effekten av arbeidsledigheten, som i tabell 8 har en gjennomgående ikke-signifikant effekt. I modell (7) ser vi en forventet, men svak negativ effekt av arbeidsledighet. Utvidelsen i modell (8) gjør at koeffisienten bytter fortegn, og vi finner en veldig svak positiv effekt. Dette tyder på at det er faktorer ved kommunens demografi som tar noe av den negative effekten på lærermangel. Ved å inkludere en forsinket periode av arbeidsledighet i (9), får vi en ytterligere positiv effekt av fokusvariabelen. Dette kan skyldes at den forsinkede arbeidsledigheten delvis fanger opp noen av faktorene som tidligere lå i arbeidsledighetsvariabelen. Den forsinkede arbeidsledigheten har en negativ ikke-signifikant effekt. Inkluderingen av kontrollvariabler endrer den totale variasjonen i dataene og forholdet mellom variablene, som kan påvirke den direkte effekten av fokusvariabelen. Det er derfor ikke overraskende at fortegnene på koeffisientene varierer ved endringer i modellspesifikasjoner og estimeringsmetoder, spesielt når effekten ikke er å finne som statistisk signifikant. R-kvadrert øker i forhold til den enkle modellspekifikasjonen når det legges til kontrollvariabler i estimering med FE.

Basert på F-testen som ble gjennomført i delkapittel 6.1.1 vurderer vi FE som den foretrukne estimeringsmetode fremfor POLS. Likevel er det interessant å undersøke hvordan POLS reagerer når vi inkluderer kontrollvariabler. Dette skyldes at estimeringsmetoden har en fordel når det gjelder å utnytte variasjonen over tid. I modell (6) er effekten av arbeidsledighet negativ, men svakere sammenlignet med arbeidsledighetskoeffisienten i modell (2) i tabell 7. Heller ikke her finner vi en signifikant effekt av arbeidsledighet.

Tabell 8: Lærermangel på skolenivå med kontrollvariabler

SKOLENIVÅ				
Metode	(7)	(8)	(9)	(10)
Variabler	FE	FE	FE	POLS
	Lærermangel _s	Lærermangel _s	Lærermangel _s	Lærermangel _s
lnArbeidsledighet_NAV	-0,058 (0,115)	0,007 (0,114)	0,020 (0,116)	-0,164 (0,133)
Barneskole	0,481** (0,244)	0,474* (0,243)	0,544** (0,256)	-1,491*** (0,151)
Ungdomsskole	-0,645* (0,340)	-0,609* (0,339)	-0,688* (0,357)	-1,740*** (0,169)
lnElever0110_sum	0,011 (0,187)	0,052 (0,191)	0,072 (0,213)	-1,250*** (0,078)
Spesandel_verk0110	0,030*** (0,004)	0,030*** (0,004)	0,030*** (0,004)	0,015*** (0,005)
Eldre		-0,110** (0,054)	-0,075 (0,057)	0,088*** (0,025)
Arbeidsalder		-0,164*** (0,053)	-0,130** (0,057)	0,316*** (0,043)
lnBefolkningstall		-2,156*** (0,618)	-1,743*** (0,664)	-0,261*** (0,061)
L.lnArbeidsledighet_NAV			-0,081 (0,117)	
Tidsfaste effekter	Ja	Ja	Ja	Ja
Skolefaste effekter	Ja	Ja	Ja	Nei
Kluster	Ja	Ja	Ja	Ja
Observasjoner	77 241	77 241	73 126	77 241
R-kvadrert	0,042	0,043	0,044	0,090
Antall skoler	3 938	3 938	3 864	-

Standardavvik i parentes
 *** p<0,01, ** p<0,05, * p<0,1

Videre ser vi på effekten av kontrollvariablene når vi benytter de ulike estimeringsmetodene. Vi får forskjellig utslag på lærermangel når vi ser på de ulike skoletypene opp mot kombinerte skoler. I den fullstendige modellen (9) finner vi at barneskoler

har 0,544 prosentpoeng høyere gjennomsnittlig lærermangel sammenlignet med kombinertskoler. For ungdomsskoler ser vi at koeffisienten har et negativt fortegn og vi finner at gjennomsnittlig lærermangel er 0,688 prosentpoeng lavere enn skoler som er kombinert. Koeffisientene er relativt like i alle modellene som er estimert med FE.

Videre finner vi at en økning på ett prosentpoeng i spesialundervisning er assosiert med en økning på 0,03 prosentpoeng i lærermangel i modell (9), når alle andre variabler holdes konstant. Denne effekten er signifikant og omtrentlig lik for alle modellene estimert med faste effekter, som tyder på at den ikke reagerer betydningsfullt på inkludering av øvrige kontrollvariablene. Det er verdt å merke seg at deskriptiv statistikk for kontrollvariablene i delkapittel 4.4 viste stor variasjon i graden av spesialundervisning mellom skolene.

I modell (8) finner vi at en økning på ett prosentpoeng i andelen av befolkningen som går under kategorien eldre i forhold til andelen barn fører til en reduksjon i lærermangel på 0,11 prosentpoeng. Denne effekten er signifikant, mens i modell (9) er den ikke lenger signifikant. Både (8) og (9) viser en forventet negativ effekt av arbeidsalder. En økning i andelen av befolkningen i arbeidsalder på ett prosentpoeng i forhold til andelen barn assosieres i (9) med en reduksjon i lærermangel på 0,13 prosentpoeng. Effekten av befolkningstall er signifikant i hele tabellen. Om befolkningen dobles, gir det en reduksjon i lærermangel med 2,156 prosentpoeng uten inkludering av forsinket arbeidsledighetsvariabel. Inkluderingen medfører en svakere negativ effekt, ved at doubling av befolkningen gir en reduksjon på 1,74 prosentpoeng i lærermangel.

Alle kontrollvariablene i modell (10) som er estimert med POLS er signifikante. De gir ulike resultater enn når vi benytter FE, og det er ikke noe gjennomgående mønster i hvordan de to estimeringsmetodene varierer. Gjennomsnittlig lærermangel på barne-skole er i likhet med ungdomsskole lavere enn ved skolene som er kombinert. Effekten av skolestørrelse skiller seg ut i modell (10). Det er grunn til å tro at større skoler opplever mindre tilstedeværelse av lærermangel. Dette fanges ikke opp i (7)-(9) på grunn av kontrollering av faste effekter, da det er rimelig å anta at skolene har omtrent lik størrelse over tid. Vi finner en negativ signifikant effekt, hvor en doubling i antall elever ved skolen fører til en reduksjon i lærermangel på 1,743 prosentpoeng. Effekten av andelen spesialundervisning er fremdeles positiv og signifikant, men mindre ved POLS-estimering enn FE. Effekten av de demografiske faktorene i kommunen er ved POLS positiv når vi ser på andelen av befolkningen som er eldre og i arbeidalder. Vi finner en negativ effekt av befolkningsstørrelse, men den er mindre enn når vi benytter FE.

6.2 Hovedmodell: Kommun nivå

Videre ser vi på lærermangel på kommunenivå. At lærermangel er på kommunenivå betyr at skolene i datasettet er slått sammen slik at det laveste enhetsnivået er kommune. Analysen følger samme fremgangsmåte som med datasettet på skolenivå, hvor den enkle modellen med de ulike estimeringsmetodene presenteres først, etterfulgt av å inkludere kontrollvariabler.

Tabell 9 viser estimering av modeller som samsvarer med modellspesifikasjon (5). I likhet med på skolenivå, er modellen først estimert med POLS-metoden, etterfulgt av FE-metoden og til slutt blir RE-metoden brukt for å estimere modell (6).

Modell (2) med POLS og tidsfaste effekter gir på kommunenivå en signifikant positiv effekt. Modell (3), med kommunefaste effekter hvor det ikke er kontrollert for tidsfaste effekter, gir en signifikant negativ effekt av arbeidsledighet på lærermangel. På samme måte som på skolenivå er ikke effekten signifikant i modell (4). Dette skyldes at vi inkluderer tidsdummier som tar hensyn til heterogenitetsfaktorer som er konstante over tid innenfor hvert panel. I modell (5) får vi samme koeffisient, men forskjellig standardavvik når vi har klustrede standardavvik. Estimering med RE i (6) gir positiv ikke-signifikant effekt av arbeidsledighet. R-kvadrert er 0,122 på kommunenivå med FE, mens den er 0,039 i tilsvarende modell på skolenivå.

Tabell 9: Lærermangel på kommunenivå uten kontrollvariabler

KOMMUNENIVÅ						
Metode	(1)	(2)	(3)	(4)	(5)	(6)
Variabler	POLS	POLS	FE	FE	FE	RE
	Lærermangel.k	Lærermangel.k	Lærermangel.k	Lærermangel.k	Lærermangel.k	Lærermangel.k
lnArbeidsledighet_NAV	0,347 (0,222)	0,815*** (0,091)	-0,536*** (0,087)	-0,054 (0,109)	-0,054 (0,166)	0,064 (0,157)
Tidsfaste effekter	Nei	Ja	Nei	Ja	Ja	Ja
Kommunefaste effekter	Nei	Nei	Ja	Ja	Ja	-
Kluster	Ja	Nei	Nei	Nei	Ja	Ja
Observasjoner	10 667	10 667	10 667	10 667	10 667	10 667
R-kvadrert	0,002	0,080	0,004	0,122	0,122	-
Antall kommuner	-	-	436	436	436	436

Standardavvik i parentes
 *** p<0,01, ** p<0,05, * p<0,1

6.2.1 Kommunenivå med kontrollvariabler

På samme måte som i analysen på skolenivå går vi videre med FE og POLS som estimeringsmetoder når det legges til kontrollvariabler. Modell (7) inkluderer skolekarakteristikker, mens modell (8) er ytterligere utvidet med kommunekarakteristikker. I modell (9) har vi lagt til forsinkelse på ledighet på én periode. Modell (10) er estimert med POLS-metoden og tidsfaste effekter, og vi har inkludert alle kontrollvariabler utenom forsinkelse av arbeidsledighet.

Vi begynner med å se på hvordan effekten av arbeidsledighet reagerer på inkludering av kontrollvariablene. Effekten av arbeidsledighet på lærermangel på kommunenivå er kun signifikant i modell (10) i tabell 10 når vi estimerer med POLS. Effekten er positiv og gir at en økning i arbeidsledighet på én prosent fører til økt lærermangel med 0,654 prosentpoeng. Dette er motsatt effekt av arbeidsledighet enn det vi fant i modell (10) i tabell 8, når lærermangel ble betraktet på skolenivå og med kontrollvariabler. Estimeringene med FE i tabell 10 gir at effekten av arbeidsledighet varierer mellom å være positiv og negativ avhengig av hvilke kontrollvariabler som er inkludert. Vi finner ingen signifikant effekt av arbeidsledighet i modellene (7)-(9).

Tabell 10: Lærermangel på kommunenivå med kontrollvariabler

KOMMUNENIVÅ				
Metode	(7)	(8)	(9)	(10)
Variabler	FE	FE	FE	POLS
	Lærermangel_k	Lærermangel_k	Lærermangel_k	Lærermangel_k
lnArbeidsledighet_NAV	-0,101 (0,162)	0,010 (0,166)	-0,030 (0,154)	0,654*** (0,245)
Barneskole_andel	0,013** (0,005)	0,015*** (0,005)	0,015*** (0,005)	-0,007 (0,005)
Ungdomsskole_andel	-0,007 (0,013)	-0,005 (0,012)	-0,004 (0,013)	-0,012 (0,010)
lnElever0110_sum	1,171* (0,642)	1,844 (1,149)	1,299 (1,228)	-3,244** (1,406)
Spesandel_verk0110	0,051*** (0,013)	0,053*** (0,013)	0,053*** (0,014)	0,048*** (0,017)
Eldre		-0,235*** (0,087)	-0,240*** (0,092)	-0,069 (0,080)
Arbeidsalder		-0,248*** (0,090)	-0,261*** (0,096)	-0,041 (0,103)
lnBefolkningstall		-4,177*** (1,527)	-3,083** (1,551)	2,566* (1,407)
L.lnArbeidsledighet_NAV			0,119 (0,138)	
Tidsfaste effekter	Ja	Ja	Ja	Ja
Kommunefaste effekter	Ja	Ja	Ja	Nei
Kluster	Ja	Ja	Ja	Ja
Observasjoner	10 667	10 667	10 178	10 667
R-kvadrert	0,130	0,137	0,139	0,136
Antall kommuner	436	436	436	-

Robuste standardavvik i parentes
*** p<0,01, ** p<0,05, * p<0,1

Videre ser vi på effekten av kontrollvariablene på lærermangel på kommunenivå, hvor vi først tar for oss modellene estimert med FE- metoden. Finner at effekten av andelen barneskoler i kommunen er signifikant og forholdsvis stabil over de tre ulike modellspesifikasjonene. Tolker koeffisienten i den fullstendige modellen (9) som at ett prosentpoeng økning i andelen barneskoler i kommunen gir en økning i lærermangel på 0,015 prosentpoeng når andre variabler holdes konstant. Effekten av en økning i andelen ungdomsskoler er negativ, men ikke-signifikant. Skolestørrelse, gitt ved elevtall, er signifikant kun når vi i modell (7) inkluderer skolekarakteristika. En doubling i antall skoleelever i kommunen medfører en økning i lærermangel på 1,171 prosentpoeng. En forklaring på at modell (8) og (9) ikke lenger gir signifikant effekt kan være at befolkningstallet i kommunen eller at andelen i arbeidsstyrken er viktige faktorer

som påvirker tilgangen på kvalifiserte lærere uavhengig av antall elever. Dermed kan disse variablene dominere i forklaringen av lærermangel i modellen, og effekten av elever kan miste sin signifikans. Sammenlignet med resultatene fra tabell 8 finner vi at effekten av andelen spesialundervisning fortsatt er signifikant og har økt.

Effekten av demografiske variabler er signifikant på kommunal lærermangel. Ett prosentpoengs økning av både eldre og arbeidsrelevant alder i befolkningen opp mot andelen barn gir en signifikant reduksjon av lærermangel med henholdsvis 0,240 og 0,261 prosentpoeng i modell (9). Større befolkningstall i helhet gir også redusert lærermangel og vi ser fra modell (9) at en dobling i kommunens befolkning reduserer den kommunale lærermangelen med 3,083 prosentpoeng. Forsinket arbeidsledighet gir en positiv, ikke-signifikant effekt på lærermangel i neste periode.

Modell (10) viser en svak negativ ikke-signifikant effekt av andelen barneskole og ungdomsskole i kommunen. Derimot har skolestørrelse signifikant negativ effekt. Om elevtallet dobles medfører det en reduksjon i lærermangel på 3,244 prosentpoeng. Videre ser vi at vi får en økning i lærermangel på 0,048 prosentpoeng dersom andelen spesialundervisning i kommunen øker med ett prosentpoeng. Verken andelen eldre eller andelen i arbeidsalder, i forhold til andelen barn, har signifikant effekt ved estimering med POLS-metoden på kommunenivå. Til forskjell fra estimeringen på skolenivå i tabell 8 finner vi en svakere negativ effekt. Effekten av en dobbelt så stor befolkning i kommunen gir derimot en positiv og signifikant effekt på 2,566 prosentpoeng.

6.3 Robusthet

En viktig del av modellering i empirisk forskning er å foreta en robusthetssjekk, også kalt en sensitivitetsanalyse. Vi ønsker å undersøke robustheten av analysen som er gjort hittil, altså se på hvor sensitive resultatene i modellen er ovenfor endringer i modellspesifikasjoner. Å foreta en robusthetssjekk er viktig for å sikre at resultatene er pålitelige og generaliserbare. Mer spesifikt ønsker vi å se hvordan effekten av arbeidsledighet påvirkes.

Alle analysene i dette kapitlet blir gjort med den enkleste modellen, tilsvarende modell (5), og modellen som inkluderer både skole- og kommunekaraktistika, tilsvarende modell (8), fra hovedanalysen. Både på skole- og kommunenivå er det kontrollert for tidsfaste effekter og det er brukt klustrede standardavvik. For lærermangel på skolenivå har vi faste skoleeffekter, mens på kommunenivå har vi faste kommuneeffekter. Fra hovedanalysen fant vi verken signifikant effekt av arbeidsledighet eller av

forsinket ledighet. Velger derfor å ikke inkludere forsinkelse i arbeidsledigheten videre.

Vi begynner robusthetssjekken med å se på effekten av å bruke den andre arbeidsledighetsvariabelen. Videre betrakter vi lærermangel i antall lærere som ikke oppfyller kravet til kompetanse, i stedet for i antall årsverk, og undersøker om resultatene holder seg robuste og stabile. Til slutt ser vi på en annen enhetsinndeling av datamaterialet ved å ta hensyn til arbeidsledighet i arbeidsmarkedsregioner fremfor i kommunene.

6.3.1 Analyse med den andre variabelen på arbeidsledighet

I dette delkapittelet benytter vi arbeidsledighetsraten fra FIVA, *Arbeidsledighet_FIVA*, som fokusvariabel. I kontrast til arbeidsledighetsraten fra NAV, brukes det i FIVA et årsgjennomsnitt av arbeidsledighetsraten i kommunen. I tillegg er alle i arbeidsdyktig alder brukt i nevneren, mens NAV bruker arbeidsstyrken. Denne forskjellen bør imidlertid påvirke resultatene i liten grad da vi kontrollerer for befolkningstallet.

Kommenterer først på hvordan effekten av arbeidsledighet på lærermangel endrer seg når den andre arbeidsledighetsraten benyttes. Tabell 11 viser at arbeidsledigheten fra FIVA gir en estimert negativ effekt både på skole- og kommunenivå, men den er ikke-signifikant. Tilsvarende gav modell (8) i tabell 8 og 10 en positiv effekt på lærermangel når det ble estimert med arbeidsledighet fra NAV.

Tabell 11: Analyse med bruk av arbeidsledighet fra FIVA

Metode Variabler	SKOLENIVÅ		KOMMUNENIVÅ	
	(1) FE Lærermangel_s	(2) FE Lærermangel_s	(3) FE Lærermangel_k	(4) FE Lærermangel_k
lnArbeidsledighet_FIVA	-0,196 (0,142)	-0,074 (0,137)	-0,264 (0,233)	-0,174 (0,236)
Barneskole		0,456* (0,243)		
Ungdomsskole		-0,613* (0,339)		
lnElever0110_sum		0,046 (0,191)		2,023* (1,159)
Spesandel_verk0110		0,030*** (0,004)		0,053*** (0,013)
Eldre		-0,100* (0,054)		-0,192** (0,090)
Arbeidsalder		-0,155*** (0,054)		-0,220** (0,092)
lnBefolkningstall		-2,118*** (0,613)		-4,201*** (1,498)
Barneskole_andel				0,014*** (0,005)
Ungdomsskole_andel				-0,005 (0,012)
Tidsfaste effekter	Ja	Ja	Ja	Ja
Skolefaste effekter	Ja	Ja	-	-
Kommunefaste effekter	-	-	Ja	Ja
Kluster	Ja	Ja	Ja	Ja
Observasjoner	77 425	77 308	10 719	10 719
R-kvadrert	0,039	0,043	0,120	0,133
Antall skoler	3 938	3 938	-	-
Antall kommuner	-	-	436	436

Standardavvik i parentes
 *** p<0,01, ** p<0,05, * p<0,1

Vi kommenterer også på hvordan kontrollvariablene reagerer på bruken av den andre arbeidsledighetsvariabelen. Ved å sammenligne funnene våre i tabell 11 med tabell 8 på skolenivå og tabell 10 på kommunenivå, observerer vi at effekten av kontrollvariablene er relativt like. R-kvadrert for de to modellene er relativt like. Det indikerer at arbeidsledighetsvariabelen fra FIVA gir tilsvarende god modell som når vi bruker arbeidsledighet fra NAV.

Oppsummert for denne robusthetstesten kan vi konkludere med at ingen av de to variablene for ledighet viser en signifikant effekt av arbeidsledighet på lærermangel i perioden 1995-2019. Koeffisienten har en verdi nært null ved begge estimeringenene, og det er derfor ikke uventet at koeffisienten varierer inkonsekvent mellom positive og negative verdier.

6.3.2 Annet mål på lærermangel

Videre ønsker vi å undersøke robustheten i våre resultater ved å gjøre endringer i den avhengige variabelen. Spesifikt benytter vi antall lærere uten kompetanse som alternativt mål for å kvantifisere lærermangel, i stede for antall årsverk gjennomført av lærere uten kompetanse. Denne tilnærmingen til lærermangel har imidlertid noen ulemper som vi må ta i betraktning.

Den første ulempen omhandler antall år med observasjoner i den forstand at vi får en kortere tidshorisont. Fra og med 2004 ble det registrert antall lærere ved skolene i GSI. I 2015 begynte i tillegg registreringen av antallet av lærerne som oppfylte kravet om kompetanse, samt hvor mange som ikke oppfylte kravet. Dette innebærer at tidsperioden vi har datamateriale på blir 2015-2019, som i denne sammenheng er en kort periode. En kort tidsperiode gjør det utfordrende å bruke FE som estimeringsmetode på grunn av lite tverrsnittsvariasjon. Dette kan føre til at vi får upålitelige estimater. POLS er ikke like avhengig av en lang tidperiode som FE, og derfor hensiktsmessig å benytte når vi bruker det alternative målet for lærermangel.

En annen ulempe ved dette målet på lærermangel er at det ikke skiller på hvor stor stillingsprosent hver lærer har. På denne måten kan vi få et mer upresist mål for hvor omfattende lærermangelen i praksis er.

Variabelen for andelen lærere uten kompetanse, *LM_andelssats*, genereres ved å dele antallet lærere uten kompetanse på totalt antall lærere. På skolenivå er det 14 154 observasjoner, og på kommunenivå er det 2 102 observasjoner. Vi har altså et mye mindre panel. Deskriptiv statistikk for andelen ansatte uten kompetanse gir at gjennomsnittlig lærermangel på skolenivå ligger på 6,9%. På kommunenivå er det i gjennomsnitt

7,4% av lærerne som er uten kompetanse. Begge disse gjennomsnittene er større enn gjennomsnittene ved definisjonen av lærermangel gitt ved andelen årsverk, som er på rundt 4% på både skole- og kommunenivå, i tabell 1. Forskjellen er forventet på bakgrunn av at det er rimelig å anta at lærere uten formell kompetanse oftere er ansatt i de mindre stillingene.

I tabell 12 finner vi at den estimerte effekten av arbeidsledighetsraten med FE er negativ, men ikke-signifikant. Det er i kontrast til tilsvarende modell (8) i tabell 8 og 10, hvor vi fant en svak positiv ikke-signifikant effekt ved bruk av det opprinnelige målet på lærermangel. Når det gjelder effekten av kontrollvariablene med FE, finner vi i tabell 12 heller ikke noe signifikant, forutenom effekten av andelen ungdomskoler i modell (5). Videre ser vi at R-kvadrert er betydelig redusert sammenlignet med hovedmodellen både på kommune- og skolenivå.

Estimering med POLS i modell (3) og (6) gir til forskjell fra FE en positiv effekt av arbeidsledighet på lærermangel både på skole- og kommunenivå. På kommunenivå observerer vi en signifikant effekt, hvor en økning i arbeidsledigheten med én prosent gir økt lærermangel med 1,253 prosentpoeng. Sammenlignet med tilsvarende modell (10) i tabell 10 er effekten sterkere med det alternative målet på lærermangel. Videre finner vi i modell (3) i tabell 12 at effekten av kontrollvariablene er signifikante, forutenom spesialundervisning og andelen eldre. I modell (6) ser vi en betydelig negativ og signifikant effekt av elevtallet i kommunen. Sammenlignet med tilsvarende modell (10) i tabell 10 er effekten sterkere. Andelen spesialundervisning er også signifikant og positiv, som forventet.

Tabell 12: Analysen med bruk av annet mål på lærermangel

Metode Variabler	SKOLENIVÅ			KOMMUNENIVÅ		
	(1) (FE)	(2) (FE)	(3) (POLLS)	(4) (FE)	(5) (FE)	(6) (POLLS)
	LM_andelssats	LM_andelssats	LM_andelssats	LM_andelssats	LM_andelssats	LM_andelssats
lnArbeidsledighet_NAV	-0,061 (0,421)	-0,059 (0,417)	0,539 (0,427)	-0,327 (0,524)	-0,386 (0,531)	1,253** (0,589)
Barneskole		0,229 (1,001)	-2,268*** (0,369)			
Ungdomsskole		-2,347 (2,794)	-2,803*** (0,422)			
lnElever0110_sum		-0,137 (0,899)	-3,478*** (0,245)		-1,036 (4,849)	-10,11* (5,647)
Spesandel_verk0110		0,027* (0,015)	-0,012 (0,015)		0,050 (0,036)	0,090* (0,046)
Eldre		-0,480 (0,419)	0,112 (0,076)		-0,417 (0,509)	-0,252 (0,303)
Arbeidsalder		-0,651 (0,412)	0,658*** (0,134)		-0,154 (0,457)	-0,079 (0,358)
lnBefolkningstall		-3,001 (3,884)	-0,793*** (0,196)		-0,573 (8,290)	8,228 (5,632)
Barneskole_andel					0,005 (0,021)	0,003 (0,012)
Ungdomsskole_andel					0,057** (0,025)	-0,034 (0,021)
Faste tidseffekter	Ja	Ja	Ja	Ja	Ja	Ja
Faste skoleeffekter	Ja	Ja	Nei	-	-	-
Faste kommuneeffekter	-	-	-	Ja	Ja	Nei
Kluster	Ja	Ja	Ja	Ja	Ja	Ja
Observasjoner	14 122	14 096	14 096	2 102	2 102	2 102
R-kvadrert	0,002	0,004	0,144	0,010	0,015	0,147
Antall skoler	2 977	2 977	-	-	-	-
Antall kommuner	-	-	-	428	428	-

Standardavvik i parentes
 *** p<0,01, ** p<0,05, * p<0,1

Når vi bruker antall lærere uten kompetanse som mål på lærermangel, får vi en betydelig reduksjon i antall observasjoner. Dette gjør det utfordrende å direkte sammenligne estimeringsresultatene, og vi kan ikke fastslå om forskjellene skyldes utvalget eller det faktum at målet på lærermangel er annerledes. Denne reduksjonen i datamaterialet gir oss grunn til å anta at det opprinnelige målet på lærermangel gir oss mer pålitelig resultater.

6.3.3 Arbeidsmarkedsregion

Analysen til Falch et al. (2009) ble gjort med arbeidsledighet på regionalt nivå. I datasettet fra FIVA har vi en identifikasjonsvariabel for hvilket arbeidsmarked de ulike kommunene tilhører. Vi benytter denne variabelen for konstruere en felles ledighetsvariabel på regionsnivå. Dette gjøres ved å summere antall arbeidsledige og antall personer i arbeidsstyrken i hver kommune innenfor den samme arbeidsmarkedsregionen. Informasjonen på antall arbeidsledige og i arbeidsstyrken finner vi i

datasettet fra NAV. Variabelen *Ledighet_id* er regional arbeidsledighetsrate i januar for det inneværende året. Vi har noen flere observasjoner for arbeidsledighet ved å bruke inndelingen av arbeidsmarkeder enn ved kommuner. Grunnen til dette er at kommuner som av ulik årsak ikke har registrert kommunal arbeidsledighetsrate vil bli tilegnet den regionale ledigheten for sin region. Gjennomsnittlig regional arbeidsledighet er 3,24%. Det er ingen regioner som har null arbeidsledighet, og vi har at den laveste arbeidsledighetsraten er på 0,66%. Den høyeste regionale arbeidsledigheten er 11,16%. Bruken av arbeidsmarkedsregioner gjør at vi kan undersøke om arbeidsledighetsraten i kommunene rundt har betydning for lærermangel på skolene.

Tabell 13 viser resultater fra estimeringen gjort på skole- og kommunenivå. Estimert effekt av regional arbeidsledighet er signifikant negativ på skolenivå. Én prosent økning i den regionale arbeidsledigheten estimeres til å redusere lærermangel med 0.35 prosentpoeng. Kontrollert for skole- og kommunekaraktistikker er effekten av regional arbeidsledighet signifikant på et 5% signifikansnivå. Videre finner vi at arbeidsmarkedsregional ledighet på kommunenivå gir ikke-signifikant negativ effekt. Effektene av kontrollvariablene er tilsvarende til estimeringene gjort i modell (8) i tabell 8 på skolenivå og i tabell 10 på kommunenivå.

Tabell 13: Analysen med bruk av regional arbeidsledighet

Metode Variabler	SKOLENIVÅ		KOMMUNENIVÅ	
	(1) (FE) Lærermangel_s	(2) (FE) Lærermangel_s	(3) (FE) Lærermangel_k	(4) (FE) Lærermangel_k
lnLedighet_id	-0,456*** (0,164)	-0,349** (0,158)	-0,439 (0,291)	-0,382 (0,278)
Barneskole		0,466* (0,243)		
Ungdomsskole		-0,607* (0,339)		
lnElever0110_sum		0,052 (0,191)		2,125* (1,167)
Spesandel_verk0110		0,030*** (0,004)		0,051*** (0,013)
Eldre		-0,109** (0,054)		-0,187** (0,092)
Arbeidsalder		-0,165*** (0,054)		-0,225** (0,092)
logBefolkningstall		-1,996*** (0,615)		-4,230*** (1,548)
Barneskole_andel				0,015*** (0,005)
Ungdomsskole_andel				-0,005 (0,012)
Tidsfaste effekter	Ja	Ja	Ja	Ja
Skolefaste effekter	Ja	Ja	-	-
Kommunefaste effekter	-	-	Ja	Ja
Kluster	Ja	Ja	Ja	Ja
Observasjoner	77 484	77 367	10 760	10 760
R-kvadrert	0,039	0,043	0,118	0,132
Antall skoler	3 938	3 938		
Antall kommuner			436	436

Standardavvik i parentes
 *** p<0,01, ** p<0,05, * p<0,1

Analysen ved bruk av regional arbeidsledighet gir den forventede negative signifikante effekten av arbeidsledighet på lærermangel på skolenivå. At effekten er negativ og signifikant samsvarer med funnene i Falch et al. som brukte et regionalt mål på arbeidsledighet. Det er derfor interessant å undersøke estimeringer med denne arbeidsledighetsvariabelen ytterligere.

6.4 Heterogenitet

Heterogenitet referer til forskjeller eller variasjon mellom enheter i et utvalg. Vi ønsker å undersøke for støy innad i utvalget ved å gjennomføre analyse på ulike deler. Det er viktig å ta hensyn til heterogenitet når det gjøres analyser og resultater tolkes, da det kan påvirke estimater og statistisk signifikans.

I likhet med delkapittel 6.3 blir alle analysene i dette delkapittelet gjort med den enkleste modellen i hovedanalysen, tilsvarende modell (5), og modellen som inkluderer både skole- og kommunekaraktistika, tilsvarende modell (8). Vi bruker estimeringsmetoden FE, kontrollerer for tidsfaste effekter og bruker klustrede standardavvik på både skole- og kommunenivå. Heller ikke her inkluderer vi forsinkelse i arbeidsledigheten av samme grunn som tidligere.

Utvalget deles opp på ulike måter, og vi starter med å dele tidsperioden i to, 1995-2007 og 2008-2019. Vi deler utvalget på samme måte også når regional arbeidsledighet benyttes. Deretter deler vi utvalget ved å skille mellom de tre skoletypene, og utfører separate analyser for hver av dem individuelt.

6.4.1 Dele i to tidsperioder

Figur 3 i delkapittel 4.3 illustrerte trenden i arbeidsledighetsratene og den aggregerte lærermangelen. I figuren ser det ut som det er et negativt forhold mellom lærermangel og arbeidsledighet i den første delen av periode, mens det i den siste delen ikke er like tydelig. Vi vil derfor undersøke hvordan resultatet blir når vi deler opp analyseperioden i to tidsperioder, 1995-2007 og 2008-2019. På denne måten får vi to like lange perioder. I tillegg ser vi at tidsperiodene samsvarer med trendene i figur 3. Med denne inndelingen kan vi også undersøke om vi finner tilsvarende funn som Falch et al., hvor analysen varte frem til 2002, når vi sentrerer analysen rundt samme tidsperiode.

Tabell 14 viser resultatene på skolenivå for hver av periodene, mens tabell 15 viser tilsvarende resultater på kommunenivå. Vi finner signifikant negativ effekt av ar-

beidsledighet på lærermangel for den første perioden på både skole- og kommunenivå. Kommenterer først skolenivå og deretter resultatene for kommunenivå.

For perioden 1995-2007 finner vi at effekten av arbeidsledighet reduserer lærermangel med 0.334 prosentpoeng ved én prosent økning i arbeidsledighet, kontrollert for skole- og kommunekaraktistikker. Effekten er signifikant på et 5% signifikansnivå. I den andre perioden er effekten derimot positiv og ikke-signifikant.

Tabell 14: Delt i to perioder på skolenivå

SKOLENIVÅ	Første periode		Andre periode	
	(1)	(2)	(3)	(4)
Periode	1995-2007	1995-2007	2008-2019	2008-2019
Variabler	Lærermangel _s	Lærermangel _s	Lærermangel _s	Lærermangel _s
lnArbeidsledighet_NAV	-0,346** (0,167)	-0,334** (0,166)	-0,100 (0,157)	0,078 (0,159)
Barneskole		0,075 (0,356)		0,090 (0,379)
Ungdomsskole		0,705 (0,596)		-0,981 (0,691)
lnElever0110_sum		0,501* (0,278)		0,403 (0,325)
Spesandel_verk0110		0,035*** (0,005)		0,024*** (0,006)
Eldre		-0,429*** (0,084)		0,027 (0,125)
Arbeidsalder		-0,496*** (0,076)		0,095 (0,112)
lnBefolkningstall		-0,131 (1,073)		-6,909*** (1,528)
Tidsfaste effekter	Ja	Ja	Ja	Ja
Skolefaste effekter	Ja	Ja	Ja	Ja
Kluster	Ja	Ja	Ja	Ja
Observasjoner	42 418	42 360	34 940	34 881
R-kvadrert	0,067	0,073	0,011	0,015
Antall skoler	3 653	3 652	3 398	3 398

Standardavvik i parentes

*** p<0,01, ** p<0,05, * p<0,1

Av kontrollvariablene over finner vi kun effekten av andelen spesialundervisning som signifikant over begge perioder. Denne er som forventet positiv. For den første perioden er andelen i arbeidsrelevant alder opp mot andelen barn signifikant, mens for den andre perioden er det befolkningstallet som er signifikant. På skolenivå har vi flere observasjoner for den første perioden på grunn av en liten reduksjon i antall skoler i den andre perioden.

På kommunenivå gir resultatene i tabell 15 en negativ signifikant effekt av arbeidsledighet i den første perioden. Dersom arbeidsledigheten øker med én prosent reduserer det lærermangel med 0.44 prosentpoeng, kontrollert for skole- og kommunekaraktetrikker. Tolker koeffisienten likt som Falch et al. gjør i sin analyse med prosentpoengs økning i arbeidsledighet. Gitt et gjennomsnittlig ledighetsnivå på 3,15%, gir en økning i arbeidsledighetsraten med ett prosentpoeng en reduksjon i lærermangel med 0,15 prosentpoeng. I likhet med analysen på skolenivå gir den siste perioden, 2008-2019, en ikke-signifikant positiv effekt. Å dele opp analysen i to deler gir et betydningsfullt innblikk i endringer som kan ha skjedd både tidligere og i løpet av tidsperioden til analysen.

Tabell 15: Delt i to perioder på kommunenivå

KOMMUNENIVÅ	Første periode		Andre periode	
	(1)	(2)	(3)	(4)
Periode	1995-2007	1995-2007	2008-2019	2008-2019
Variabler	Lærermangel_k	Lærermangel_k	Lærermangel_k	Lærermangel_k
lnArbeidsledighet_NAV	-0,488** (0,237)	-0,444* (0,230)	0,085 (0,202)	0,161 (0,214)
Barneskole_andel		0,011 (0,007)		0,013* (0,007)
Ungdomsskole_andel		0,003 (0,013)		-0,031** (0,013)
lnElever0110_sum		5,927*** (1,448)		0,339 (2,145)
Spesandel_verk0110		0,038*** (0,015)		0,065*** (0,018)
Eldre		-0,198 (0,131)		-0,436** (0,177)
Arbeidsalder		-0,334** (0,131)		-0,311** (0,155)
lnBefolkningstall		-2,846 (1,942)		-6,517* (3,659)
Tidsfaste effekter	Ja	Ja	Ja	Ja
Kommunefaste effekter	Ja	Ja	Ja	Ja
Kluster	Ja	Ja	Ja	Ja
Observasjoner	5 602	5 602	5 065	5 065
R-kvadrert	0,219	0,245	0,044	0,061
Antall kommuner	434	434	432	432

Standardavvik i parentes

*** p<0,01, ** p<0,05, * p<0,1

Effekten av kontrollvariablene varierer mellom de to periodene. I den første perioden er effekten av skolestørrelse gitt ved elevtall signifikant og gir stor utslagseffekt. Dersom elevtallet i kommunen dobber, vil det føre til en økning i lærermangel med

5,92 prosentpoeng. I den andre perioden finner vi derimot ingen signifikant effekt av skolestørrelse, men av en annen størrelsesvariabel. I perioden mellom 2008 og 2019 er det størrelsen på befolkningen i kommunen som gir signifikant effekt. Dersom størrelsen på befolkningen dobles, estimeres det at lærermangelen vil reduseres med 6,5 prosentpoeng.

6.4.2 To perioder med arbeidsmarkedsregional ledighet

Analysen over hele tidsperioden for paneldatasettet gjort med arbeidsmarkedsregional ledighet i delkapittel 6.3.3 gav signifikant negativ effekt på lærermangel. Vi ønsker undersøke hvordan effekten av regional ledighet er dersom vi deler opp i to perioder, tilsvarende som i delkapittel 6.4.1.

Tabell 16 viser at det er forskjell på effekten og signifikansen av arbeidsledighet i arbeidsmarkedsregionen i de to periodene, tilsvarende funn som ved den kommunale arbeidsledigheten fra NAV. I den første perioden fra 1995-2007 gir økning i den regionale arbeidsledigheten med én prosent en reduksjon i lærermangel på 0,77 prosentpoeng på et 1% signifikansnivå. Tolkning ved ett prosentpoengs økning i arbeidsledighet, og med et regionalt ledighetsnivå på 3,24%, reduseres lærermangel med 0,24 prosentpoeng. Vi finner også den samme endringen i signifikans i de to periodene når analysen gjøres på kommunenivå, se tabell A3 i appendiks.

Tabell 16: Delt i to perioder på skolenivå med regional arbeidsledighet

SKOLENIVÅ	Første periode		Andre periode	
	(1)	(2)	(3)	(4)
Periode	1995-2007	1995-2007	2008-2019	2008-2019
Variabler	Lærermangel_s	Lærermangel_s	Lærermangel_s	Lærermangel_s
lnLedighet_id	-0,785*** (0,222)	-0,768*** (0,221)	-0,155 (0,212)	0,077 (0,212)
Barneskole		0,082 (0,356)		0,089 (0,378)
Ungdomsskole		0,695 (0,596)		-0,996 (0,691)
lnElever0110_sum		0,501* (0,278)		0,377 (0,327)
Spesandel_verk0110		0,035*** (0,005)		0,023*** (0,006)
Eldre		-0,421*** (0,084)		0,023 (0,125)
Arbeidsalder		-0,496*** (0,075)		0,102 (0,112)
lnBefolkningstall		-0,116 (1,070)		-6,680*** (1,551)
Tidsfaste effekter	Ja	Ja	Ja	Ja
Skolefaste effekter	Ja	Ja	Ja	Ja
Kluster	Ja	Ja	Ja	Ja
Observasjoner	42 460	42 402	35 024	34 965
R-kvadrert	0,067	0,073	0,011	0,014
Antall skoler	3 653	3 652	3 398	3 398

Standardfeil i parentes

*** p<0,01, ** p<0,05, * p<0,1

6.4.3 Skille mellom skoletyper

Gjennom hovedanalysen så vi at det var ulik effekt av barneskole og ungdomsskole i forhold til kombinerte skoler på lærermangel. Vi ønsker derfor å skille på lærermangel og effekt av arbeidsledighetsrate på de ulike skoletypene for å undersøke for variasjon mellom skoletypene. Ser da på effekten av arbeidsledighet og kontrollvariabler på lærermangel registrert på barneskole, ungdomsskole og de kombinerte skolene hver for seg. Skolekarakteristikene som brukes vil gjelde for de aktuelle skoletrinnene som tilsvarer skoletypen.

Det er 2 609 barneskoler, 685 ungdomsskoler og 1 276 skoler som er kombinerte.

Barneskoler utgjør den største andelen av skoler i Norge. Effekten av arbeidsledighet på lærermangel på de ulike skoletypene er ikke-signifikant, se tabell A4 i appendiks. Det er de kombinerte skolene som får utslag for en negativ effekt av arbeidsledighet når alle kontrollvariabler er inkludert. Resultatet tolkes som at det ikke er noe signifikant forskjell mellom skoletypene og hvordan lærermangel påvirkes av arbeidsledighet.

7 Avslutning

I denne masteroppgaven har vi undersøkt hvilken effekt arbeidsledighet har på graden av lærermangel i grunnskolen. Oppgaven er et forsøk på å oppdatere tidligere analyse av Falch, Strøm og Johansen i 2009. Vi har benyttet oss av et paneldatasett for norske kommuner og skoler i tidsperioden 1995-2019 for å gjennomføre analysen. Paneldatasettet er satt sammen av data som vi har hentet fra ulike kilder. Skoleinformasjon har vi fra Grunnskolens Informasjonssystem (GSI). Kommunestatistikk og arbeidsledighet er hentet fra "Local Government Dataset" fra Fiva, Halse og Natvik, i tillegg til at vi har data på månedlig kommunal arbeidsledighet fra NAV.

Hypotesen er at en nedgang i arbeidsledighetsraten resulterer i økt lærermangel. Dette kan forklares ved at lærere, under gitte lønnsbetingelser, vil finne andre jobber utenfor skolesystemet når arbeidsmarkedet er stramt. Lærermangel har vi definert som andelen årsverk gjennomført av lærere som ikke oppfyller kravene til formell kompetanse. Vi har gjennomført analysen med arbeidsledighetsraten i kommunene, og skilt mellom skolenivå og kommunenivå for å fange opp skole- og kommuneeffekter. Vi begynte med en helt enkel modellspesifikasjon hvor kun fokusvariabelen, arbeidsledighet, var inkludert. Modellen ble estimert ved bruk av forskjellige metoder, og ved hjelp av ulike tester valgte vi å fortsette analysen med faste effekter (FE) som hovedestimeringsmetode. Likevel valgte vi også å estimere utvalgte modeller med sammenslåtte minste kvadreters metode (POLS) videre i analysen. Dette gjorde vi for å illustrere sammenhenger i data som eventuelt har gått tapt når estimeringsmetoden med FE benyttes. I hovedanalysen finner vi ikke den forventede signifikante negative effekten av arbeidsledighet på graden av lærermangel, verken på skole- eller kommunenivå.

Videre inkluderte vi kontrollvariabler for skole- og kommunekarakteristika, hvor vi finner noen av dem signifikante. Blant annet gir andelen spesialundervisning signifikant økning i lærermangel. Befolkningstall og alderssammensetning er også å finne som signifikante. At kommunens populasjonssammensetning har innvirkning på graden av lærermangel kan tolkes som at tilgang på arbeidskraft reduserer lærermangel. En høyere andel eldre i befolkningen i forhold til barn i skolealder kan indikere et redusert behov for lærere.

For å undersøke påliteligheten til resultatene våre har vi gjennomført en robusthetsanalyse. Først benyttet vi et annet mål på arbeidsledighet, ved at vi så på årsgjennomsnittlig rate i kommunene i stedet for arbeidsledigheten i januar. Vi finner at begge arbeidsledighetsvariablene gir ikke-signifikant effekt. Det andre vi gjorde var å benytte et annet mål på lærermangel, gitt ved andelen av det totale antallet lærere som ikke hadde formell kompetanse. Vi vurderer det til at det gir mindre

pålitelige resultater, fordi tidsperioden blir redusert som følge av at antall lærere har blitt rapportert i GSI i færre år.

Deretter utvidet vi inndelingen av hvor arbeidsledigheten gjelder til å omfatte arbeidsmarkedsregionene. Her finner vi den forventede negative effekten som signifikant på skolenivå. At vi fikk signifikant effekt gir oss grunn til å tro at arbeidsmarkedet i kommunene rundt har betydning for lærermangel på skolene. Læreryrket er et geografisk fleksibelt yrke, i og med at det er behov for lærere på skoler i alle kommuner. Det gjør at lærere kan velge å arbeide i en nabokommune og pendle, slik at arbeidsmarkedsregional ledighet er relevant.

Trenden med negativt forhold mellom arbeidsledighet og lærermangel var noe tydelig i figur 4 i kapittel 4.3. Ved å dele opp analysen i to tidsperioder, 1995-2007 og 2008-2019, finner vi at arbeidsledigheten har signifikant negativ effekt i den første perioden. Effekten er signifikant både når arbeidsledighet måles i kommunene og arbeidsmarkedsregionene. Det indikerer at det er noe som har skjedd i løpet av analyseperioden og i forhold til tidligere funn. Kontrollert for ulike skoletyper finner vi at det ikke er noe forskjell på effekten av arbeidsledighet og lærermangel knyttet til ulike skoletyper.

7.1 Endring fra tidligere funn

Denne oppgaven har basert seg på Falch, Strøm og Johansen (2009) sin analyse på lærermangel og konjunktursvingninger over tidsperioden 1981-2002. Definisjonen vi har brukt på lærermangelmålet er lik deres, men vår analyse er basert på årsverk i motsetning til Falch et al. (2009) som kun så på lærere i heltidsstilling. Ved å bruke årsverk inkluderer vi også deltidsansatte uten formell lærerkompetanse. Det er rimelig å anta at ukvalifiserte lærere oftere blir ansatt i deltidsstillinger sammenlignet med kvalifiserte lærere. Ved å inkludere deltidsansatte i målet på lærermangel får vi dermed et mer realistisk bilde av situasjonen og et bedre mål for å fange opp den faktiske effekten av arbeidsledighet. Mens Falch et al. brukte regional arbeidsledighet har vi i hovedsak brukt kommunal arbeidsledighet.

Vi finner negativ signifikant effekt i tidsperioden 1995-2007, hvorav 7 av årene overlapper de de siste årene fra Falch et al. sin analyse. De fant at om arbeidsledighet øker med ett prosentpoeng ble det estimert til å redusere lærermangel med 0,4 prosentpoeng. Vår analyse gir at én prosent økning i arbeidsledighet reduserer lærermangel med 0,44 prosentpoeng. Tolket likt som Falch et al., med en økning på ett prosentpoeng i arbeidsledighet, finner vi en reduksjon med 0,15 prosentpoeng på lærermangel. Effekten vi har funnet er mindre enn hva Falch et al. fant. Videre gir vår analyse at

effekten av arbeidsledighet er ikke-signifikant i den andre perioden, 2008-2019.

Analysen vår over hele tidsperioden estimerer ikke-signifikant effekt av kommunal arbeidsledighet. Når vi brukte arbeidsmarkedsregional ledighet, finner vi signifikant negativ effekt på skolenivå. Ved å dele analysen i to får vi, tolket likt som Falch et al., en reduksjon på 0,24 prosentpoeng i lærermangel ved en ett prosentpoengs økning i regional ledighet på skolenivå. Også her forsvinner effekten i den andre perioden. Dette støtter opp om hovedfunnene våre om at effekten av arbeidsledighet på lærermangel i løpet av tidsperioden har blitt borte.

7.1.1 Endringer i perioden

I delkapittel 6.4 delte vi opp datasettet i to tidsperioder, 1995-2007 og 2008-2019. Da fant vi at det er ulik effekt av arbeidsledighet på lærermangel i løpet av tidsperioden vi undersøker. I tidsperioden som Falch et al. undersøkte forhandlet lærere lønn med staten. Fra 2004 gikk det over til å være lønnsforhandlinger med kommunen, KS. En mulig årsak for at effekten forsvinner kan være at kommunens mulighet for lokal lønnsrespons absorberer den tidligere effekten av arbeidsledighet på lærermangel. Samtidig har vi brukt tidsfaste effekter for å korrigere for nasjonale variasjoner.

Arbeidsledighetsraten har over store deler av tidsperioden vært relativt stabil. Figur 4 i kapittel 4.3 viste at de største variasjonene i arbeidsledigheten skjedde i årene før 2010. Bevegelsene etter 2010 varierer mindre i absolutt størrelse. En stabil arbeidsledighetsrate kan ha ført til at lærere har blitt mindre opptatt av markedssituasjonen i valget om de skal forbli i yrket, fordi det har vært en mindre varierende faktor. I likhet med Falch og Strøm (2005) fant også vi at ikke-økonomisk faktorer, som andelen spesialundervisning, har signifikant effekt på lærermangel.

Vi har estimert en gjennomsnittseffekt for alle skoler og kommuner. Vi har vist at det er stor variasjon i lærermangel mellom kommuner, og utviklingen i lærermangel og arbeidsledighetsratene er ikke like. Det kan være sammenhenger noen steder selv om det ikke er slik i gjennomsnitt. Spesielt grafene for Oslo kommune viser klar negativ korrelasjon mellom arbeidsledighet og lærermangel, særlig i den første delen av tidsperioden.

7.2 utfordringer med oppgaven

Vi har hatt noen utfordringer med oppgaven. For det første har vi et potensielt problem med analysen som omhandler utelatte variabler. I datasettet har vi variabler for kommunens budsjett og for hvor stor andel som går til skoleformål. Å inkludere en slik variabel kunne kontrollert for kommunens respons og prioritering av skoler og lærere. Budsjettering har likevel blitt utelatt da det kan gi endogenitetsproblemer. Det kommer av at kommunens budsjettprioritering kan påvirkes av om det eksisterer lærermangel i kommunen. Dermed ville ikke budsjettvariabler vært uavhengig og eksogent gitt, men avhengig av den avhengige variabelen i analysen.

En annen utfordring handler om mangel på sensitive variabler fra GSI. Vi har i analysen kunnet kontrollere for andelen av undervisningen som går til spesialundervisning, som vi har brukt som en indikator på mer utfordrende undervisningsforhold ved skolene. I GSI registrerer rektorene også hva spesialundervisningen går til. Å inkludere disse variablene ville gitt en enda mer presis kontrollering for spesialforhold ved skoler. Dette datamaterialet ligger ikke offentlig ute på skolenivå, begrunnet med at det er sensitiv data. Til tross for forsøk med å søke om dispensasjon for taushetsplikt, fikk vi ikke dette materialet til vår oppgave. Annen sensitiv data fra GSI, som vi ikke fikk, var elevenes etnisitet. En videre effekt av det er at vi ikke har kunnet kontrollere for skolenes behov for morsmåls- og norskopplæring. Vi synes det er synd at data fra utdanningsdirektoratet, med formål om å brukes til forskning, ikke ble gjort tilgjengelig for oss.

I tillegg er registreringen i GSI gjort av rektorene ved skolene. Det åpner opp for at det kan skje feilrapportering og dermed gi målefeil i datasettet. Figur 3 i kapittel 4.2 viste tegn til feilrapportering i de første årene i forbindelse med årsverk. Det har vi utelatt ved at vi startet analysen noen år senere. Det er likevel en mulighet for at det har skjedd flere feilrapporteringer i løpet av årene i analysen som vi ikke har fanget opp.

En siste utfordring handler om klustringsnivået. Når OLS benyttes på tverrsnittsdata, er standarden en antagelse om at standardavvikene er identisk og uavhengig distribuert. Dette er ikke gjeldende når vi har paneldata. Når vi benytter paneldata må vi derfor ta hensyn til mulige kompliserte standardavvikstrukturer innenfor enhetene. Vi rapporterte derfor klustrede standardavvik også når modellen ble estimert med faste effekter (FE). En mulig utfordring oppstår når vi benytter (FE) på skolenivå, samtidig som vi ønsker å rapportere klustrede standardavvik på kommunenivå. Utfordringen skyldes en mulig korrelasjon mellom observasjonene innenfor samme kommune som ikke fanges opp når vi grupperer på skolenivå. I slike tilfeller eksisterer det en risiko for at standardavvikene undervurderes og resultere i feilaktig signifikans. For å

håndtere denne korrelasjonen er det teoretisk mulig å gruppere på det høyere nivået, altså på kommunenivå, samtidig som vi kontrollerer for skolefaste effekter. Imidlertid er ikke dette teknisk gjennomførbart i det dataprogrammet vi benytter for å estimere regresjonene (STATA).

7.3 Konklusjon

Vi har undersøkt effekten av arbeidsledighet på graden av lærermangel i grunnskolen i tidsperioden 1995-2019. Lærermangel er definert som andelen årsverk utført av lærere uten formell kompetanse. Fra vår analyse kan vi konkludere med at arbeidsledigheten i kommunene og i arbeidsmarkedsregionene gir signifikant negativ effekt på lærermangel i tidsperioden 1995-2007. Vi finner omtrent halvparten av effekten som ble funnet av Falch et al. (2009). Effekten av arbeidsledighet er ikke lengre signifikant i 2008-2019. Denne endringen tyder på at noe har skjedd siden Falch et al. gjorde sin analyse. De fant at bedre arbeidsmarked utenfor skolen, målt ved arbeidsledighet, økte lærermangelen. Vi finner at denne effekten er blitt borte.

Referanser

Andersson, C., Johansson, P. & Waldenström, N. (2010). Do you want your child to have a certified teacher? *Economics of Education Review*, 30(1). 65-78.

Bacolod, M. P.. (2007). Do alternative opportunities matter? The role of female labor markets in the decline of teacher quality. *The Review of Economics and Statistics*, 89(4). 737-751.

Boardman, A. E., Darling-Hammond, L. & Mullin, S. P. (1982). A framework for the analysis of teachers' demand and supply. *Economics of Education Review*, 2(2). 127-155.

Chevalier, A., Dolton, P. & McIntosh, S. (2005). Recruiting and Retraining Teachers in the UK: An Analysis of Graduate Occupation Choice from the 1960s to the 1990s. *Economics*, 74(293). 69-96.

Falch, T., Johansen, K. & Strøm, B. (2009) Teacher shortage and the business cycle. *Labor Economics*, 16, 648-658.

Falch, T. & Strøm, B. (2005). Teacher turnover and non-pecuniary factors. *Economics of Education Review*, 24., 611-631.

Fiva, J. H., Halse, A. & Natvik, G. J. (2020) Local Government Dataset. Tilgjengelig på <https://www.jon.fiva.no/data.htm>

Fraenkel, R. C. (2022) Local labor markets and job match quality: Teachers. *Labour Economics*, 78.

Goldhaber, D. & Theobald, R. (2022). Teacher Attrition and Mobility Over Time. *Educational Researcher*, 51(3). 235 - 237.

Grunnskolen Informasjonssystem. Hentet fra <https://gsi.udir.no/>

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review*, 30(2011). 466-479.

NAV. (2022). *Historisk statistikk om arbeidsmarkedet*. Tilgjengelig på <https://www.nav.no/no/nav-og-samfunn/statistikk/arbeidssokere-og-stillinger-statistikk/historisk-statistikk>

- Opplæringslova. (1998). *Lov om grunnskolen og den vidaregåande opplæringa* (LOV-1998-07-17-61). Lovdata. Hentet 24.februar 2023 fra <https://lovdata.no/dokument/NL/lov/1998-07-17-61>
- Regjeringen. (2022). *Evaluering av seksårsreformen*. <https://www.regjeringen.no/no/aktuelt/evaluering-av-seksarsreformen/id2950959/>
- Regjeringen. (2022). *Tvungen lønnsnemnd i lærerstreiken*. <https://www.regjeringen.no/no/aktuelt/Tvungen-lonnsnemnd-larerstreiken/id2929135/>
- Rivkin, S. G., Hanushek, E. A. & Kain, J. F. (2005). Teachers, Schools, and Achievement. *Econometrica*, 73(2). 417 - 458.
- Senter for økonomisk forskning. (2021) *Ulikheter i lærerstatistikk* (SØF-rapport 4/21). <https://samforsk.no/uploads/files/Ulikheter-i-laererstatistikk-WEB.pdf>
- St.meld. nr.33 (2002-2003). *Om ressursituasjonen i grunnopplæringen m.m.* Utdannings- og forskningsdepartementet. <https://www.regjeringen.no/contentassets/a5900f92b95f4d58b07cde941e5bb56/no/pdfs/stm200220030033000dddpdfs.pdf>
- Stokke, T. A., Stoltz, G. & Barth, E. (2023, 5.januar). Arbeidsledighet. I *Store norske leksikon*. Hentet 24.februar 2023 fra <https://snl.no/arbeidsledighet>
- Sparrman, V. (2012). Arbeidsledighet som konjunkturindikator og forklaringsfaktor i makromodeller. *Økonomiske analyser*, 2012(5) https://www.ssb.no/a/publikasjoner/pdf/oa_201205/sparrman.pdf
- UNESCO. (2015). *Education 2030: Incheon Declaration and Framework for Action for the implementation of Sustainable Development Goal 4*. https://uis.unesco.org/sites/default/files/documents/education-2030-incheon-framework-for-action-implementation-of-sdg4-2016-en_2.pdf
- Utdanningsforbundet. (2021, 26.februar). *Femårige lærerutdanninger og undervisningskompetanse*. <https://www.utdanningsforbundet.no/medlemsgrupper/universitet-og-hogskole/ny-master-i-grunnskolelærerutdanning-og-undervisningskompetanse/>
- Verbeek, M. (2017). *A guide to modern econometrics* (Fifth Edition). Wiley.
- Wooldrige, J. M. (2019). *Introductory Econometrics. A modern Approach* (Seventh Edition). Cengage Learning, Inc.

A Appendiks

A.1 Deskriptiv statistikk

Tabell A1: Deskriptiv statistikk for annet mål på lærermangel

Variabelnavn	Variasjon	Gjennitt	Standardavvik	Min	Maks	Observasjoner
LM_andelssats_s	Overordnet	6,980	10,73	0	100	N = 14 154
	Mellom		9,251	0	100	n= 2 980
	Innenfor		6,155	-53,02	69,98	T-bar = 4,750
LM_andelssats_k	Overordnet	7,368	7,087	0	62,50	N = 2 126
	Mellom		5,942	0	40,69	n = 430
	Innenfor		3,844	-15,44	37,97	T-bar = 4,944

Tabell A2: Deskriptiv statistikk for regional arbeidsledighet

Variabelnavn	Variasjon	Gjennitt	Standardavvik	Min	Maks	Observasjoner
Ledighet_id_s	Overordnet	3,243	1,226	0,658	11,16	N = 77 529
	Mellom		0,860	0,966	10,55	n= 3 940
	Innenfor		0,973	-0,175	9,060	T-bar = 19,677

A.2 Tabeller

Tabell A3: Delt i to perioder på kommunenivå med regional arbeidsledighet

KOMMUNENIVÅ	Første periode		Andre periode	
	(1)	(2)	(3)	(4)
Metode	1995-2007		2008-2019	
Periode	1995-2007		2008-2019	
Variabler	Lærermangel_k	Lærermangel_k	Lærermangel_k	Lærermangel_k
lnLedighet_id	-0,766** (0,341)	-0,721** (0,340)	-0,070 (0,424)	0,012 (0,425)
Barneskole_andel		0,011 (0,007)		0,014** (0,007)
Ungdomsskole_andel		0,004 (0,013)		-0,031** (0,013)
lnElever0110_sum		5,889*** (1,523)		1,508 (2,104)
Spesandel_verk0110		0,037** (0,015)		0,064*** (0,020)
Eldre		-0,166 (0,139)		-0,350* (0,191)
Arbeidsalder		-0,343*** (0,132)		-0,205 (0,157)
lnBefolkningstall		-3,045 (1,906)		-7,233** (3,670)
Tidsfaste effekter	Ja	Ja	Ja	Ja
Kommunefaste effekter	Ja	Ja	Ja	Ja
Kluster	Ja	Ja	Ja	Ja
Observasjoner	5 631	5 631	5 129	5 129
R-kvadrert	0,214	0,239	0,040	0,057
Antall kommuner	434	434	432	432

Standardfeil i parentes

*** p<0,01, ** p<0,05, * p<0,1

Tabell A4: Skille på alle skoletypene

Metode Variabler	BARNESKOLE		UNGDOMSSKOLE		KOMBINERTSKOLE	
	(1) (FE)	(2) (FE)	(3) (FE)	(4) (FE)	(5) (FE)	(6) (FE)
lnArbeidsledighet_NAV	-0,056 (0,154)	0,017 (0,153)	0,041 (0,260)	0,043 (0,252)	-0,141 (0,217)	-0,025 (0,216)
lnElever0107_sum		0,115 (0,243)				
Spesandel_verk0107		0,038*** (0,005)				
Eldre		-0,053 (0,072)		-0,249** (0,114)		-0,120 (0,113)
Arbeidsalder		-0,073 (0,071)		-0,248** (0,116)		-0,230** (0,105)
lnBefolkningstall		-2,090** (0,845)		-1,014 (1,325)		-2,810** (1,431)
lnElever0810_sum				0,470 (0,411)		
Spesandel_verk0810				0,011 (0,008)		
lnElever0110_sum						0,336 (0,466)
Spesandel_verk0110						0,031*** (0,007)
Tidsfaste effekter	Ja	Ja	Ja	Ja	Ja	Ja
Skolefaste effekter	Ja	Ja	Ja	Ja	Ja	Ja
Kluster	Ja	Ja	Ja	Ja	Ja	Ja
Observasjoner	45 783	45 777	12 447	12 424	19 128	19 039
R-kvadrert	0,040	0,043	0,032	0,035	0,051	0,056
Antall skoler	2 609	2 607	685	685	1 276	1 275

Standardavvik i parentes
 *** p<0,01, ** p<0,05, * p<0,1

