

Sensitivitet for boliglånsrenter: Hvilke kundegrupper har størst risiko for avslutning ved en renteøkning?

Ole Hjørnegård Hannestad

Masteroppgave

Masteroppgaven er levert for å fullføre graden

Master i samfunnsøkonomi

Universitetet i Bergen, Institutt for økonomi

Februar 2021



UNIVERSITETET I BERGEN

Forord

Masteroppgaven setter punktum for min utdanning ved Universitetet i Bergen og har gitt meg muligheten til å fordype meg i et tema som ikke bare er svært interessant for min egen del, men som jeg også kan dra stor nytte av i arbeidslivet i årene som kommer. Det har vært en viktig erfaring å få arbeide selvstendig med en oppgave som har bydd på mange utfordringer underveis ettersom jeg har kombinert oppgaveskriving med fullt arbeid i Sparebanken Øst og en baby på hjemmebane.

Bakgrunnen for valg av tema er at jeg alltid har hatt et ønske om å arbeide innenfor bankbransjen. Da jeg ble ansatt i en 100% stilling i Sparebanken Øst i 2018 spurte jeg tidlig om det var muligheter for å skrive en masteroppgave som baserer seg på en problemstilling banken kunne dra nytte av samtidig som jeg kunne få tilgang til datamateriale som ikke er tilgjengelig for alle. Banken har gjennom hele prosessen vært positiv til oppgaven og bidratt med forslag til problemstillinger.

Jeg ønsker å rette en stor takk til min veileder Hans K. Hvide for gode råd, tilbakemeldinger og hans store dose tålmodighet. Jeg vil også takke Sparebanken Øst for tilgang til relevant datamateriale samt deres fleksibilitet slik at jeg har hatt muligheten til å fullføre oppgaven i en travel arbeidshverdag. Til slutt ønsker jeg å takke min kone Linda-Louise som har vært en fantastisk støttespiller og tilrettelegger slik at jeg har fått muligheten til å bruke mange sene kvelder med masterarbeid.

Sammendrag

Formålet med denne masteroppgaven har vært å gjennomføre en analyse av hvilke boliglånskunder som er mest sensitive for en renteøkning og har derfor høyere sannsynlighet for å avslutte boliglånet påfølgende måned etter renteøkningen. Oppgaven benytter en lineær sannsynlighetsmodell med interaksjonsvariabler som predikerer sannsynligheten for at en kunde avslutter boliglånet påfølgende måned, med utgangspunkt i kundene til Sparebanken Øst. Oppgaven presenterer to modeller som fokuserer på mersannsynligheten for avslutning av boliglån etter en renteøkning, der den første utelukkende ser på lånestørrelser som forklaringsvariabler og den andre kontrollerer for sosialdemografiske variabler.

4727 boliglånsengasjement ble observert i tre måneder etter en renteøkning i 2017 for å se hvilke kunder som avsluttet boliglånene sine. Oppgaven sammenligner en treatmentgruppe som ble utsatt for en renteøkning mot en kontrollgruppe som antas ikke påvirkes av renteøkningen.

Resultatene varierer i signifikans der resultatene for lånestørrelse blir påvirket av at kundegruppen deles i ni persentilgrupper der de andre variablene deles i tre, noe som reduserer antall observasjoner for hver gruppe. Likevel indikerer resultatene at det er de med størst boliglån som er mest sensitiv for renteøkning. Oppgaven finner også at alder er i negativ sammenheng med rentesensitivitet, selv om studien heller ikke finner signifikante resultater for dette. Studien finner derimot, noe overraskende, at særlig kundene som i oppgavens datasett representerer dem som har vært i banken kortere enn tre år og lengre enn ti år har signifikant høyere sannsynlighet for exit etter renteøkning enn resten. Oppgaven finner også en signifikant mersannsynlighet for kundene som tilhører lokalbanken, der det antas Sparebanken Øst har markedsrett. En kunde som verken tilhører lokalbanken eller Oslo doubler sannsynligheten for låneavslutning etter en renteøkning dersom kunden hadde tilhørt lokalbanken, alt annet likt.

Oppgaven bidrar til at en bedre kan forstå hvilke kunder som er særlig rentesensitive. Kim Klinger og Vale (2001) og deres estimering av byttekostnader innenfor banksektoren er nærliggende denne oppgavens problemstilling, men tar ikke for seg hvordan kunder reagerer på en renteøkning. Det at jeg i denne studien besitter et datasett som skiller på en treatmentgruppe og kontrollgruppe bidrar til å isolere renteffekten slik at man ser avslutningssannsynligheten når økt rente er hovedmotivasjonen for avslutning.

Innholdsfortegnelse

Sammendrag	3
1. Introduksjon	7
1.1 Bakgrunn for oppgaven	7
1.2 Problemstillingen.....	7
1.3 Kort oppsummering av eksisterende litteratur.....	8
1.4 Kort beskrivelse av data	9
1.5 Hva ønsker oppgaven å teste?	10
1.6 Sentrale antagelser.....	10
2. Bankmarkedet i Norge	11
2.1 Krav til nye utlån med pant i bolig.....	13
2.2 Kort om Sparebanken Øst	13
3. Teori	14
3.1 Eksisterende litteratur	14
3.2 Byttekostnader.....	16
4. Databeskrivelse	18
4.1 Datainnsamling og sortering.....	18
4.2 Modellens variabler	20
4.2.1 Avslutning	20
4.2.2 Variabel for måned	20
4.2.3 Variabel for lånestørrelse	22
4.2.4 Variabel for alder.....	22
4.2.5 Variabel for kundeforholdets lengde	22
4.2.6 Variabel for bostedsområde.....	23
5. Metode	23
5.1 Forskningsdesign.....	23
5.2 Økonometrisk metode	24
5.2.1 Lineær sannsynlighetsmodell	24
5.2.2 Interaksjonsvariabler	26
5.3 Modellens hensiktsmessighet	27
6. Økonometrisk analyse	28
6.1 Fremgangsmåte i Stata	28
6.2 Regresjonsanalyse	29
7. Modellen	30
7.1 Modell 1	30

7.2	Modell 2	32
7.3	Modell 3	33
7.4	Potensielle feil i modellen	34
	Skjevhet fra utelatte variabler	34
	Simultan kausalitet.....	34
	Reliabiliten til datamaterialet.....	35
8.	Resultater	36
8.1	Modell 1 – Det gjøres ikke forskjell på treatmentgruppe og kontrollgruppe	36
8.2	Modell 2 - Modell for mersannsynlighet for exit ved ulike lånestørrelser.	37
8.2.1	Tolkning av resultater	39
8.3	Modell 3 – Utvidet modell der det kontrolleres for sosialdemografiske forhold.	39
8.4	Valg av modell	41
8.4.1	Test for multikollinearitet.....	42
8.5	Drøfting av resultatene fra modell 3.....	43
8.5.1	Variabel for lånestørrelse	45
8.5.2	Variabel for alder	46
8.5.3	Variabel for lengde på kundeforhold.	47
8.5.4	Variabel for bostedsområde	48
9	Oppsummering og konklusjon	49
9.1	Hovedfunn	49
9.2	Kritikk, begrensninger og videre forskning.....	50
	Litteraturliste.....	51
	Appendiks	54

Figuroversikt

Figur 1: Fusjonering mellom sparebanker tilbake til 1923. Kilde: Sparebankforeningen (2019).....	11
Figur 2: Markedsandelene i det norske utlånsmarkedet 2017. Kilde: Finans Norge (2020).....	12
Figur 3: Viser predikert exitsannsynlighet for de ulike lånestørrelsegruppene i modell 2.....	38
Figur 4: Predikert mersannsynligheten for exit for treatmentgruppen sammenlignet med kontrollgruppen for alle variabler i modell 2	38
Figur 5: Viser predikert exitsannsynlighet for alle forklaringsvariabler i modell 3	40
Figur 6: Viser predikert mersannsynligheten for exit for treatmentgruppen sammenlignet med kontrollgruppen for alle variabler i modell 3	41
Figur 7: Viser predikert exitsannsynlighet for de ulike persentilnivåene når kunden verken tilhører lokal eller Oslo, men tilhører A2 og Lk2	45
Figur 8: Viser mersannsynligheten for exit for treatmentgruppen sammenlignet med kontrollgruppe, der kun lånestørrelse blir observert.	46
Figur 9: Viser predikert exitsannsynlighet for de to aldersgruppene som også tilhører gruppe P5, Lk2 og verken hører til lokal eller Oslo.....	46
Figur 10: Viser predikert exitsannsynlighet for de to gruppene for kundeforholdlengde. Kundene tilhører også gruppe P5, A2 og tilhører verken Oslo eller lokalbanken.	47
Figur 11: Viser predikert exitsannsynlighet for lokalbanken vs ikke-lokalbanken. Den predikerte exitsannsynligheten gjelder i tillegg til lokal/Oslo for kunder som tilhører gruppen p5, a2 og lk2.	48

1. Introduksjon

I introduksjonen presenteres oppgavens problemstilling, en kort oppsummering av eksisterende litteratur, forklaring rundt oppgavens hensikt og relevante antagelser slik at leseren tidlig får en forståelse av hva oppgaven ønsker å utrette.

1.1 Bakgrunn for oppgaven

Volumutvikling i utlånsporteføljene til banker står sentralt når banker vurderer renteendringer. Hovedfordelen av en renteøkning, alt annet likt, er økte renteinntekter. Den tilsvarende ulempen er tap av utlånsvolum som følge av at lånekunder velger å flytte lånet sitt til konkurrerende banker. Dette dilemmaet står beslutningstakere overfor hver gang en renteøkning utføres. Ønsker man en vekst i utlånsporteføljen eller vekst i netto renteinntekter?

Problemstillingen er utformet i samsvar med Sparebanken Øst der det ble konkludert med at dette ville gi informasjon som er nyttig for banken i tillegg til at det tilfredsstiller den nødvendige metodikken for å avlegge en masteroppgave i Samfunnsøkonomi.

1.2 Problemstillingen

Formålet med oppgaven er å identifisere hvilke kundegrupper som er mest rentesensitive når boliglånsrenten øker, slik at de velger å avslutte boliglånene sine. Oppgaven besitter ikke informasjon om lån avsluttes eller flyttes til andre banker. Derfor blir alle lån som innfris innbefattet av notasjonen «avslutning», selv om det intuitivt kan tenkes at de aller fleste flytter boliglånet til annen bank. Den presenterte sannsynligheten vil da gjelde for en periode på en måned etter en renteøkning. Videre i oppgaven vil «exit» bli ofte brukt som betegnelse for låneavslutning.

Denne oppgaven tar stilling til en enkelt renteøkning som ble gjennomført av Sparebanken Øst ved månedsskifte februar/mars 2017, mot alle utlånskunder i privatmarkedet som hadde boliglånsrente under 3%. Dette var 5680 individuelle boliglån. Dette skjedde på et tidspunkt der rentemarkedet for utlån i stor grad stod stille og perioden bar preg av få hendelser som gav større effekter på utlånsmarkedet som helhet (Haraldsdottir, Wettre, og Tveita, 2018). Dette presiseres fordi kundene til Sparebanken Øst antageligvis hadde tilgang til de samme

tilbudene fra andre banker før og etter renteøkningen. Dersom det eksisterte andre banker som satt opp renten på samme tidspunkt, har ikke gjennomgangen av litteratur funnet noen, noe som i så fall tyder på at dette ikke var en stor bank, og markedet i liten grad ble påvirket.

Mer formelt kan problemstillingen også presenteres som «Determinanter på elasticiteten av exit som følge av renteøkning». Elastisitet defineres som den prosentvise endringen i en variabel når en annen variabel endres med 1%, alle andre forhold likt (Wooldridge, 2014, s. 847). Uttrykket «elastisitet» blir i oppgaven ofte brukt for å fortelle hvordan kundegrupper endrer etterspørselen etter boliglån når prisen endres. Det er vanlig å omtale en kundes etterspørsel som elastisk dersom etterspørselen endres mer enn 1% når prisen endres med 1%. Hvis etterspørselen endres mindre enn 1% ved den samme prisendringen, kalles kundens etterspørsel for «inelastisk».

1.3 Kort oppsummering av eksisterende litteratur

Dagens eksisterende litteratur forteller hva man kan forvente å finne av resultater. Mye av forskningen er utenlandsk som gjør at studien må vurdere eventuelle forskjeller i markedene for utlån mellom de gjeldene landene. På tross av dette er det mye studien kan ta med seg fra utenlandsk forskning når det vurderes hva man forventer å finne samt ved drøfting av resultatene. DeFusco og Paciorek publiserte en artikkel i 2017 som eksempelvis er relevant (DeFusco og Paciorek, 2017). De finner en elastisk etterspørsel etter boliglån for førstegangskjøpere når rentene øker. Lo (2017) støtter opp under funnene og finner også en sterk sammenheng mellom etterspørselen etter boliglån og renteøkning. Hun kontrollerer i tillegg for kredittscoren til kundene og finner at det er de med best kredittscore som har mest elastisk etterspørsel.

Alan og Loranth (2013) finner som Lo (2017) at det er dem med best kredittscore som har mest elastisk etterspørsel når renten endrer seg, men med et datasett bestående av usikrede lån og ikke boliglån. De finner også en sammenheng for lånestørrelse der de med høyest kredittsum endrer etterspørselen mest når renten endrer seg. Karlan og Zinman (2019) fokuserer, i motsetning til de andre som er nevnt her, på kunder med dårlig kredittscore, nemlig kunder i Mexico som etterspør kreditt for å starte firma. Forfatterne presiserer selv at den typiske kunden i forskningen har dårlig kredittscore. De finner en inelastisk etterspørsel hvor kundene i større grad er villige til å ta lånet uavhengig av rentetilbudet.

I Norge har vi særlig Bent Vale med medforfattere som har publisert forskning som omhandler rentesensitivitet for norske boliglånskunder. Dette er nærliggende denne oppgaven og Vales arbeid blir særlig sammenlignet med resultatene fra denne oppgaven. Vale har blant annet utarbeidet en modell for å estimere byttekostnadene for boliglånskunder.

Det som er interessant med denne oppgaven og som gjør den unik fra den eksisterende litteraturen er datasettet som den baserer seg på. Til sammenligning med Vale, Kim og Kliger (2001) som utledet en modell for byttekostnader i det norske bankmarkedet så besitter denne oppgaven data i form av en kontrollgruppe som jeg antar at avslutter lånene sine hver måned. For å forklare nærmere: Det er nærliggende å anta at kunder med store lån oftere bytter boliglånstilbyder etter en renteøkning. Det er derimot også nærliggende å anta at kunder med store lån bytter boliglånstilbyder uten at det skjer en renteøkning, eksempelvis fordi de oppdager at det eksisterer et billigere tilbud i markedet. Så hvor mye av kundeflukten etter en renteøkning skylles renteøkningen, og hva skylles at noen kunder avslutter lånene sine oftere?

Denne oppgaven besitter data som både ser på kunder som avslutter etter en renteøkning og de som avslutter i en normal måned, som også virker å gjøre datasettet unikt. Det blir derfor interessant å se hvilke kunder som har størst «mereffekt» av en renteøkning og sammenligne funn med spesielt resultatene fra Vale, Kliger og Kim (2001).

1.4 Kort beskrivelse av data

Oppgaven baserer seg på et datasett som inneholder alle boliglånskundene til Sparebanken Øst som ble utsatt for en renteøkning 28.02.2017, med noen ekskluderinger som blir forklart i kapittel 4.1. Datasettet er todelt. Det ene er treatmentgruppen som består av 4727 låneengasjement hvor det observeres antall låneavslutninger den første måneden. Det andre datasettet er kontrollgruppen som består av de samme kundene som for treatmentgruppen, men nå er kundene som avsluttet lånene de to første månedene selektert bort, og datasettet består av 4296 låneengasjement. Jeg antar at det eksisterer en renteøkningseffekt for treatmentgruppen, og at denne effekten har utgått for kontrollgruppen i måned 3. Å finne forskjellene mellom kundene som avslutter boliglånet fra de to gruppene er hovedessensen i oppgaven. En oversikt over alle observasjoner i datasettet finnes i tabell 1, kapittel 4.2. Det blir ikke lagt til nye lån som produseres i løpet av de tre månedene.

1.5 Hva ønsker oppgaven å teste?

Studien ønsker å teste rentesensitivitetsforskjeller mellom treatmentgruppen og kontrollgruppen med utgangspunkt i de variablene som oppgaven besitter og som gjennomgås utdypende i delkapittel 4.2. Mer presist kan jeg si at oppgaven ønsker å finne hvilke grupper av bankens kunder som reagerer mest på en renteøkning og derfor har mer elastisk etterspørsel etter boliglån som følge av prisøkning. Jeg besitter kun data ved en renteøkning, og derfor kommer oppgaven kun til å teste hvilke kunder som forsvinner ved en renteøkning, og ikke hvilke kunder som kommer til banken ved en rentenedgang.

1.6 Sentrale antagelser

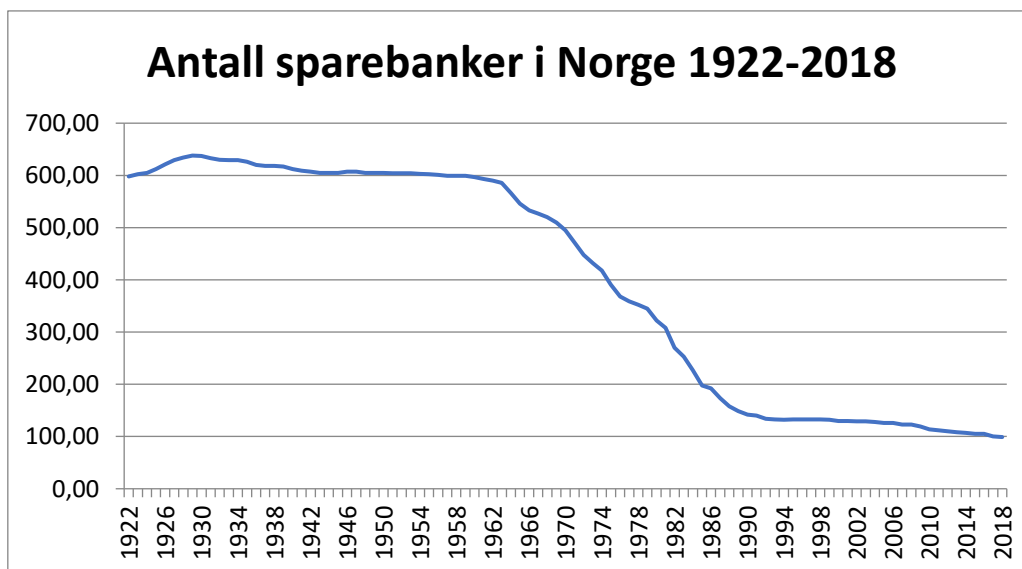
For å kunne estimere mersannsynligheten for exit som følge av renteøkningen er studien nødt til å ha en treatmentgruppe og en kontrollgruppe. Som tidligere nevnt er treatmentgruppen de lånekundene som ble utsatt for en renteøkning og kontrollgruppen de lånekundene som fikk økt rente, men som ikke avsluttet lånene i løpet av de to første månedene. Oppgaven antar at det eksisterer en renteeffekt der kunder i løpet av måned en velger å bytte bank hovedsakelig på grunn av renteøkningen. Dette ser ut til å bli støttet av antallet lån som faktisk blir avsluttet. Oppgaven antar videre at denne renteeffekten utløper i løpet av den andre måneden slik at kontrollgruppen ikke blir påvirket og avslutter lånene sine av andre grunner enn renteøkningen.

Jeg antar i denne oppgaven at effekten av renteøkningen utgår før vi kommer til måned tre. Dette vurderes som en sterk antagelse og avviker nok noe fra virkeligheten. Eksempelvis vil det være nærliggende å anta at mange kunder ikke aktivt følger med på renten sin slik at de ikke får med seg renteøkningen, eller de av individuelle grunner ikke refinansierer umiddelbart, men avventer til senere selv om motivasjonen bak avslutningen av boliglånet er grunnet renteøkningen. Derfor kan jeg tidlig kommentere at kontrollgruppen ikke er perfekt. Likevel observeres en nesten tredobling i antall avslutninger i løpet av den første måneden mot de som avslutter i løpet av den tredje måneden. Det er derfor vurdert som fornuftig å benytte seg av måned tre som kontrollgruppe og jeg forventer å finne virkelighetsnære resultater som kan anvendes for å vurdere rentesensitiviteten for ulike kundegrupper.

Oppgaven antar videre at Sparebanken Øst har en større markedsrett i lokalområdet sitt, som er gamle Buskerud, enn i andre områder på Østlandet. Denne antagelsen er svakere og stemmer nok i større grad overens med virkeligheten. Denne antagelsen er sentral for å kunne drøfte resultatene rundt bostedsområde.

2. Bankmarkedet i Norge

Bankmarkedet preges i stor grad av mange mindre sparebanker og enkelte aktører med mye markedsrett. Gjennom fusjoner kan en tydelig se at banksektoren har blitt mer konsentrert. I 2019 eksisterte det 132 norske banker eller norske datterbanker av utenlandske banker (Bankenes sikringsfond, 2019). Vi kan se at antallet tilbydere av boliglån har blitt kraftig redusert, spesielt mellom 60 og 90-tallet. I 1960 eksisterte det 597 banker i Bankenes sikringsfond sammenlignet med 142 i 1990. Siden har det stabilisert seg noe, selv om vi ser en nedgående trend. Dersom en medlemsbank benytter ulike merkenavn innenfor samme juridiske enhet, regnes alle bankene som en i relasjon til bankgarantien, som er grunnen til at vi har 132 norske banker i 2019 på tross av at den grafiske presentasjonen viser at vi har omtrent 100.



Figur 1: Fusjonering mellom sparebanker tilbake til 1923. Kilde: Sparebankforeningen (2019)

I henhold til markedsandeler er DnB den klart største banken i Norge med 28,9% av den samlede utlånsverdien i 2017.



Figur 2: Markedsandelene i det norske utlånsmarkedet 2017. Kilde: Finansnorge (2020)

I dagens utlånsmarked er det stadig færre barrierer en boliglånskonsument møter i jakten på å bytte bank. Bankene utvikler seg raskt for å tilby kunder selvbetjente og enkle løsninger, og i dag er det mulig å flytte boliglånet sitt når man sitter på bussen eller hjemme foran tv-en. Gjennom signering med BankID kan bankene hente lønnslipper og skattemeldinger gjennom Altinn, og kundene trenger aldri å møte fysisk i banken dersom de ikke ønsker dette. Det finnes flere banker i dag som ikke en gang tilbyr kundemøter. Sbanken (2020) og Nybygger (2020) er eksempler på boliglånstilbydere som har kunder fra hele Norge uten at de har kontorer kundene kan oppsøke. Det at bankene har muligheten til å tilby boliglån uten å være i kontakt med kundene gjør at bankene ikke lenger bare konkurrerer om kundene i sitt lokalmarked, eller med konkurrerende banker i nærområdet. Følgende medfører at konsumentene har et stort utvalg av banker som tilbyr boliglån der en kan refinansiere eksisterende boliglån fra dagens bank til dem. Det finnes også stadig flere løsninger på nett som bidrar til at boliglånskonsumenter enkelt kan få god informasjon om hvilke tilbydere som eksisterer i markedet. På finansportalen.no kan man for eksempel finne over hundre tilbydere av boliglån, prisingen av de ulike tilbyderne samt hvordan man kan komme i kontakt med dem. På tross av at Norge har fått færre og større banker og mindre barrierer som følge av å flytte boliglån mellom banker er det fortsatt naturlig å anta at mange banker fremdeles har et

geografisk satsingsområde. Sparebank 1 Nord-Norge skriver på sine hjemmesider at deres forretningside er å levere finansielle løsninger til kunder med basis i det nordnorske markedet, og deres selverklærte visjon er «for Nord-Norge» (Sparebank 1 Nor-Norge, 2019). Det er derfor naturlig å anta at Sparebank 1 Nord-Norge og Sparebanken Sør ikke konkurrerer om de samme kundene.

2.1 Krav til nye utlån med pant i bolig

Utlånsmarkedet er i dag regulert av forskriften om krav til nye utlån med pant i bolig (boliglånsforskriften). Boliglånsforskriften stiller krav til bankene om hvilke forutsetninger som må være til stede for at banken under normale omstendigheter skal få kunne innvilge lån med pant i bolig. De mest begrensende bestemmelsene stiller krav til at banken maksimalt kan innvilge lån opptil 85% av boligens verdi. Kundenes samlede gjeld skal heller ikke overstige fem ganger brutto årsinntekt. Kundene skal også være i stand til å kunne betjene en renteøkning på 5 prosentpoeng fra det gjeldende rentenivået (Finansdepartementet, 2020). Ved å implementere gjeldende regelverk i utlånsmarkedet begrenses bankenes mulighet til å konkurrere på risiko og bankene ønsker typisk å tilby refinansiering av boliglån til andre bankers boliglånskunder. Dette presiseres fordi dersom du som boliglånskunde kvalifiserer til boliglån i en bank så er sjansen svært stor for at flere banker også ønsker å tilby boliglån. Dette er relevant fordi boliglånskunder i dag har store valgmuligheter dersom de ønsker å bytte bank. Forskriften har gjennomgått noen endringer siden den først ble introdusert i 2015, men begrensningene innenfor egenkapital for primærbolig, maksimal gjeldsgrad i forhold til egen inntekt og likviditet for fremtidige renteøkninger gjaldt i 2017 som beskrevet over (Finansdepartementet 2020).

2.2 Kort om Sparebanken Øst

Sparebanken Øst er en bank med dype røtter i Buskerud og særlig Øvre Eiker. I Buskerud har Sparebanken Øst en større markedsandel enn de har i resten av Østlandsområdet. De er synlige i lokalmiljøet og godt kjent. Banken har i dag 29 kontorer fordelt ut over østlandsområdet der 10 av kontorene befinner seg i det området som tidligere het «Buskerud». Konkurransesituasjonen Sparebanken Øst opplever varierer naturlig nok mellom

lokalmarkedet og «nisje»-markedet. Nisje beskriver her områder der banken har kontorer, men lite markedsrett. Dette er eksempelvis Fredrikstad, Oslo eller Tønsberg. Banken er et resultat av flere fusjoner de senere årene, men har sin oppstart i Skotselv i Øvre Eiker for 188 år siden. Banken fokuserer på være personlig der kundene får bistand av en fast rådgiver.

3. Teori

I dette kapittelet presenteres relevant forskning og litteratur som kan sammenlignes med denne oppgavens innhold. Den eksisterende litteraturen blir videre brukt til å anta hva studien kan forvente å finne av resultater samt drøfte de faktiske resultatene for å vurdere om de samsvarer med relevant forskning.

3.1 Eksisterende litteratur

Det finnes mye forskning fokusert mot elastisitet innenfor utlånsmarkedet, men ettersom dette markedet som regel er sterkt regulert og varierer mellom land blir nøkkelen å hente ut det som er relevant for denne problemstillingen fra den tilgjengelige teorien.

Karlan og Zinman (2019) gjennomførte eksempelvis en 29-måneders studie av mikroutlånsmarkedet i Mexico for å undersøke hvordan låntakere oppfører seg når utlånsrentene endres. Mikrolån defineres her som mindre lån der låntaker typisk er en uattraktiv lånekunde med dårlig kredittscore og uten noe særlig sikkerhet som ønsker å starte eget firma og må gjøre så til relativt høye renter. Forfatterne finner her at låntakerne har en inelastisk etterspørsel etter kreditt og tolererer renteendringer uten at etterspørselen endrer seg i vesentlig grad (Karlan og Zinman, 2019). Forskningen er relevant fordi de også ser på effekten av renteendring, selv om kundene for mikrolån i Mexico antageligvis avviker elastisitetmessig fra boliglånskunder i Norge.

Alan og Loranth (2013) undersøkte rentesensitiviteten for kredittkortlånetakere i Storbritannia. De fant at rentesensitiviteten i stor grad avhenger av hvor gode lånekundene er, målt i kredittrisiko. Kunder som representerte en lav kredittrisiko for banken hadde mer elastisk etterspørsel enn de som representerte høy kredittrisiko for banken. Det samme gjelder for lånebeløp der de fant en positiv sammenheng mellom lånebeløp og rentesensitivitet der

kunder med større lånebeløp etterspurte en relativt mindre kreditt etter renteøkning enn de med lavt lånebeløp. På tross av at usikrede lån i forbindelse med kredittkort og langsiktige boliglån er forskjellige produkter, så er det interessant å se den positive sammenhengen mellom rentesensitivitet og lånestørrelse samt den negative sammenhengen mellom kredittrisiko og rentesensitivitet.

Lo (2017) har forsket på rentesensitiviteten til boliglånstakere i det amerikanske boliglånsmarkedet fra 2008 til 2014 og undersøker forskjeller i rentesensitiviteten for låntakere. Hun finner en sterk sammenheng mellom etterspørselen etter boliglån når rentene endres og kundenes «FICO-score». FICO er den vanligste kredittratingen av personkunder i USA der høyere score tilsvarer bedre rating med høyere gjeldsevne og lavere sannsynlighet for mislighold (FICO, 2020). For kunder med høye FICO-nivåer finner hun eksempelvis en 50% større sjans for valg av bank dersom banken reduserer boliglånsrenten med 25 basispunkter (Lo, 2017). Dette tilsvarer et rentekutt fra 2,5% til 2,25%.

DeFusco og Paciorek (2017) presenterte en artikkel som undersøkte etterspørselen etter boliglån for førstegangskjøpere når rentenivået øker. De fant en 1,5-2% nedgang i etterspørsel etter boliglån når lånerenten økte med et prosentpoeng, noe som støtter opp under funnene til Lo (2017) som også finner negativ sammenheng mellom rentenivå og etterspørsel etter boliglån (DeFusco og Paciorek, 2017). Datasettet som er brukt er rent amerikansk der utlånsmarkedet i større grad er preget av fastrentelån i motsetning til det norske markedet der andelen fastrentelån til norske husholdninger generelt er lave. I september 2019 stod andelen fastrentelån for 8,1% av boliggjelden for norske husholdninger (SSB, 2019), mot rundt 90% i USA (Beckett, 2017). Dette er relevant fordi transaksjonskostnadene ved å bytte bank ved et fastrentelån som regel er høyere enn ved flytende rente. Dersom rentemarkedet har gått ned i fastrenteperioden må du betale banken en «overkurs», som representerer den forventede summen banken taper på at lånet avsluttes. Siden fastrentelån i gjennomsnitt representerer en høyere transaksjonskostnad ved exit enn flytende rente, og høye transaksjonskostnader ses i sammenheng med inelastisk etterspørsel (Kim, Kliger og Vale, 2001), så er dette relevant.

I Norge så har Bent Vale som til daglig jobber for Norges Bank sin avdeling for finansiell stabilitet presentert relevant forskning innenfor det norske bankmarkedet. Han har publisert flere artikler om konkurransen innenfor den norske banksektoren der særlig hans arbeid innenfor byttekostnader er sentralt som teori for denne oppgaven. I samarbeid med Moshe Kim og Doron Kliger (2001), har de presentert en modell for å beregne byttekostnadene innenfor banksektoren. Modellen baserer seg på et datasett av Norske banker fra 1988 til 1996

og gjør da resultatene særlig relevant her. Forfatterne finner at «lock-in»-effekter som løpende genereres ved kundeforhold, og bankens størrelse har stor forklaringskraft for byttekostnadene til en enkelt kunde. Her er bærer bankens størrelse preg av å ha negativ sammenheng med byttekostnader da kundegruppen virker å være relativt mobil (Kim, Kliger og Vale, 2001). «Lock-In»-effekter er i hovedsak en ikke-økonomisk kostnad som kommer av at kunder blir vant til å benytte seg av de produktene og de rutinene de kjenner (Kim, Kliger og Vale, 2001). Det å bytte bank innebærer å lære seg nye åpningstider, kontorlokaler, priser, kontonumre og ofte pinkoder og passord, for å nevne noe. «Lock-in»-effekter bidrar altså til at det er mer krevende for kunder å bytte bank.

3.2 Byttekostnader

Byttekostnader er en sentral del av den eksisterende litteraturen som denne besvarelsen baserer seg på og er både relevant og vesentlig innenfor bankmarkedet (Kim, Kliger og Vale, 2003). Byttekostnader representerer både den faktiske kostnaden ved å bytte bank målt i kroner, ved gebyrer og omkostninger, men gjelder også den subjektivt opplevde kostnaden av å forholde seg til nye rutiner, lære seg nye passord og bestille nye produkter.

Klemperer (1987) presenterte en artikkel som omhandlet konkurransen i markeder med byttekostnader. Her ser han på konsumenter som er delvis låst i konsumet av et produkt der markedet også tilbyr substitutter av det samme godet. Han presenterer flere former for byttekostnader som er svært relevante i denne oppgaven. Boliglånkonsumenter i Norge i dag møter nemlig flere hindringer i jakten på å flytte boliglånet til en ny bank, og størrelsen på kostnadene forbundet med dette vil ha effekt for om en konsument velger å gjennomføre byttet eller ikke. Transaksjonskostnad for bytte av tilbyder er noe som er svært relevant ved bytte av boliglånstilbyder. Denne typen kostnad forbindes med de direkte kostnadene et slikt bytte utløser (Klemperer, 1987). Ved å undersøke DnB sin hjemmeside for priser ved låneopptak ser man at DnB tar et etableringsgebyr fra 1.900 i tillegg til depotomkostninger på kroner 1000,- (DnB, 2019). Legger man til offentlig tinglysningsgebyr for refinansiering på kroner 200,- medfører det en direkte transaksjonskostnad på kroner 3100,- for å flytte et eksisterende boliglån til DnB ifølge DnB selv. Det er nærliggende å tenke at slike transaksjonskostnader vil ha varierende effekt over lånestørrelser. Et lån med kort løpetid og relativt lav lånesum vil ha vesentlig lavere rentekostnader enn et stort lån med lang løpetid der potensialet for å spare rentekostnader er større.

En annen type byttekostnad er ifølge Klemperer kostnaden forbundet med å tilegne seg ny kunnskap (Klemperer, 1987). Det vil for enkelte være krevende å lære seg nye rutiner, finne frem i en ukjent nettbank og forholde seg til nye kontonumre. Det er også en form for byttekostnad dersom en kjenner rådgiveren sin i nåværende bank der en har over tid bygget opp et tillitsforhold.

Byttekostnader er vesentlige og i en annen artikkel Klemperer (1995) presenterer foreslår han at byttekostnader representerer en betydelig reduksjon i velferd da de generelt sett hever priser og videre reduserer inngangsmuligheter for nye bedrifter til å etablere seg i markeder slik at markedskonkurransen reduseres (Klemperer, 1995). Byttekostnader reduserer også ifølge Klemperer produktutvalget kundene har tilgjengelig fordi det hemmer bedrifters insentiver til å diversifisere produktene kundene tilbys.

Moshe Kim, Doron Kliger og Bent Vale har bygget videre på det som Klemperer utarbeidet og presenterer i 2001 en modell for å beregne byttekostnader i det norske bankmarkedet, noe som gjør denne litteraturen relevant for denne oppgaven. De benyttet seg av et datasett basert på alle norske banker mellom 1988 til 1996 og finner at byttekostnadene i det norske bankmarkedet er vesentlig (Kim, Kliger og Vale 2001). Av byttekostnadene er omtrent en fjerdedel som følge av «lock-in»-effekter. Forfatterne beregner byttekostnader til å gjennomsnittlig stå for så mye som 35% av en banks markedsmakt.

Dette forteller oss noe om hvilke variabler som kan ha effekt på sannsynligheten for exit. Det vil være nærliggende å tro at med dagens elektroniske bankverden er alder en faktor som bidrar til byttekostnader. Dette støttes opp under forskning gjort av SSB som viser at unge og utdannede mennesker er flinkere digitalt (Fjørtoft, 2017). Oppgaven har ikke data som representerer utdanning, men har for alder. Det vil være nærliggende videre å anta at byttekostnadene for unge er lavere enn byttekostnadene for eldre fordi terskelen for unge ved bytte av bank virker lavere da de i større grad forstår det som kreves av dem digitalt.

Med utgangspunkt i at «lock-in»-effekter står for en fjerdedel av byttekostnadene vil det også være nærliggende å anta at lengden på kundeforholdet vil være relevant for byttekostnadene der det bør foreligge en positiv sammenheng mellom byttekostnader og lengden på kundeforhold.

4. Databeskrivelse

I dette kapittelet opplyses det først om hvordan data er innhentet og sortert i tillegg til utdypende beskrivelse av hvilke variabler som benyttes, hvordan de er gruppert og deres hensikt i modellen.

4.1 Datainnsamling og sortering

Dataene som er brukt som utgangspunkt for oppgaven er sekundære og kvantitativ. Det betyr at de ikke ble produsert med den enkelte hensikt å besvare problemstillingen (Saunders, Lewis og Thornhill, 2009), men er hentet fra bankens elektroniske kjernesystem. Det er Sparebanken Øst sin IT-avdeling som har bistått med uthenting og sortering av data slik at relevante data blir anvendt og sortert på en oversiktlig måte. Banken sitter på det som med rette kan kalles «utfyllende data» og noe av det første som ble avklart var hvilke variabler som var ønsket i datasettet.

En tidlig avgrensning som ble gjort er ekskluderingen av bedriftsengasjement. Det er nærliggende å anta at prediktorer for exit varierer mellom privat og bedrift og det er i denne oppgavens interesse å se på de private låneengasjementene. Bankens porteføljer i selvbetjente kanaler ble også ekskludert. Det er derfor kontorkanalens portefølje for privatmarkedet med boliglånsrente under 3% som er brukt som datagrunnlag. Banken valgte å ikke innføre renteheving på kunder med rente over 3% og derfor er disse også fjernet fra datasettet. Videre er personinformasjon som vil gjøre det mulig å identifisere kunder fjernet fra datasettet før det er oversendt meg. Dette er da navn, personnumre og adresser.

Etter den initiale produksjonen av data fra Sparebanken Øst sin IT-avdeling ble det oversendt 5 680 boliglån. En tidlig utfordring som måtte tas hensyn til er hvordan flere lån per kunde påvirker resultatene. I henhold til bankens rutiner så deles boliglån som overstiger 75% av boligens verdi der hovedlånet settes innenfor 75% og resten som et topplån. Det vil altså medføre at en låntaker med seks millioner i låneengasjement med topplån vil bidra med et stort og et lite lån, noe som vil påvirke resultatene uheldig. Heldigvis er datasettet satt opp slik at lån fra samme låneengasjement kommer sortert etter hverandre. Derfor ble alle lån med samme oppstartdato, har samme alder, samme kommune som i tillegg ligger etter hverandre i datasettet vurdert som så sannsynlig at var samme låntaker at en med særdeles liten risiko for

å gjøre feil kan slå disse sammen. Feilkilden er at lån som ikke eies av samme låntaker slås sammen og gir større lån enn hva som er realiteten. Denne feilkilden er derimot vurdert som så liten at den ikke veier opp for feilen ved at alle låneengasjement med topplån teller dobbelt, og derfor er sammenslåingen gjennomført.

En annen potensiell feilkilde ligger i de minste lånene der en kan anta at byttekostnadene vil være så store at lån ikke vil avsluttes og at dette kan påvirke resultatene. Mindre lån vil også antageligvis ha større mulighet for å bli innfridd av kunder uten at dette er grunnet bytte av bank. Det ble derfor bestemt at alle lån med størrelse mindre enn 300.000 norske kroner ble ekskludert fra datasettet. Dette er derimot gjort etter at topplån er slått sammen med hovedlån, slik at topplån med lånestørrelse mindre enn 300.000 er hensyntatt. Etter denne initiale databehandlingen består datagrunnlaget av 4727 låneengasjement.

Datoen for datafilen er 28.02.2017 og datafilen viser alle lån over en periode på tre måneder. Nye lån er derfor ikke hensyntatt, og det er kun lånene som eksisterte 28.02.2017 som er brukt. I datasettet så følger studien låneporteføljen over tre måneder og ser hvilke lån som avsluttes i løpet av denne perioden. Nye lån som utbetales av banken i denne perioden tas ikke med i datasettet.

Det er antatt at kundene som påvirkes av renteeffekten avslutter lånet i løpet av den første måneden. Når oppgaven sammenligner kundene som avsluttet lånene tas ikke kundene fra måned to hensyn til. Dette er fordi jeg antar at renteeffekten utgår i løpet av denne måneden og det vil være vanskelig å vite hvem som avslutter grunnet renteøkning og hvem som avslutter av andre grunner. Kontrollgruppen er ment å representere bankens boliglånsportefølje når det ikke eksisterer en renteøkningseffekt, og har låneavslutninger selv om det ikke er grunnet renteøkningen. Oppgaven sammenligner derfor kundene som avslutter mellom de to gruppene for å finne hvilke variabler som medfører økt rentesensitivitet. Når oppgaven videre ser på mersannsynligheten for exit så sammenlignes disse to underpopulasjonene for å se hvilke kunder som har en mersannsynlighet for exit etter en slik renteheving. Antagelsen om at renteeffekten har gått bort i løpet av to måneder vil nok ikke være 100% korrekt, men den gir et sammenligningsgrunnlag som vurderes som tilfredsstillende for denne utredningen. Dette er diskutert i kapittel 1.6.

Studien ender til slutt med et datasett der 308 lån avsluttes i løpet av den første måneden fra treatmentgruppen mot 112 som avslutter i løpet av den tredje måneden fra kontrollgruppen. I

løpet av den andre måneden avsluttes 123 lån. 4184 låneengasjement blir værende igjen i porteføljen og avsluttes ikke.

4.2 Modellens variabler

Datasettet som oppgaven baserer seg på besitter følgende informasjon: Hvilken måned et lån er avsluttet i, dersom det avsluttes. Alderen til hovedlåntaker, lånestørrelse, lengden på kundeforhold og i hvilket fylke kundene bodde ved renteøkningstidspunktet. Det er variabler som gjerne skulle vært med, men som av personvern hensyn ikke er hentet ut. Dette er eksempelvis data om kjønn eller utdanning som antageligvis ville resulterte i interessante resultater. På tross av at banken sitter med denne informasjonen, så ville denne informasjonen måtte bli hentet ut ved å manuelt søke opp kundeforhold og dette ville brutt med bankens retningslinjer i henhold til personvern.

4.2.1 Avslutning

Modellens avhengige variabel er en dummyvariabel med verdi null eller en og navn «avslutning» som representerer om en kunde avslutter lånet sitt eller ikke. Modellen predikerer sannsynligheten for suksess ($\text{avslutning}=1$), slik at verdien til den avhengige variabelen vil være det samme som sannsynligheten for at en kunde avslutter lånet påfølgende måned etter renteøkning. I måned en avsluttes 308 lån fra treatmentgruppen. I måned to avsluttes 123 lån der jeg antar at renteeffekten utgår i løpet av denne måneden. I måned tre avsluttes 112 lån der jeg antar at renteeffekten har utgått og at kundene avslutter lånene av andre årsaker enn renteøkningen. Kundene som er igjen ved måned tre blir oppgavens kontrollgruppe og sammenlignes med treatmentgruppen for å finne relevante forskjeller. Kontrollgruppen er selektert og inneholder ikke lånene som ble avsluttet i løpet av måned en eller måned to. Den avhengige variabelen observeres altså 420 ganger i datasettet, $\text{mnd1} + \text{mnd3}$.

4.2.2 Variabel for måned

Studien har tre månedsdummyer som representerer om et lån avsluttes i måned en, to eller tre, med henholdsvis 308, 123 og 112 observasjoner. Månedsdummy for måned en kalt «mnd1» representerer de kundene som avsluttet lånet i løpet av den første måneden etter renteøkningen. Denne underpopulasjonen er sentral for oppgaven for det er blant disse kundene mersannsynligheten for exit grunnet renteøkning vil foreligge. Månedsdummy for måned to kalt «mnd2» representerer kundene som avslutter lånet i løpet av den andre

måneden etter renteøkning. Ettersom oppgaven antar at effekten av renteøkningen er borte når vi kommer til måned tre, men at den er til stede for måned en, så representerer måned to en periode der renteeffekten avtas fra å gjelde for måned en og ikke være relevant i måned tre. Ettersom det er problematisk å skille kundene i måned to mellom hvem som avslutter som følge av renteøkningen så blir ikke denne månedsdummyen en uavhengig variabel som brukes til å predikere avslutning. Månedsdummy for måned tre er kalt «mnd3» og representerer de kundene som avslutter lånet i løpet av den tredje måneden etter renteheving. Jeg antar at renteeffekten har utløpt. Selv om det ikke lenger er en renteeffekt så vil det til enhver tid være avslutninger av boliglån til fordel for andre banker, og denne gruppen med mnd3 = 1 representerer låneavslutningene for kontrollgruppen. Når oppgaven sammenligner mersannsynligheten for exit så identifiseres derfor forskjeller mellom de to ulike underpopulasjonene i mnd1 og mnd3.

I tabellen under oppsummeres alle forklaringsvariablene som datasettet består av.

Variabel	Totalt antall observasjoner	Variabelnavn i modellen	Låneavslutninger i treatmentgruppen	Låneavslutninger i kontrollgruppen
Totalt	4727		308	112
Måned 1	308	<i>mnd1</i>	308	0
Måned 3	112	<i>mnd3</i>	0	112
Persentil 10	524	<i>p1</i>	27	7
Persentil 20	525	<i>p2</i>	31	19
Persentil 30	525	<i>p3</i>	30	8
Persentil 40	525	<i>p4</i>	26	7
Persentil 50	525	<i>p5</i>	34	14
Persentil 60	525	<i>p6</i>	35	13
Persentil 70	525	<i>p7</i>	41	18
Persentil 80	525	<i>p8</i>	40	14
Persentil 90	528	<i>p9</i>	44	12
Aldersgruppe1	1146	<i>a1</i>	92	29
Aldersgruppe2	2887	<i>a2</i>	182	74
Aldersgruppe3	694	<i>a3</i>	34	9
Lengde kundeforhold1	1123	<i>lk1</i>	83	31
Lengde kundeforhold2	1111	<i>lk2</i>	44	30
Lengde kundeforhold3	2039	<i>lk3</i>	153	44
Oslo	651	<i>oslo</i>	39	16
Lokal	2520	<i>lokal</i>	196	58

Tabell 1: Summary statistics

Note: Tabellen viser antallet observasjoner for hver forklaringsvariabel. En lineær sannsynlighetsmodell som oppgaven benytter, presenterer sannsynlighetskoeffesienter som direkte blir påvirket av antallet observasjoner innenfor hver variabel for treatmentgruppen og kontrollgruppen sammenlignet med totalt antall observasjoner. Variablene er sortert slik at

det er likt antall observasjoner innenfor hver persentilgruppe for lånestørrelse. For alder og lengde på kundeforhold er ikke dette gjort fordi det er vurdert som mer hensiktsmessig å sortere etter likhet blant kundene og ikke for å ha et likt antall kunder i hver gruppe.

4.2.3 Variabel for lånestørrelse

Variabelen som representerer størrelsen på kundens lån kalles i datasettet for *spersentil*. Dette er en kategorisk og uavhengig variabel som er delt opp i ni persentiler fra 10 til 90 for å sammenslå lånestørrelsene og lettere observere koeffisienten som gir mersannsynlighet for exit for denne lånepersentilen. Her er representert verdien 10 at en kunde er i gruppen med minst lånestørrelse. Ettersom studien har ni ulike verdier for lånestørrelse så er derfor ikke en kunde med *spersentil* = 10 blant de 10% med lavest lånestørrelse, men heller avrundet 11,1%. I regresjonen noterer px hver persentil der x varierer mellom 1 og 9. Hver persentil blir derfor en dummyvariabel med navn px . Alle kundene tilhører en av persentilgruppene.

4.2.4 Variabel for alder

Denne variabelen representerer intuitivt nok hvilken alder hovedlåntaker har datert 28.02.2017. Oppgaven sitter ikke med informasjon om det eksisterer medlåntakere på lånene, så det er hovedlåntakers alder som er førende. Det er også slik at i løpet av de tre månedene hvor lånene observeres for avslutning så endres ikke alder. Det betyr at alderen som brukes har økt for omtrent en fjerdedel av lånene etter den tredje måneden. Det antas at dette er tilnærmet irrelevant for resultatene i oppgaven og blir ikke vektlagt ytterligere.

Siden jeg ønsker å ha alle variabler på dummyform så er alder fordelt over tre variabler med navn a_1 , a_2 , og a_3 . Variablene benyttes, fra lavest til høyest, når hovedlåntaker har alder 21-35, 36-60 eller 61-92.

4.2.5 Variabel for kundeforholdets lengde

Variabelen er som med alle variablene som brukes, også på dummyform. Her deles kundene inn i de ulike underpopulasjonene basert på hvor lenge hovedlåntaker har vært kunde i banken. Det er derfor ikke låneoppstart som er førende, men når kundeforhold ble etablert.

Denne variabelen er håndtert likt som aldersvariabelen og fordelt utover tre dummyvariabler fra kortest til lengst lengde på kundeforholdet, med navn lk_1 , lk_2 , og lk_3 .

lk_1 brukes når kundeforholdet har eksistert kortere enn 01.01.2014. Ved kundeforhold som ble inngått mellom 01.01.2007 og 01.01.2014 benyttes lk_2 , og lk_3 brukes for alle lån der kundeforhold er eldre enn 01.01.2007. Siden datafilen er fra 28.02.2017 så stemmer det ikke

nøyaktig når jeg omtaler kunder i lk_2 som at de har hatt kundeforhold i mellom 3-10 år, men jeg omtaler dem likevel slik senere i oppgaven fordi det er enklere.

4.2.6 Variabel for bostedsområde

I datasettet eksisterer det to dummyvariabler for å kontrollere for bosted. Studien har en dummy som heter *lokal* og gjelder dersom hovedlåntaker har adresse i det som tidligere var Buskerud. Grunnen til at dette vurderes som relevant er at dette er lokalmarkedet til Sparebanken Øst. Det er her banken har markedsrett og er kjent blant befolkningen, og det vurderes som relevant for rentesensitiviteten. Dummy for lokal brukes for omtrent 53% av låneengasjementene.

Den andre månedsdummyen heter *oslo* og benyttes dersom hovedlåntaker har adresse i Oslo. Det er besluttet å ta med en slik dummy fordi individer som bor i Oslo har høyest utdanningsnivå som kan ha en effekt på resultatet (Holøien, Nygård og Holseter, 2017). I datasettet har omtrent 14% av låneengasjementene adresse i Oslo. Kunder som derimot ikke tilhører lokal eller Oslo blir derfor den tredje kundegruppen, og de inngår i konstanttermen i alle regresjonene oppgaven senere benytter. Dette er eksempelvis kunder fra Fredrikstad, Asker, Tønsberg eller Jessheim.

5. Metode

I metodekapittelet presenteres metoden som oppgaven bruker for å gjennomføre utredningen og videre besvare problemstillingen.

5.1 Forskningsdesign

Oppgavens formål beskriver hva en ønsker å oppnå. Normalt deles formålet inn i tre hovedtyper, der en ønsker å enten utforske, forklare eller beskrive de problemstillingene som oppgaven tar stilling til (Grønmo, 2016, s375). Studien må ikke utelukkende tilhøre en av de tre typene, men det er relevant å definere hvilken kategori den hører til ettersom hvilke teknikker og forskningsmetoder som anvendes normalt varierer mellom de tre typene. Utforskende studier beskriver konsekvenser eller betydninger studiens resultat har for omgivelsene. Forklarende studier ønsker å finne virkninger og årsaker som forklarer

sammenhenger. Beskrivende studier søker å beskrive de ulike prosessene eller fenomenene som observeres (Grønmo, 2016).

Oppgaven har som mål å sammenligne treatmentgruppen og kontrollgruppen for å se forskjeller i rentesensitivitet for ulike variabler, dem imellom. Oppgaven kategoriseres derfor som en «forklarende studie» (Grønmo, 2016).

5.2 Økonometrisk metode

Dette kapitlet beskriver den økonometriske metoden som oppgaven benytter seg av for å nå resultatene. Økonometriske analyser baseres på statistiske metoder for å estimere økonomiske forhold og videre teste økonomiske teorier (Wooldridge, 2014, s. 01).

5.2.1 Lineær sannsynlighetsmodell

Oppgaven benytter seg av en regresjon der den avhengige variabelen er binær, nemlig at den kun har to mulige verdier. I oppgavens modell er disse verdiene 0 eller 1, hvor 0 noterer kunder som ikke avslutter lån og 1 noterer avsluttede lån. En slik modell kalles en lineær sannsynlighets modell (Wooldridge, 2014, s. 248).

Hvis vi antar at den forventede verdien til restleddet $u = 0$, vil den forventede verdien til y være:

$$E\{y|X\} = \alpha + \beta_1 X_1 + \dots + \beta_k X_k, \quad (5.2.1)$$

hvor X representerer alle de uavhengige variablene. Når y er binær ved å enten ha verdien 0 eller 1, så er $P\{y = 1|X\} = E\{y|X\}$, slik at den forventede verdien til y er den samme som sannsynligheten for at $y = 1$. Vi får derfor ligningen:

$$P\{y = 1|X\} = \alpha + \beta_1 X_1 + \dots + \beta_k X_k, \quad (5.2.2)$$

Ligning (5.2.2) sier at sannsynligheten for at $y = 1$ for $p(x_j) = P\{y = 1|x_j\}$, er en lineær funksjon av x_j . Denne typen modell kalles som tidligere nevnt en «lineær sannsynlighetsmodell», fordi den endrede sannsynligheten for $P\{y = 1|X\}$, er lineær for parameterne β_j (Wooldridge, 2014, s. 249). I denne oppgavens modell som med den lineære sannsynlighetsmodellen som vises i (5.2.2) så representerer koeffisienten β_j den endrede sannsynligheten for $y = 1$ når X_j endres.

Det som gjør oppgavens modell spesiell er at den utelukkende benytter dummyvariabler som uavhengige variabler. En dummyvariabel er en variabel som kun har verdi 0 eller 1. Et eksempel på en dummyvariabel i vår oppgave er *lokal* som definerer om en kunde tilhører lokalbanken eller ikke. Variabelen *lokal* vil ha en koeffisient β_j hvor verdien representerer mersannsynligheten for at en kunde avslutter lånet sitt dersom de tilhører lokalbanken i forhold til konstanttermen. Dersom kunden ikke tilhører lokalbanken vil $\Delta y = \beta_j(0) = 0$ dersom *lokal* er modellens eneste uavhengige variabel. Ved å videre legge til alle variablene på dummyform kan vi holde alle variablene lik null, bortsett fra en. Den endrede sannsynligheten for avslutning vil være tilsvarende koeffisienten β_j til variabelen og representerer mersannsynligheten for exit sammenlignet med konstanttermen. Oppgaven bestemmer selv hvilke kunder som benyttes i konstanttermen.

Eksempelvis kan vi tenke oss at vi i konstanttermen har en kundegruppe med lånestørrelse $s = m$. Ved å legge til en dummyvariabel der vi har $s = h, h > m$, så kan vi observere den nye sannsynligheten for exit ved en høyere lånestørrelse, s , som vil være lik koeffisienten til dummyvariabelen til den tillagte lånestørrelsen. La oss vise formelt ved å anta at vi kun har en forklaringsvariabel i modellen som representerer lånestørrelse h , notert S_h . La oss også anta at konstanttermen definerer sannsynligheten for exit ved lånestørrelse S_m , hvor $S_h > S_m$

Vi kan da sammenligne følgende ligninger for å finne mersannsynligheten for exit for kundegruppe S_h i forhold til S_m . Husk at vi har at $S_m = \alpha$

$$\text{Regresjonsligningen for eksempelmodellen: } P\{y = 1|X\} = \alpha + \beta_1 S_h + u \quad (5.2.3)$$

Her er: $S_m = \alpha, E(u) = 0$, og S_h kan enten ha verdi 0 eller 1

Sannsynligheten for exit for kundegruppe S_m presenteres slik:

$$P\{y = 1|S_m\} = \alpha + \beta_1(S_h * 0) \quad (5.2.4)$$

Sannsynligheten for kundegruppe S_h presenteres slik:

$$P\{y = 1|S_h\} = \alpha + \beta_1(S_h * 1) \quad (5.2.5)$$

Ved å observere den ulike sannsynligheten mellom $P\{y = 1|S_m\}$ og $P\{y = 1|S_h\}$ vil vi observere den endrede sannsynligheten for de to kundegruppene med ulik lånestørrelse. Ved å følge denne metoden utvikler oppgaven regresjonsligningen ved å legge til variabler. For å finne den endrede sannsynligheten holdes enkelt og greit alle variablene lik 0 bortsett fra en variabel som får verdi 1. Dette kan gjøres for alle variablene og vil dermed resultere at vi

finner den endrede sannsynligheten for alle variabler sammenlignet med konstanttermen når alle andre variabler holdes konstante. Den endrede sannsynligheten for exit når $X_j = 1$, $\Delta \frac{P\{Y=1\}}{X_j}$ vil derfor være tilsvarende koeffisienten β_j .

Det kan eksistere variabler som har en påvirkning på y , men som ikke er inkludert spesifikt i modellen. Disse befinner seg i restleddet og har forventet verdi lik null. Ved å tilføye flere variabler som har effekt på y så gis modellen større forklaringskraft, og usikkerheten i modellen reduseres (Wooldridge, 2014, s. 69). I modell 1 som vises formelt i ligning (7.1.1) benyttes en regresjonsligning som er utformet likt som ligning (5.2.2).

5.2.2 Interaksjonsvariabler

Ved å benytte en ligning tilsvarende (5.2.2) vil man finne den endrede sannsynligheten for exit for de ulike variablene sammenlignet med konstanttermen. Denne oppgavens problemstilling ser derimot på forskjellen på den endrede sannsynligheten mellom treatmentgruppen og kontrollgruppen, noe (5.2.2) ikke forklarer. Dette løser oppgaven ved å inkludere noe som kalles interaksjonsvariabler i regresjonsligningen. En interaksjonsvariabel ser på hvordan en forklaringsvariabel endres når en annen forklaringsvariabel endres (Wooldridge, 2014, s. 198). Vi kan eksempelvis se på effekten av å variere lånestørrelse fra S_m til S_h slik som over, mellom treatmentgruppen og kontrollgruppen. La oss fortsette fra ligning (5.2.5) der vi kun har S_h som forklaringsvariabel, men vi utvider ved å ha forklaringsvariabler for måned 1 i tillegg. Dette får nemlig observasjoner fra måned 3 inn i konstanttermen α , slik at konstanttermen nå representerer sannsynligheten for lånestørrelse S_m for kundene i måned 3.

$$P\{y = 1|S_h\} = \alpha + \beta_1 S_h + \beta_2 mnd1 + \beta_3 (S_h * mnd1) \quad (5.2.6)$$

Sannsynligheten for avslutning blir når $S_h = 1$

$$\frac{\Delta P\{Y = 1\}}{\Delta S_h} = \beta_1 + \beta_3$$

Koeffisient β_1 vil vise sannsynligheten for kontrollgruppen med lånestørrelse S_h sammenlignet med konstanttermen. β_3 viser mersannsynligheten for lånestørrelse S_h for treatmentgruppen sammenlignet med kontrollgruppen. Dersom $\beta_3 < 0$ har kontrollgruppen en høyere sannsynlighet for exit for lånestørrelse S_h enn treatmentgruppen.

Med interaksjonsvariabler blir regresjonsligningen presentert slik:

$$P\{y = 1|X\} = \alpha + \beta_1 X_1 + \dots + \beta_k X_k + \beta_{k+1}(X_1 * mnd1) + \dots + \beta_{k+k}(X_k * mnd1)$$

(5.2.7)

Ligning (5.2.7) viser hvordan regresjonsligningen som denne oppgavens modell benytter, fungerer. En kan se fra modellene oppgaven benytter for lånestørrelse i ligning (7.2.1) og kontrollert for sosialdemografiske variabler i (7.3.1), at ligningene er utformet likt som ligning (5.2.7).

En slik modell har fordeler og ulemper. Fordelene er at modellen vil estimere koeffisienter som er enkle å tolke samt koeffisienter som tydelig viser oss mersannsynligheten mellom treatment og kontrollgruppen. Ulempen med en slik modell er at modellen kan estimere sannsynligheter over 1 og under 0, noe som i realiteten er umulig. En lineær sannsynlighetsmodell har også heteroskedastisitet i restleddet som må tas høyde for ved vurdering av standardavvik (Wooldridge. 2014, s. 251). Dette gjøres ved å bruke robuste standardavvik i regresjonen vår. Alternativer til å bruke en lineær sannsynlighetsmodell er å bruke en probitmodell med kumulativ standard normalfordeling eller en logitmodell med kumulativ standard logistisk fordeling. Etersom vi har relativt få observasjoner for noen av variablene og grunnet den enklere tolkningen av koeffisienter så er det vurdert som mest hensiktsmessig å benytte en lineær sannsynlighetsmodell.

5.3 Modellens hensiktsmessighet

Modellens hensikt er å finne hvilke variabler som gir en økt mersannsynlighet for exit som følge av en renteøkning. Modellen baserer seg på et datasett som virker, sammenlignet med den eksisterende litteraturen, å være unikt ved at studien både har en treatment og en kontrollgruppe som lar meg se på elastisiteten som følge av renteøkningen. Den mest nærliggende forskningen som er gjennomgått er Vale, Kim og Kliger (2001) sin estimering av byttekostnader i banksektoren. De finner estimater for hvilke kunder som har lettest for å bytte bank, men sier ikke noe om effekten en renteøkning har for kunders sannsynlighet for å avslutte boliglån. Modellen kan derfor brukes av banker for å vurdere effekten en isolert renteøkning har på porteføljen deres og kan muligens bli brukt til å prise kunder på en slik måte at bankene sparer frafall i utlånsporteføljene sine og boliglånskonsumenter slipper å bytte bank.

Det ville videre vært interessant å observere hvilke kunder som flytter til en bank som velger å prise seg lavere enn konkurrentene sine og se om resultatene sammenfaller med resultatene fra denne oppgaven.

6. Økonometrisk analyse

I dette kapittelet forklares hvordan Stata er benyttet som verktøy for det økonometriske arbeidet før intuisjonen bak regresjonsmetoden presenteres.

6.1 Fremgangsmåte i Stata

Oppgaven har brukt Stata som verktøy for den økonometriske analysen. Stata er et utberedt og anerkjent statistikkprogram innenfor økonometri og er vurdert som mest hensiktsmessig i denne oppgaven. Selve analysen er gjennomført i Stata, men datasorteringen er gjort både i Stata og i Excel siden enkelte prosesser og sorteringsutfordringer lettere lar seg utføre i Excel.

Det er som tidligere nevnt vurdert som mest hensiktsmessig å ha alle variablene og underpopulasjonene på dummyform med verdi null eller en. Dette er gjort for å lettere vurdere den partielle effekten ved å gå fra 0 til 1, eller motsatt. Det betyr at for de kategoriske variablene *spersentil*, *aldergruppe* og *lengdekundeforhold* er det laget en dummy for hver underpopulasjon som det gis et oversiktlig bilde av i tabell 1. Det vil si at for den kategoriske variabelen *spersentil* som har ni ulike verdier så produseres det istedenfor ni ulike dummyvariabler, en for hver enkelt persentil. Variabelen er derfor endret slik. Initialt: $spersentil_{i=1,2,\dots,9} = 10(i)$. Og etter endring: $spersentil_{i=1,2,\dots,9} = p1, p2, \dots, p9$ der px representerer en enkelt dummyvariabel for et enkelt persentilnivå. Tilsvarende er gjort for variablene *aldersgruppe* og *lenggekundeforhold*.

For å mest presist tolke sannsynlighetskoeffisientene som modellen produserer så er det produsert to datafiler som er slått sammen i Stata ved hjelp av *append*, (Stata, 2020). Den ene datafilen inneholder alle bankens lån utsatt for renteøkning (treatment) der kun de som avsluttes i løpet av måned en registreres med en avslutning. I den andre datafilen er det ekskludert ut alle lån som ble avsluttet i måned en og måned to og blir da kontrollgruppen. I denne datafilen er det kundene som avslutter lån i måned tre som registreres med avslutning.

Datafil en inneholder derfor 431 observasjoner mer enn datafil to. Kommandoen *append* legger til datafil to med datafil en. Oppgaven vil derfor i den lineære sannsynlighetsmodellen evne å se både sannsynligheten for måned en og måned tre ved bruk av regresjonsligninger som (5.2.7) som presentert i kapittel 5.2.

Regresjonen som benyttes har en dummyvariabel som avhengig variabel. Dette innebærer at dersom den avhengige variabelen blir lik 1 i modellen så representerer det en 100% sannsynlighet for avslutning, og tilsvarende ved verdi lik 0, som representerer en 0% sannsynlighet for avslutning. Det er derfor svært nærliggende å anta at den avhengige variabelen vil ha en verdi mellom 0 og 1.

For å benytte seg av en slik modell i Stata, så kreves det å ha en tydelig definert konstanterterm. Dette er fordi studien trenger en definert gruppe fra alle variablene som kan sammenlignes med den variabelen som endres. Eksempelvis kan jeg ikke velge å ha en konstanterterm uten spesifisert aldersgruppe, fordi Stata ikke vet hvilken aldersgruppe den skal sammenligne med når variabelen aldersgruppe endres.

6.2 Regresjonsanalyse

For å få et bilde av hvilke variabler som virker å påvirke sannsynligheten for exit er det først laget en modell som ser på sannsynligheten for exit for alle de avsluttede lånene over en tremånedersperiode. Videre er modellen utviklet ved at det er lagt på interaksjonsvariabler for å finne den partielle effekten ved den avhengige variabelen med hensyn på de forklarende variablene (Wooldridge, 2014, s. 198). Dette er steget som i hovedsak er det essensielle i denne oppgaven ettersom det er her vi ser mersannsynligheten ved hver variabel når oppgaven ser på forskjeller mellom treatmentgruppen og kontrollgruppen. I førsteomgang presenteres en modell som utelukkende bruker lånestørrelser som uavhengig variabler, fordelt utover de ni persentilgruppene hvor en av dem plasseres i konstantertermen. Denne modellen utvikles videre der det kontrolleres for alder, bostedsområde og lengde på kundeforhold. Videre i denne oppgaven vil disse tre variablene bli kalt «sosialdemografiske variabler» når de omtales samlet.

7. Modellen

I dette kapittelet presenteres de tre modellene som benyttes for å presentere resultatene oppgaven ønsker å finne. Modellene som benyttes fokuserer på interaksjonsvariabler for å finne mersannsynligheten for exit i treatmentgruppen sammenlignet med kontrollgruppen og hvordan denne oppfører seg for de ulike variablene.

Den første modellen brukes for å gi en oversikt over hvem det er som har avsluttet boliglånet sitt i løpet av de tre månedene etter renteøkningen. Modellen tar ikke hensyn til effekten for renteøkningen. Den blander sannsynligheten for exit for de som avslutter på grunn av den økte renten med de som avslutter boliglånet sitt av andre grunner. Derfor er denne modellen kun tatt med for å gi et bilde av hvem som har avsluttet lånene i løpet av de tre månedene og vil ikke påvirke oppgavens konklusjon.

Videre utvikles modellen ved bruk av interaksjonsvariabler der jeg tar hensyn til renteeffekten og skiller mellom treatmentgruppen og kontrollgruppen. Jeg antar at renteeffekten utløper i løpet av måned to slik at gruppen som avslutter i måned en er påvirket, men at gruppen som avslutter i måned tre ikke er det. I modell to sammenlignes de to gruppene og ser på den observerte sannsynligheten for exit dem imellom. Denne modellen ser utelukkende på mersannsynligheten med lånestørrelser som uavhengige variabler. Videre utvikles modellen for å kontrollere for sosialdemografiske forhold. I vår oppgave er dette alder, bostedsadresse i form av fylke før fylkessammenslåing og lengden på kundeforholdet. Jeg får da en lineær sannsynlighetsmodell som estimerer sannsynlighetskoeffisienter for de ulike variablene hvor alt annet holdes likt sammenlignet med konstanttermen.

7.1 Modell 1

Selv om oppgavens problemstilling er å identifisere hva som kjennetegner kunder med mersannsynlighet for exit etter renteheving, så er det interessant å vurdere hvilke variabler som generelt gir økt sannsynlighet for exit i datasettet vårt. Her hensyntas alle kundene som avsluttet boliglånene sine i løpet av en tremånedersperiode etter rentehevingen.

Regresjonsresultatene fra modellen kan sees i appendikset som *tabell 1*.

Som tidligere nevnt så er alle forklaringsvariabler dummyer. Derfor kan en teste effekten en variabel har på sannsynligheten for exit ved å endre fra å kun se på konstanttermen ved å

legge til en variabel og observere hvordan sannsynligheten endrer seg. Konstanttermen er i denne modellen sannsynligheten for exit ved måned 1 + måned 2 + måned 3 for kundene med lånestørrelse i *persentil 50*, alder i aldergruppe 2, lengden på kundeforholdet er i gruppe 2, og kunden tilhører verken Oslo eller lokalområdet. La oss anta at sannsynligheten for exit ved denne kundegruppen er 5%. Modellens konstantterm blir da 0.05. For aldersgruppe 1 så holdes de andre variablene lik 0, men variabel for aldersgruppe 1 får verdi 1. La oss anta at koeffisienten som modellen produserer blir -0.01. Dette betyr at for denne gruppen er sannsynligheten for exit $0.05 - 0.01 = 0.04 \rightarrow 4\%$

Slike koeffisienter blir produsert for alle dummyvariablene og viser altså hvordan sannsynligheten for exit endres når en variabel får verdi 1, men holder alle andre variabler konstante. Modellen kan sees at fungerer likt som ligning (5.2.2) som er presentert i metodekapittelet.

Modellen utledes ved følgende regresjon:

$$P\{y = 1|X\} = \alpha + \beta_1 p_1 + \beta_2 p_2 + \beta_3 p_3 + \beta_4 p_4 + \beta_6 p_6 + \beta_7 p_7 + \beta_8 p_8 + \beta_9 p_9 + \beta_{10} lokal + \beta_{11} oslo + \beta_{12} a_1 + \beta_{13} a_3 + \beta_{14} lk_1 + \beta_{15} lk_3 + u \quad (7.1.1)$$

For tolkning av variabler vises det til kapittel 4.2.

Ettersom p_5 , a_2 , og lk_2 befinner seg i konstantleddet er de ikke med i selve regresjonsligningen.

Med utgangspunkt i den gjennomgåtte litteraturen så har man noen forventinger til resultatene. Alan og Loranth (2013) og Lo (2017) presenterer funn som tilsier at lånestørrelse står i positiv sammenheng med elastisk etterspørsel. Det er derfor nærliggende å tenke at en finner en positiv sammenheng mellom relativt store lån og sannsynlighet for exit. Med utgangspunkt i teorien rundt byttekostnader som er gjort rede for av Klemperer (1995) og videre utviklet mot det norske markedet og banksektoren av Kim, Kliger og Vale (2001) tyder litteraturen på at alder, lengden på kundeforholdet og lånestørrelse vil være relevante. Lengden på kundeforhold forventes å være i negativ sammenheng med sannsynlighet for exit grunnet «lock-in»-effekter. Desto lenger en har vært i banken, desto tyngre oppleves det å tilegne seg ny kunnskap, anskaffe nye produkter og generelt forholde seg til ny bank. Alder forventes også å være i negativ sammenheng med sannsynlighet for exit. Det å bytte bank i dag gjøres lettere av digital kompetanse noe Fjørtoft i SSB (2017) viser at unge mennesker innehar i forhold til eldre mennesker og vil påvirke byttekostnadene som de eldre opplever.

Litteraturen tyder også på at byttekostnadene reduseres med lånebeløp ettersom lånekunder med større boliglån sparer mer ved å bytte bank enn lånekunder med mindre boliglån gitt samme rente før og etter exit og at gebyrer typisk ikke er avhengig av lånestørrelse.

Det er derimot mer usikkert hva en kan forvente av dummyene *lokal* og *oslo*. I lokalbanken så antar studien at det eksisterer en markedsrett som er større enn utenfor lokalmarkedet.

Eksempelvis kan en tenke seg at i lokalmarkedet så vurderer kundene andre banker i mindre grad ettersom Sparebanken Øst oppleves som mer kjent og tryggere. Det kan også ha motsatt effekt, nemlig at ved at kunden har valgt banken fordi den er nærmest, eller fordi den er mest synlig, men at dette ikke skaper noen lojalitet slik at kundene lettere bytter til andre banker dersom de mottar bedre tilbud der. I Oslo antas det at kundene har motsatt effekt. Kundene har mange alternativer til Sparebanken Øst, men kundene er typisk i banken fordi de har et godt forhold til rådgiveren sin, eller de har fått banken anbefalt av bekjente. De har altså ikke valgt å være kunder i banken fordi de kjenner til Sparebanken Øst og fordi banken er synlig i lokalmiljøet, men av andre grunner.

7.2 Modell 2

I modell 2 så benyttes interaksjonsvariabler for å estimere sannsynligheten for exit for treatmentgruppen vs kontrollgruppen som gjort rede for i kapittel 5.2. Lånestørrelse er fordelt over ni persentilgrupper og er de eneste forklaringsvariablene. Persentilgruppene er hver representert ved en dummyvariabel. I denne modellen selekteres det mellom måned en og måned tre der oppgaven antar renteeffekten gjelder i måned en og ikke i måned tre. I måned en er som tidligere nevnt alle lån utsatt for renteøkning hensyntatt og studien observerer en avslutning av 308 lån. I måned tre gjenstår de samme lånene, men nå er de 308 lånene fra treatmentgruppen selektert bort i tillegg til 123 lån som ble avsluttet i måned to.

Modellen utledes ved å bruke følgende regresjon, som fungerer likt som ligning (5.2.7) presentert kapittel 5.2.2:

$$P\{y = 1|X\} = \alpha + \beta_0 mnd_1 + \beta_1 p_1 + \beta_2 p_2 + \beta_3 p_3 + \beta_4 p_4 + \beta_5 p_6 + \beta_6 p_7 + \beta_7 p_8 + \beta_8 p_9 + \beta_9 p_1 mnd_1 + \beta_{10} p_2 mnd_1 + \beta_{11} p_3 mnd_1 + \beta_{12} p_4 mnd_1 + \beta_{13} p_6 mnd_1 + \beta_{14} p_7 mnd_1 + \beta_{15} p_8 mnd_1 + \beta_{16} p_9 mnd_1 + u \quad (7.2.1)$$

Resultatene fra modellen tolkes som den økte sannsynligheten for hver variabel i forhold til konstanttermen der koeffisientene for hver variabel representerer mersannsynligheten for

måned 3-gruppen. Dette er bestemt ved å inkludere $\beta_0 mnd_1$ i regresjonsligningen, slik at mnd_3 plasseres i konstanttermen. Koeffisientene for de enkelte persentilgruppene representerer derfor sannsynligheten for exit for kundegruppen til gitt persentilnivå som også tilhører måned tre, sammenlignet med konstanttermen. Koeffisientene for interaksjonsvariablene representerer mersannsynligheten for exit for treatmentgruppen sammenlignet med kontrollgruppen. Konstanttermen representerer her kontrollgruppen med lånestørrelse lik persentil 50, noe som kan sees ved at den ikke er inkludert i regresjonsligningen og er derfor lik α . Ved bruk av interaksjonsvariablene vil jeg få koeffesienter både for kontrollgruppen og for treatmentgruppen når variabel p_i endres.

7.3 Modell 3

I modell 3 så kontrollerer studien for de sosialdemografiske variablene i tillegg til å benytte interaksjonsvariabler som i modell 2, forklart i kapittel 5.2.2.

Modellen utledes ved å bruke følgende regresjon:

$$\begin{aligned}
 P\{y = 1|X\} = & \alpha + \beta_0 mnd_1 + \beta_1 p_1 + \beta_2 p_2 + \beta_3 p_3 + \beta_4 p_4 + \beta_5 p_6 + \beta_6 p_7 + \beta_7 p_8 + \beta_8 p_9 + \\
 & \beta_9 a_1 + \beta_{10} a_3 + \beta_{11} lk_1 + \beta_{12} lk_3 + \beta_{13} lokal + \beta_{14} oslo + \beta_{15} p_1 mnd1 + \beta_{16} p_2 mnd1 + \\
 & \beta_{17} p_3 mnd1 + \beta_{18} p_4 mnd1 + \beta_{19} p_6 mnd1 + \beta_{20} p_7 mnd1 + \beta_{21} p_8 mnd1 + \beta_{22} p_9 mnd1 + \\
 & \beta_{23} a_1 mnd1 + \beta_{24} a_3 mnd1 + \beta_{25} lk_1 mnd1 + \beta_{26} lk_3 mnd1 + \beta_{27} (lokal * mnd1) + \\
 & \beta_{28} (oslo * mnd1) + u
 \end{aligned}
 \tag{7.3.1}$$

Dette er en også en lineær sannsynlighetsmodell som estimerer sannsynlighetskoeffisienter for de ulike dummyvariablene. Modellen fungerer likt som modell 2, men har nå hensyntatt de sosialdemografiske variablene for alder, bostedsområde og kundeforholdets lengde. I konstanttermen er sannsynligheten for exit for kundene som avsluttet lånet i måned tre med aldersgruppe 2, kundeforholdlengde 2 og verken bosatt i lokalområdet eller i Oslo. Når oppgaven tolker koeffisientene så observeres sannsynligheten for kontrollgruppen sammenlignet med konstanttermen, for koeffisient β_1 til β_{14} . Koeffisient β_{15} til β_{28} viser sannsynligheten for treatmentgruppen sammenlignet med kontrollgruppen. Negative fortegn betyr at variabelen har en lavere sannsynlighet enn konstanttermen eller kontrollgruppen, avhengig av hvilken parameter en ser på.

7.4 Potensielle feil i modellen

Skjevhet fra utelatte variabler

I vårt datasett besitter studien noen variabler som jeg forventer gir oss noen spennende koeffisienter gjeldende mersannsynligheten for exit for treatmentgruppen sammenlignet med kontrollgruppen. Det som derimot er viktig er å diskutere den mulige korrelasjonen mellom forklaringsvariabler i modellene og utelatte variabler som ikke besittes i datasettet. Dette kalles «omitted variable bias» (Wooldridge, 2014, s. 90). Grunnen til at det er viktig å diskutere denne skjevheten fra utelatte variabler er fordi en må ta hensyn til denne potensielle effekten når jeg vurderer studiens resultater. Oppgaven besitter data om lånestørrelse, alder, lengden på kundeforhold, bostedsadresse og ved hvilket tidspunkt lån ble avsluttet i løpet av tremånedersperioden etter renteøkningen. Det er ikke utenkelig at det eksisterer korrelasjoner mellom utelatte variabler som påvirker resultatene og sammenligningen mellom treatmentgruppen og kontrollgruppen. Dette er nok særlig sannsynlig fordi kontrollgruppen vår ikke er perfekt og skulle gjerne vært basert på et tidspunkt som er lenger fra treatmentgruppen, enn det som benyttes i denne oppgaven. Eksempelvis kan det hende at lånestørrelse i stor grad påvirkes av om låntakerne er høyt utdannede eller ikke. Dersom vi antar en positiv sammenheng mellom lengden utdanning og lånestørrelse kan det være at resultatene for lånestørrelse ville vært forandret dersom vi tok høyde for utdanningsnivået til låntakerne, hvor muligens det er utdanningsnivået som gir økt sannsynlighet for exit og ikke lånestørrelse. Her har jeg en mulig skjevhet der den uavhengige variabelen lånestørrelse kan korrelere med utelatte variabler.

Simultan kausalitet

Dette er en effekt som kan oppstå og som påvirker resultatenes validitet og kan være problematisk for denne oppgaven. Validitet handler om hvor relevant datamaterialet er for å besvare problemstillingen (Grønmo, 2016, s 241) Det at kontrollgruppen og treatmentgruppen består av så mange av de samme kundene samt at studien har et så kort tidsmellomrom mellom dem kan påvirke resultatenes validitet gjennom simultan kausalitet. En slik effekt kommer av at de uavhengige variablene påvirkes av den avhengige variabelen, i vårt tilfelle, at kunder avslutter boliglånet.

Eksempler på mulige tilfeller av kausalitet kan være ved resultatene for sannsynligheten for exit ved ulike lånestørrelser. Ved en renteøkning tyder forskning på at det er de med størst lån

som sparer mest på å bytte bank. Det kan tenkes at de som avslutter lånene i løpet av måned en påvirker kontrollgruppen ved at særlige rentesensitive og store lånekunder forsvinner og trekker med seg sannsynligheten for exit fra kontrollgruppen. Kontrollgruppen representerer den kundegruppen som ikke blir utsatt for en renteøkning (jeg antar effekten har gått ut i løpet av to måneder). Vi ser fra tabell 1 at 44 lånekunder av 525 mulige avslutter lånet sitt fra gruppen med de største lånene i måned en, men at kun 12 gjør det i måned tre. Datasettet er selektert slik at de som avslutter lån i måned en ikke er med i måned tre. Simultan kausalitet inntreffer dersom sannsynligheten for exit er større for gruppen med de største lånene før renteøkningen enn for oppgavens kontrollgruppe. Dersom jeg har en kundegruppe som «trekker opp sannsynligheten» og de forsvinner i løpet av måned en eksisterer en slik effekt. Det er derfor viktig at drøftingen av resultatet at slike effekter blir tatt hensyn til.

Reliabiliten til datamaterialet

Oppgavens reliabilitet beskriver hvor troverdig studiets datamateriale er. Høy reliabilitet beskrives ved at man vil observere de samme dataene uansett når en samler dem inn (Grønmo, 2016, s240). Denne oppgaven benytter sekundære og kvantitative data. Kvantitative data er som regel tallfestede og representerer ofte et større utvalg enn det kvalitative data gjør. Kvantitative data blir derfor brukt til å generalisere et utvalg i en populasjon (Saunders, Lewis og Thornhill, 2009). Datagrunnlaget er produsert selv av banken. Dataene som benyttes kan hentes ut fra bankens arkiv og vil være like, uavhengige av når dataene hentes ut. Det vil derimot være mulig at det er gjort feil enten av IT-avdelingen i uttrekket, i forbindelse med datasortering eller i forbindelse med sensurering av personinformasjon. Det er derimot sik at datamaterialet gjerne skulle bestått av flere observasjoner, særlig av kundene som avslutter lånene sine. Det vurderes til at det er høy reliabilitet til at datasettet fordi det antageligvis ville sett likt ut om det ble hentet ut på nytt, men dersom en hadde endret hvilken renteheving en så på, eller man vurderte en annen bank, så ville vi sannsynligvis observert noe varierende resultater. Dette reflekteres også i resultatenes signifikansnivå fordi studien har få observasjoner. Spesielt variabel $p1$ i tabell tre i appendikset er mangelfull da jeg kun har syv observasjoner i kontrollgruppen.

8. Resultater

Formålet med oppgaven er å observere mersannsynligheten for at ulike kundegrupper avslutter boliglånet sitt når utlånsrenten økes. Oppgaven observerer hvordan exitsannsynligheten endres fra konstanttermen og videre fra kontrollgruppen når lånestørrelse, alder, bostedsområde og lengden på kundeforholdet varierer. Formålet med dette kapittelet er å presentere resultatene som er funnet. Gjennom lineære sannsynlighetsmodeller det legges til dummyvariabler og holder alle andre variabler konstante finner studien mersannsynligheten for den dummyvariabelen som får endret verdi fra 0 til 1, der de andre variablene har verdi lik 0. Gjennom modell 1 observeres hvilke variabler som gir økt sannsynlighet for exit, uten at det tas høyde for effekten av renteøkningen. Jeg ser kun på hva som kjennetegner kundene som avsluttet boliglånet i løpet av tremånedersperioden etter renteøkningen.

Videre utvides modellen ved å sammenligne treatmentgruppen med kontrollgruppen. Modellen finner den observerte exitsannsynligheten for begge grupper. Denne sammenlignes i første omgang mellom gruppene hvor lånestørrelse er den eneste forklaringsvariabelen. Jeg ser fra resultatene at enkelte av lånepersentilene har små forskjeller og/eller få observasjoner som resulterer i at noen av variablene blir lite signifikante.

Modellen blir utviklet for å ta høyde for de andre variablene som ikke er lånestørrelse, nemlig alder, lengde på kundeforhold og bostedsområde. Disse kalles i denne oppgaven som tidligere nevnt «sosialdemografiske variabler». Resultatene fra denne modellen viser at modellen styrkes ved å legge til de ekstra variablene, ved at forklaringskraften øker. Jeg ser også at den observerte sannsynligheten endres i varierende grad for de ulike lånepersentilene ettersom konstanttermen endres.

8.1 Modell 1 – Det gjøres ikke forskjell på treatmentgruppe og kontrollgruppe

Resultatene kan observeres i tabell 1 i appendikset.

Jeg ser at variablene for de ulike lånestørrelsene har lavere signifikansnivå enn de andre forklaringsvariablene, noe som antageligvis skyldes færre observasjoner. En kan likevel se en trend i styrkende sannsynlighet med økende lånestørrelse. Den mest signifikante variabelen er *lokal* som gir et bilde om at det faktisk er i lokalbanken det avsluttes flest lån. Dette er

interessant med tanke på at det antas at bankens markedsrett er sterkere her sammenlignet med bankens andre områder.

Resultatene for alder er i tråd med det som kan forventes med bakgrunn i den eksisterende litteraturen. Resultatene studien her legger frem tyder på at lavere alder gir høyere sannsynlighet for å avslutte boliglånet. Dette støtter opp under antagelsen om at ung alder er i sammenheng med lavere byttekostnader. En av de mulige årsakene kan være at unge er dyktigere på pc (Fjørtoft, 2017) og dette reduserer terskelen for å søke og lære seg en ny bank.

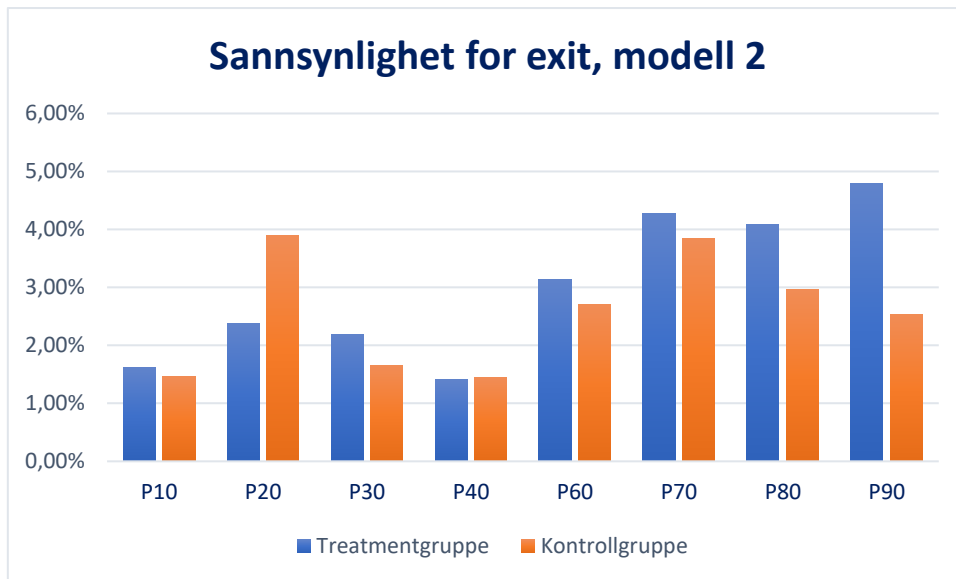
En interessant observasjon er at både de med korte og lange kundeforhold har økt sannsynlighet for exit, noe som virker lite intuitivt og som ikke støttes av den eksisterende litteraturen.

8.2 Modell 2 - Modell for mersannsynlighet for exit ved ulike lånestørrelser.

Modellen sammenligner sannsynligheten for avslutning av boliglån etter en renteøkning (treatmentgruppe) med sannsynligheten for avslutning når rentenivået holdes stille (kontrollgruppe). Modellen finner derfor mersannsynligheten for exit etter renteøkning for kunder som tilhører de ulike variablene som vist i kapittel 5.2. Variablene *Persentil 10* til *Persentil 90* representerer grupperingen av en enkelt boliglånskundes lånestørrelse. Ettersom studien har ni persentilgrupper så representerer eksempelvis *Persentil 10* de 11,1% med lavest lånestørrelse. Alle variablene er som tidligere nevnt på dummyform. Variabelen *Persentil 50* for kontrollgruppen er ikke tatt med i regresjonslikningen fordi den er plassert i konstanttermen. Jeg får derfor heller ikke estimert koeffesienten til *Persentil 50* for treatmentgruppen. Konstanttermen representerer sannsynligheten for exit for kundene som avsluttet lånet i måned tre og med lånestørrelse til *Persentil 50*. Tabellen som viser regresjonsresultatene fra modell 2 sees i appendikset under navnet «tabell 2».

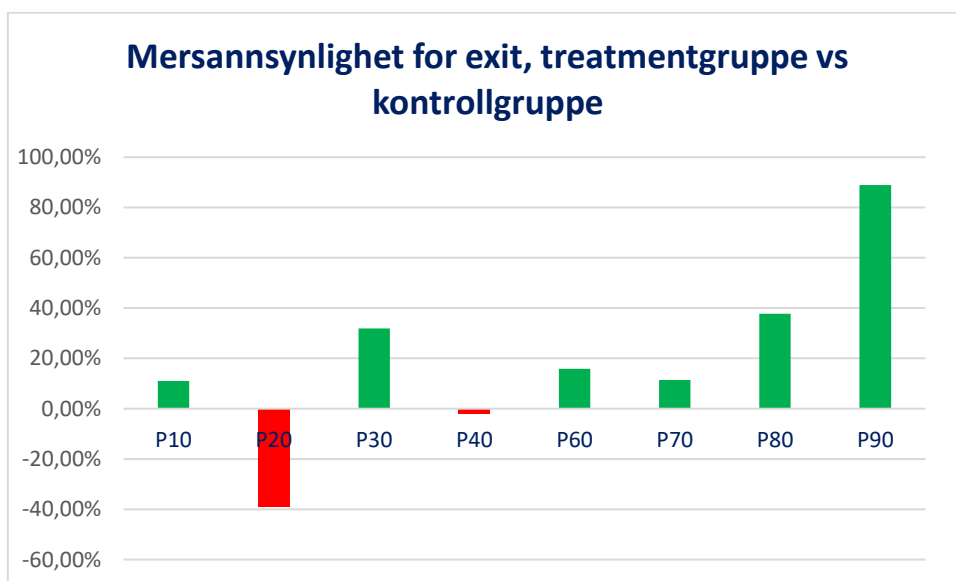
I kolonnen til høyre i tabell 2 presenteres mersannsynligheten for hver persentilgruppe. Jeg ser at signifikansen for en del av persentilgruppene er lav, noe som er et resultat av få observasjoner. Resultatene for *Persentil 60* ser en eksempelvis er lite signifikant. Kombinert med få observasjoner er grunnen at sannsynligheten for exit ligger nært sannsynligheten for konstantleddet, nemlig *Persentil 50* for kontrollgruppen.

Grafisk ser resultatene fra modellen slik ut.



Figur 3: Viser predikert exitsannsynlighet for de ulike lånestørrelsegruppene i modell 2

Jeg ser at særlig P20 skiller seg ut med noe som vurderes til en unormalt høy exitsannsynlighet for kontrollgruppen. Det observeres også en trend hvor høyere lånestørrelse predikerer økt exitsannsynlighet. Noe interessant er at det for kontrollgruppen predikeres en reduserende exitsannsynlighet etter P70.



Figur 4: Predikert mersannsynligheten for exit for treatmentgruppen sammenlignet med kontrollgruppen for alle variabler i modell 2

Modell 2 predikerer kun lavere exitsannsynlighet for P20 og P40 for treatmentgruppen enn kontrollgruppen. Studien viser at det er de største lånene som modellen predikerer den høyeste mersannsynligheten for exit.

8.2.1 Tolkning av resultater

Man ser fra tabell 2 i appendikset at kundegruppen som tilhører *persentil 90* både har den største sannsynligheten for exit i treatmentgruppen og mersannsynligheten sammenlignet med kontrollgruppen. Modellen presenterer også en positiv sammenheng mellom mersannsynligheten for exit for treatmentgruppen for de større persentilgruppene. Dette er i samsvar med hva som forventes å finne basert på den eksisterende litteraturen som er gjennomgått. Litteraturen tyder på at større boliglån har lavere byttekostnad. Det er også nærliggende å tenke at kunder med store boliglån har relativt god kredittscore. Dette begrunnes med at de norske boligforskriftene stiller strenge krav til sikkerhet og inntekt for at bankene skal kunne innvilge kundene boliglån. God kredittscore har studien også sett fra den eksisterende litteraturen at står i sammenheng med elastisk etterspørsel etter boliglån.

For de lavere persentilgruppene er svaret mindre intuitivt. Jeg ser at *persentil 20* har den største sannsynligheten for exit i kontrollgruppen, noe som strider med hva en kan forvente basert på den eksisterende litteraturen der byttekostnader burde medført en relativt dyr bytteprosess for kunder i *persentil20*-gruppen sammenlignet med høyere lånestørrelser. Videre har studien ingenting som tyder på at det bare er tilfeldig at det observeres en såpass stor exit for kontrollgruppen innenfor *persentil 20*. Det som er interessant er at det er akkurat denne persentilen som gir store avvik fra hva en kan forvente ettersom *Persentil 10*, *Persentil 30* og *Persentil 40* estimeres til å ha mer intuitiv sannsynlighet for exit. Intuitivt brukes fordi det forventes en stigende sannsynlighet for exit når lånestørrelsene økes.

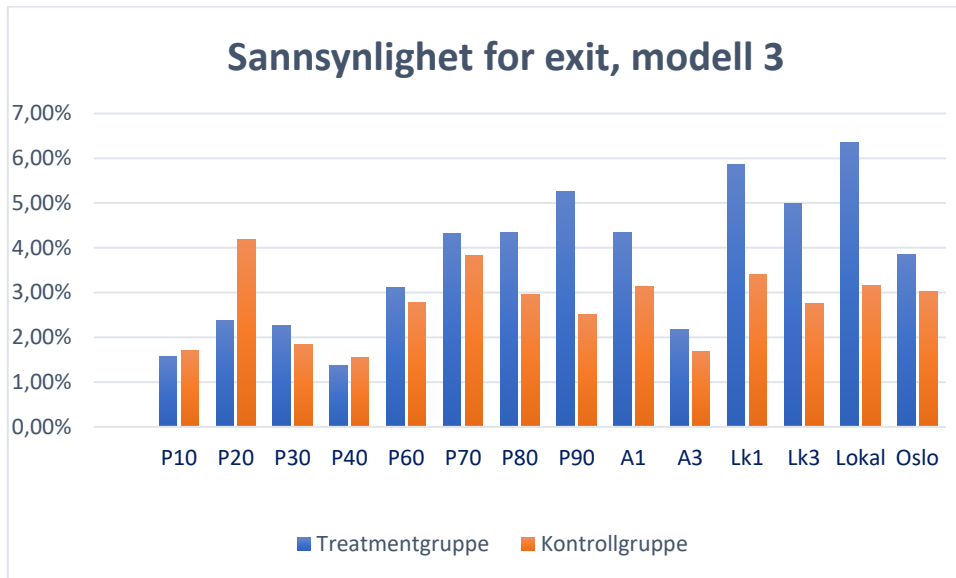
8.3 Modell 3 – Utvidet modell der det kontrolleres for sosialdemografiske forhold.

Modell 3 er bygget videre på modell 2 der det er kontrollert for alder, lengden på kundeforhold og bosted. Kontrolltermen representerer kundene som avsluttet lånet i måned tre, som tilhører aldersgruppe 2, lengde på kundeforhold gruppe 2 og som verken bor i Oslo eller tilhører lokalbanken. For videre opplysninger om variablene vises det til kapittel 4.2. Kontrolltermen er fra modell 2 nå flyttet opp i egen rad i tabell 3 som finnes i appendikset og som presenterer regresjonsresultatene fra modell 3.

Modell 3 har en høyere forklaringskraft enn modell 2 og jeg ser også hvordan de tilføyde variablene påvirker variablene for lånestørrelse. Dette er fordi konstanttermen nå i tillegg til å

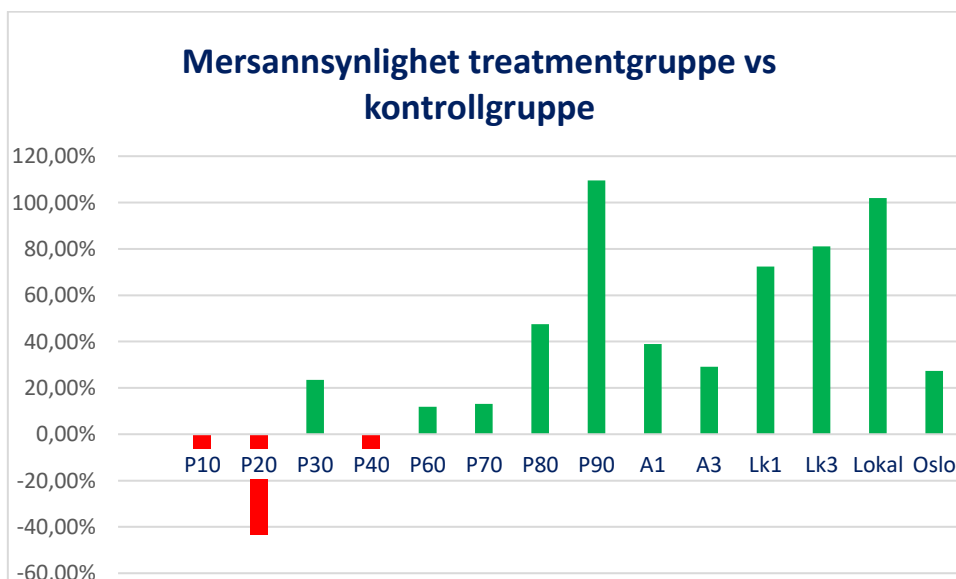
inneholde *persentil 50* for kontrollgruppen også inneholder lengde på kundeforhold 2 og aldersgruppe 2.

I figur 4 ser man exitsannsynligheten for alle forklaringsvariabler i modell 3 for både treatmentgruppen og kontrollgruppen.



Figur 5: Viser predikert exitsannsynlighet for alle forklaringsvariabler i modell 3

Man ser hvordan sannsynligheten for exit ser ut til å stige med lånestørrelse. Man ser også hvordan særlig lengde på kundeforhold og variabelen *lokal* gir en økt sannsynlighet for exit. Sannsynligheten trender ganske likt mellom treatmentgruppen og kontrollgruppen, med unntak av en redusert exitsannsynlighet for kontrollgruppen for *P80* og *P90* sammenlignet med *P70*. Videre er det interessant å vurdere mersannsynligheten for de to gruppene, som gjort i figur 5.



Figur 6: Viser predikert mersannsynligheten for exit for treatmentgruppen sammenlignet med kontrollgruppen for alle variabler i modell 3

Man ser at exitsannsynligheten er kun lavere for treatmentgruppen for $P10$, $P20$ og $P30$. Grunnen til at så mange variabler gir større sannsynlighet for treatmentgruppen forklares ved at det i konstanttermen finnes variabelgrupper som tydeligvis gir lav sannsynlighet for exit for treatmentgruppen sammenlignet med kontrollgruppen. Hadde det i konstanttermen blitt plassert variablene $P90$, $A1$, $Lk3$ og $lokal$, hadde man observert en lavere sannsynlighet for de fleste om ikke alle variablene for treatmentgruppen vs kontrollgruppen. Hvis man sammenligner figur 6 med figur 4 ser man at de minste lånene i $P10$ har gått fra å gi økt mersannsynlighet til lavere mersannsynlighet for treatmentgruppen. Årsaken er de tillagte sosialdemografiske variablene og som påvirker resultatene.

8.4 Valg av modell

Oppgaven har nå produsert tre modeller, men jeg velger å bare ta hensyn til modell 2 og modell 3 med tanke på valg av modell. Dette er som tidligere nevnt fordi modell 1 ikke tar høyde for forskjellen mellom treatmentgruppen og kontrollgruppen, og er kun tatt med for å gi leseren et tidlig bilde av hvilke kundegrupper som har høyere sannsynlighet for exit.

Tabell 3 som viser resultatene fra regresjonsanalysen ved modell 3 viser at modell 3 har en sterkere forklaringskraft enn modell 2. Forklaringskraften er sammen med AIC og BIC blant de mest anerkjente metodene var å gjennomføre variabelseleksjon (Hastie & Witten, 2017). AIC og BIC, som henholdsvis står for «Akaike information criterion» og «Bayesian information criterion» og tar utgangspunkt i å predikere verdienes avvik fra de faktiske verdiene ved at de regner ut mean squared error. Modellens avvik fra de faktiske verdiene ønsker en å minimere, og det er dette variabelseleksjon baserer seg på. Det vil derimot ikke være hensiktsmessig å gjennomføre AIC og BIC ettersom testene krever at en deler datasettet i to, noe oppgaven ikke har nok observasjoner til å gjøre. Eksempelvis ser man at kontrollgruppen kun har syv observasjoner som avslutter lånet for Persentil 10.

Forklaringskraften R^2 kan derimot estimeres direkte i modellen. Ved å legge vekt på justert R^2 tar man hensyn til tillagte variabler ved å redusere antall frihetsgrader i modellen når en legger ved forklaringsvariabler. Slik vil justert forklaringskraft kun øke dersom variablene som legges til modellen er signifikante. Den justerte forklaringskraften er derfor et bedre mål på den reelle forklaringskraften (Hastie et al., 2017).

Forskjellen på ordinær og justert forklaringskraft kan presenteres slik.

$$\text{Justert forklaringskraft: } R^2 = 1 - \frac{\frac{RSS}{(n-d-1)}}{\frac{TSS}{(n-1)}}$$

$$\text{Ordinær forklaringskraft: } R^2 = \frac{TSS-RSS}{TSS} = 1 - \frac{RSS}{TSS}$$

Her er n lik antall observasjoner, TSS er total sum of squares, RSS er residual sum of squares og d er antall frihetsgrader.

Jeg observerer at justert R^2 for modell 2 er 0,0091 og R^2 for modell 3 er 0,0134. Ettersom modell 3 har høyere forklaringskraft velger vi å inkludere de sosialdemografiske variablene og konkluderer med at modell 3 er mest hensiktsmessig å benytte.

8.4.1 Test for multikollinearitet

I valget av modell er det viktig å passe på at ikke multikollineariteten er for høy.

Multikollinearitet oppstår dersom studien har høy korrelasjon mellom en eller flere av de uavhengige forklaringsvariablene. Dersom det eksisterer multikollinearitet kan dette medføre misvisende resultater. I en lineær sannsynlighetsmodell som jeg har, er normalt ikke dette et problem fordi man kun varierer en variabel og holder de andre konstante. Likevel kan det eksistere multikollinearitet mellom forklaringsvariabler og konstantermen, så det er hensiktsmessig å teste for en slik effekt. Multikollinearitet kan testes ved å gjennomføre en VIF-test. Dersom studien har verdier over 10, som er vanlig å bruke som standard, så tyder dette på relevant kollinearitet som normalt krever at modellen endres (Wooldridge, 2014, s. 98).

VIF-test for Multikollinearitet

Variabel	Treatmentgruppen	Kontrollgruppen
Persentil 10	4,01	3,79
Persentil 20	4	3,79
Persentil 30	3,97	3,76
Persentil 40	3,96	3,75
Persentil 60	3,96	3,75
Persentil 70	3,98	3,75
Persentil 80	3,99	3,76
Persentil 90	4,05	3,82

A1	2,69	2,34
A3	2,48	2,29
Lk1	3,33	2,89
Lk3	4	2,92
Lokal	4,56	2,95
Oslo	2,82	2,61
Gjennomsnitt	3,83	

Tabell 2 – VIF-test for multikollinearitet.

Fra tabell 2 kan man se at det ikke er noen enkelte uavhengige variabler som har en VIF-verdi over 10. Modellen har en gjennomsnittskollinearitet på 3,83 og kan derfor konkludere med at det ikke foreligger stor nok korrelasjon mellom forklaringsvariablene til at dette krever en endring av modellen eller at oppgaven burde benytte modell 2 istedenfor modell 3.

8.5 Drøfting av resultatene fra modell 3

De sosialdemografiske variablene som er lagt ved ser man gir sterke koeffisienter sammenlignet med lånepersentilene, noe som kan forklares ved at kundegruppen er delt opp i færre dummyvariabler som gir flere observasjoner ved hver enkelt variabel. Studien observerer også koeffisienter som gir interessante resultater ved sammenligning av exitsannsynlighet mellom treatmentgruppen og kontrollgruppen. Disse resultatene for de sosialdemografiske variablene ser man også er mer signifikante enn resultatene for de ulike lånestørrelsepersentilene. Dette kan nok forklares med at ved oppdeling av lånestørrelse i flere persentiler enn hva studien har for de sosialdemografiske variablene så fordeles også observasjonene over flere variabler som påvirker signifikansen.

Man ser fra regresjonsresultatene i tabell 3 at studien finner til 5%-signifikansnivå, signifikante koeffisienter for treatmentgruppen ved variablene for kundeforholdlengde og for lokalbanken. Det er interessant å se at lokalbanken gir en signifikant økt sannsynlighet for exit, og det samme for kundeforholdlengde der det virker som gruppen som etablerte kundeforhold i perioden mellom 01.01.2007 og 01.01.2014 har signifikant lavere sannsynlighet for exit sammenlignet med kundene med kortere og lengre kundeforhold. Variablene for lengden på kundeforhold kan særlig drøftes mot Kim, Kliger og Vale (2001) der «lock-in effects» er en vesentlig del av byttekostnadene. En kunne derfor forventet at

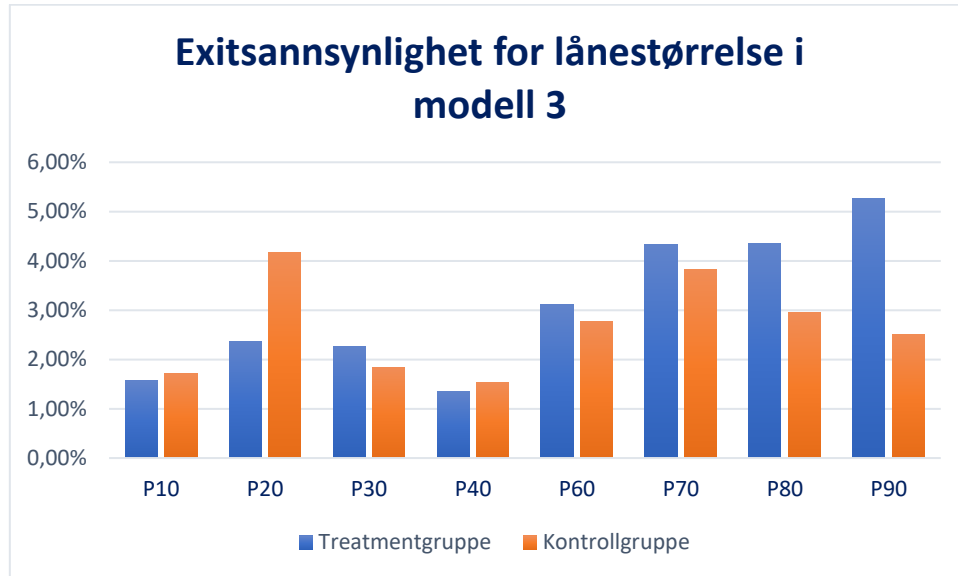
resultatene viste en sammenheng mellom korte kundeforhold og høy sannsynlighet for exit, men det finner studien altså ikke.

I modell 3 som i modell 2 så observeres en økt mersannsynlighet for de høyere persentilnivåene, men de er noe mer ekstreme i modell 3 enn i modell 2 når det kontrolleres for sosialdemografiske forhold. For alder så observerer jeg en høyere sannsynlighet for exit og også en sterkere mersannsynlighet for treatmentgruppen for de yngre lånetakerne. Dette virker å stå i samsvar med hva en kan forvente basert på den eksisterende litteraturen.

Modellen presenterer også en mersannsynlighet på 29% for den eldre kundegruppen, men sannsynligheten for exit er høyest for den yngre kundegruppen.

For variabelen som representerer bosted så ser man at det er kundene i lokalbanken som har høyest mersannsynlighet for exit for treatmentgruppen som dobles i forhold til kontrollgruppen. Kim, Kliger og Vale (2001) forklarer at større banker har mer mobile kunder. Det kan tenkes at kundene i lokalbanken har etablert kundeforhold fordi banken er synlig eller at banken er vanlig å bruke i nærmiljøet og at dette skaper en mer elastisk etterspørsel enn kundene utenfor lokalbanken som jeg antar har valgt å inngå kundeforhold av andre grunner. Variabelen som representerer kunder bosatt i Oslo gir også en økt mersannsynlighet for exit for treatmentgruppen, men man ser at denne er lavere enn resultatene for lokalbanken.

8.5.1 Variabel for lånestørrelse

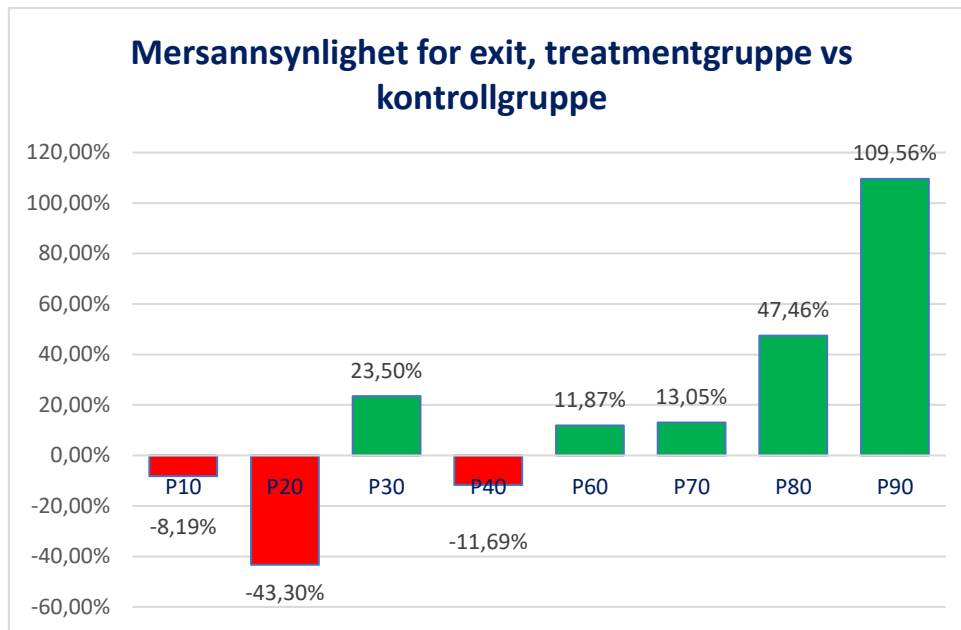


Figur 7: Viser predikert exitsannsynlighet for de ulike persentilnivåene når kunden verken tilhører lokal eller Oslo, men tilhører A2 og Lk2

For treatmentgruppen predikeres *Persentil 10*, *Persentil 20* og *Persentil 40* til lavere exitsannsynlighet for treatmentgruppen enn for kontrollgruppen. For kontrollgruppen er særlig *P20* høy, sammenlignet med de nærliggende lånenivåene og treatmentgruppen.

Resultatene tyder på at det eksisterer en positiv sammenheng mellom lånestørrelse og sannsynligheten for exit, og at dette også gjelder for mersannsynligheten mellom treatmentgruppen og kontrollgruppen, som sees i figur 8.

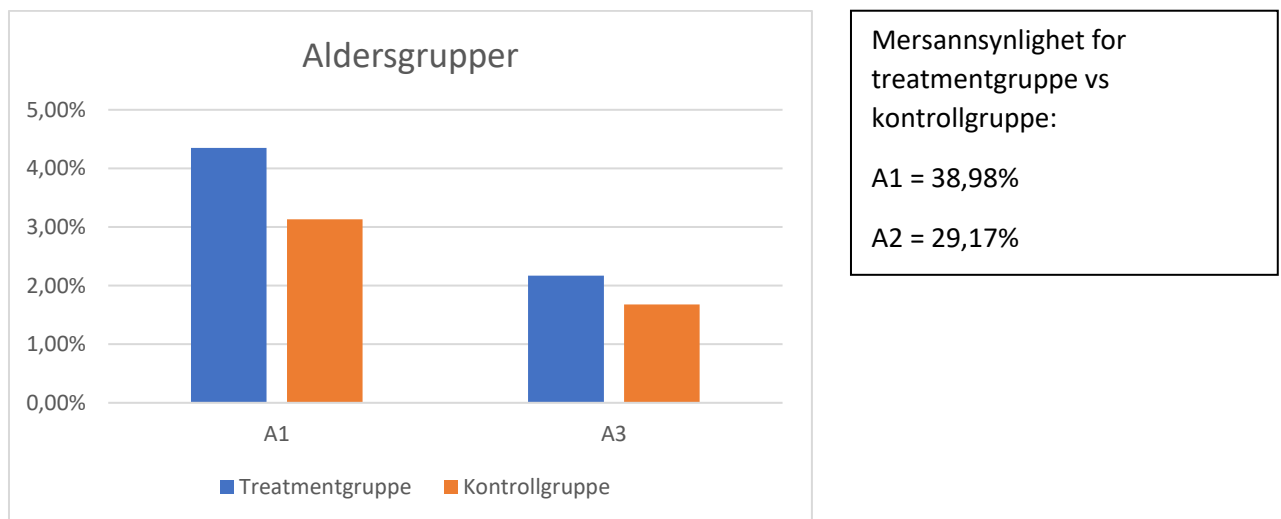
Man ser at sannsynligheten for exit for kontrollgruppen for *persentil 80* og *persentil 90* virker å være lavere enn det som kan forventes og det kan være et resultat av at oppgaven ikke har en perfekt kontrollgruppe som påvirkes av effekten av at måned en-gruppen avslutter lånene sine og videre påvirker sannsynligheten for den resterende gruppen, slik at studien muligens har en effekt fra simultan kausalitet. Likevel vurderes det til at oppgaven har bevis nok til å påstå at det eksisterer en mersannsynlighet for økt lånestørrelse etter renteøkning. Dette står i samsvar med forskning fra Alan og Loranth (2013), Lo (2017) og DeFusco og Paciork (2017) som alle presenterer en positiv sammenheng mellom elastisk etterspørsel etter lån med renteøkning.



Figur 8: Viser mersannsynligheten for exit for treatmentgruppen sammenlignet med kontrollgruppe, der kun lånestørrelse blir observert.

8.5.2 Variabel for alder

For aldersgruppene ser man at modellen predikerer en fallende exitsannsynlighet når alderen øker. Resultatene virker å være i tråd med eksisterende litteratur.



Figur 9: Viser predikert exitsannsynlighet for de to aldersgruppene som også tilhører gruppe P5, Lk2 og verken hører til lokal eller Oslo.

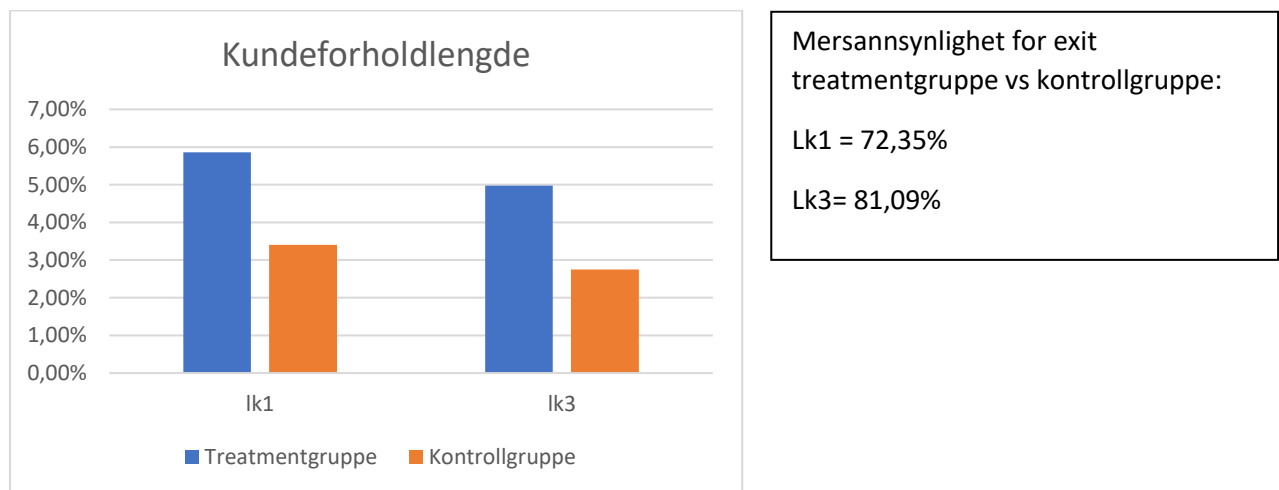
Man ser at sannsynligheten er fallende med alder, noe som støttes av den eksisterende litteraturen. Resultatene fra modellen viser også at mersannsynligheten for exit faller med alder. Det virker å være troverdige funn basert på forskningen som er gjennomgått.

Resultatene fra modellen viser at det er de yngste som har mest elastisk etterspørsel og som i

størst grad påvirkes av renteøkningen i forhold til avslutningssannsynlighet. Ettersom studien har flere observasjoner per dummyvariabel for alder enn for lånestørrelse så virker det mindre sannsynlig at simultan kausalitet påvirker resultatene. Det virker derfor som at oppgaven kan påstå at det eksisterer en negativ sammenheng mellom mersannsynligheten for exit i treatmentgruppen med hensyn på alder.

8.5.3 Variabel for lengde på kundeforhold.

Resultatene fra denne variabelen virker vanskeligere å tolke. Tabell 3 i appendikset viser signifikante koeffisienter med relativt stor mersannsynlighet mellom treatmentgruppen og kontrollgruppen. Problemet er at sannsynligheten ikke er lineær for kontrollgruppen med lavest sannsynlighet for exit for gruppe to som er kunder med 3-10 år lengde på kundeforholdene. Selv om studien ikke får presentert koeffisienten ved å endre til lk2 i modellen som man får med lk1 og lk3, så kan man se fra tabell 1 i kapittel 4.2.2 at det avsluttes relativt få lån i kundegruppe lk2 for treatmentgruppen. Det hadde vært naturlig å anta at sannsynligheten for exit var lineær over lengden på kundeforhold der en med utgangspunkt i Kim, Kliger og Vale (2001) og deres arbeid med byttekostnader og lock-in effekter burde gitt oss en negativ sammenheng for exit der sannsynligheten var høyest for gruppe lk1 med kortest kundeforhold.



Figur 10: Viser predikert exitsannsynlighet for de to gruppene for kundeforholdlengde. Kundene tilhører også gruppe P5, A2 og tilhører verken Oslo eller lokalbanken.

Basert på resultatene fra modellen ser jeg at det eksisterer en svakt positiv mersannsynlighet for exit for treatmentgruppen når kundeforhold blir lengre. Man ser at exitsannsynligheten er lavere for lk3 sammenlignet med lk1, noe som er i samsvar med Kim, Kliger og Vale (2001),

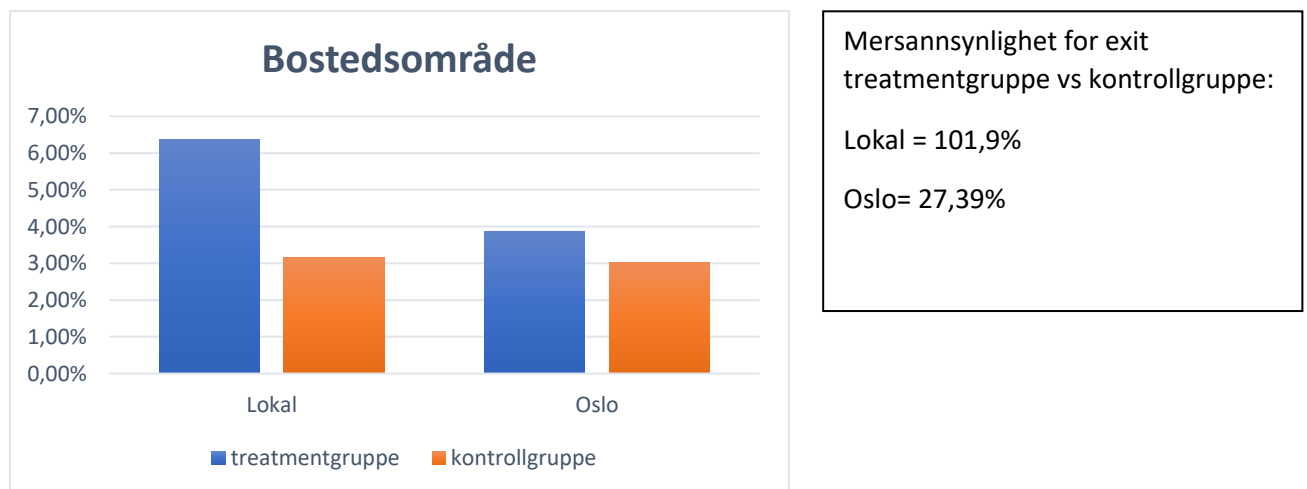
men oppgaven har ikke funnet litteratur som forklarer hvorfor det ser ut til at lk2 for treatmentgruppen har så lav exitsannsynlighet.

Det er interessant å se at lengden på kundeforhold har en så stor effekt på exitsannsynligheten og særlig mersannsynligheten for treatmentgruppen, der exitsannsynligheten mellom kontrollgruppen og treatmentgruppen varierer stort for lk1 og lk3. Resultatene tyder derfor på at rentesensitiviteten er størst for de med lengst og de med kortest kundeforhold, og at modellen predikerer en redusert effekt fra renteøkningen for kundene som har vært i banken fra 3 til 10 år.

8.5.4 Variabel for bostedsområde

Man ser fra tabell 3 i appendikset at resultatene for lokalbanken er svært interessante. Med 196 observasjoner og høy signifikans virker også resultatet pålitelig.

Sannsynlighetskoeffisienten som modellen vår beregner gir oss en dobling av sannsynlighet dersom modellen endrer fra konstantertermen (ikke Oslo, ikke lokal), til lokal.



Figur 11: Viser predikert exitsannsynlighet for lokalbanken vs ikke-lokalbanken. Den predikerte exitsannsynligheten gjelder i tillegg til lokal/Oslo for kunder som tilhører gruppen p5, a2 og lk2.

Hvordan dette resultatet står seg sammenlignet med den eksisterende litteraturen avhenger av hvordan man tolker de eksisterende resultatene. Den mest relevante forskningen å sammenligne med er igjen Vale, Kliger og Kims (2001) sin beregning av byttekostnader i den norske banksektoren. Der viser de til funn som sier at kunder av de største bankene er mest mobile og har lavest byttekostnader. En kan tenke at dette er sammenlignbart med lokalbanken der Sparebanken Øst er «større» enn de er i Oslo eller andre steder på Østlandet.

Det kan derfor virke som at kundene i lokalbanken har mest elastisk etterspørsel og derfor mest rentesensitive for en renteøkning.

Kundene som bor i Oslo ser man også har en noe forhøyet sannsynlighet for exit sammenlignet med kontrollgruppen. Mersannsynligheten for Oslogruppen er lavere, med 27,39% mersannsynlighet for treatmentgruppen mot hele 101,90% for lokalgruppen. Jeg vurderer derfor at studien har funnet resultater som lar oppgaven påstå at det eksisterer økt sannsynlighet for begge grupper, men at mersannsynligheten er størst for lokalgruppen som virker særlig sensitiv for renteøkningen.

9 Oppsummering og konklusjon

Formålet med studien har vært å predikere mersannsynligheten for at boliglån avsluttes for ulike kundegrupper etter en renteøkning sammenlignet med kundene som avslutter boliglånene når det ikke eksisterer en renteøkningseffekt. Slik finner studien hvilke kundegrupper som virker å være mest sensitive for en renteøkning. Relevant litteratur er gjennomgått for å finne hva som kan forventes av resultater, samt hva jeg kan sammenligne resultatene med.

Oppgaven benytter et datasett hentet fra Sparebanken Øst sine interne systemer som etter nødvendige ekskluderinger representerer 4727 boliglånengasjement. Disse lånene følges over en tremånedersperiode etter en renteøkning. Oppgaven sammenligner hvem som avslutter boliglånet den påfølgende måneden etter renteøkningen med en kontrollgruppe som det antas at avslutter boliglånet av andre grunner enn renteøkningen. Slik observeres hvilke boliglånskunder som er mest sensitive for en renteøkning. Oppgaven benytter en lineær sannsynlighetsmodell med interaksjonsvariabler for å skille mellom treatmentgruppen og kontrollgruppen og predikerer sannsynlighetskoeffisienter for de ulike kundegruppene når en variabel endres, men resten holdes konstante.

9.1 Hovedfunn

Den lineære sannsynlighetsmodellen oppgaven benytter seg av viser at jeg får resultater som peker på ulike mersannsynligheter der enkelte kundegrupper er mer rentesensitive enn andre.

Oppgaven finner resultater som indikerer en positiv sammenheng mellom lånestørrelse og sannsynlighet for exit. På tross av lite signifikante resultater, som ser ut til å komme av at oppgaven har delt opp lånestørrelse i ni persentilgrupper som medfører relativt få observasjoner i hver gruppe, sees det en tydelig trend der kunder med store boliglån er mest rentesensitiv etter en renteøkning. Oppgaven konkluderer også med at det eksisterer en negativ sammenheng mellom alder og rentesensitivitet. Resultatene fra alder og lånestørrelse virker å være i samsvar med eksisterende forskning, noe som er vanskeligere å si om lengden på kundeforhold. Her ser man at kundegruppene med både lengst og kortest kundeforhold er de som er mest sensitive for renteøkning sammenlignet med de som har vært kunder fra 3-10 år. Dette står i kontrast med det som er den mest nærliggende litteraturen for vurdering av denne variabelen, nemlig Kim, Kliger og Vales estimering av byttekostnader (2001). Det ville derfor vært interessant å gjennomføre lignende studier for å se om resultatene fra lengden på kundeforhold støttes av annen forskning. Oppgaven finner også at lokalbanken er de som reagerer mest på renteøkningen, der sannsynligheten for avslutning over dobles for kundene av lokalbanken sammenlignet med resten av boliglånskundene som heller ikke tilhører Oslo.

9.2 Kritikk, begrensninger og videre forskning

Det ville vært interessant å benytte modellen for å se om resultatene står seg når data hentes fra andre banker, eller andre tidsrom enn det som er gjort i denne oppgaven. Det hadde også vært hensiktsmessig å inneha en bedre kontrollgruppe der man er mer sikker på at kontrollgruppen representerer et riktig utvalg av kundene som avslutter boliglånene sine uten at renten er økt i nær fortid. Det ville også vært interessant å besitte flere variabler som kan være forklarende og videre styrke modellen. Slike variabler kunne vært inntekt, kjønn, og utdanning. I tillegg kunne det vært hensiktsmessig å gjennomføre en tilsvarende undersøkelse, men i en større bank der en kunne basert modellen på flere observasjoner.

En påbygning av utredningen kan være å gjennomføre en tilsvarende undersøkelse, men bruke en eller flere andre banker, samt benytte en annen variabelsammensetning enn det som denne oppgaven har gjort.

Avslutningsvis kunne det også vært naturlig å gå mer i dybden i bankens inntektsside hvor man kunne utredet forskjellene i bankens gevinst ved å øke renten for de ulike kundegruppene og se ved hvilke kunder banken tjener mest når renten settes opp.

Litteraturliste

Alan, Sule og Loranth, Gyongyi. (2013). Subprime Consumer Credit Demand: Evidence from a Lender's Pricing Experiment. *The Review of Financial Studies*. 26 (9), 2353-2374.

Alsberg, Ola og Lorentzen, Marius (2020). Skuffet over DNBs nye rentekutt: - Avspiser kundene. Tilgjengelig: <https://e24.no/boers-og-finans/i/JoyPkP/skuffet-over-dnbs-nye-rentekutt-avspiser-kundene>. Sist besøkt 11.12.2020.

Bankenes sikringsfond. (2019). *Medlemsbanker*. Tilgjengelig: <https://www.bankenessikringsfond.no/medlemsbanker/category820.html>.

Beckett, Sean. (2017). Why America's Homebuyers & Communities Rely on the 30-Year Fixed-Rate Mortgage. Tilgjengelig: http://www.freddiemac.com/perspectives/sean_beckett/20170410_homebuyers_communities_fixed_mortgage.page. Sist besøkt 13.12.2020.

DeFusco, Anthony A. og Paciorek, Andrew. (2017). The Interest Rate Elasticity of Mortgage Demand: Evidence from Bunching at the Conforming Loan Limit. *American Economic Journal: Economic Policy*. 9 (1), 210-240.

DnB. (2019). DnB priser for privatkunder. Tilgjengelig: <https://www.dnb.no/privat/priser/laan.html>.

Fico. (2020). FICO® Score. Tilgjengelig: <https://www.fico.com/en/products/fico-score>. Sist Besøkt: 15.12.2020

Finansdepartementet. (2020). Forskrift om krav til nye utlån med pant i bolig (boliglånsforskriften). Tilgjengelig: <https://www.regjeringen.no/no/tema/okonomi-og-budsjett/finansmarkedene/boliglansforskriften-1.-januar-202031.-desember-2020/id2679449/>

Finans Norge (2019). Bankstatistikk, Markedsandeler - forvaltningskapital. [Finansnorge.no](https://finansnorge.no)

Fjørtoft, Torstein Otterlei. (2017). Unge og høyt utdannede er flinkest foran PC-en. Tilgjengelig: <https://www.ssb.no/teknologi-og-innovasjon/artikler-og-publikasjoner/unge-og-hoyt-utdannede-er-flinkest-foran-pc-en>. Sist besøkt 17.12.20.

Grønmo, S. (2016). Samfunnsvitenskapelige metoder. 2. utg. Oslo: Fagbokforlaget,

Haraldsdottir Gudrun, Wettre, Sigbjørn Ragnar og Tveita Ola. (2018). Boliglånsrentene faller fortsatt. Tilgjengelig: <https://www.ssb.no/bank-og-finansmarked/artikler-og-publikasjoner/boliglansrentene-faller-fortsatt>. Sist besøkt: 12.12.2020

Hastie, T. James, G. & Witten, D. (2017). An Introduction to Statistical Learning. With Applications in R. (8. utgave). New York: Springer.

Karlan, Dean og Zinman, Jonathan. (2019). Long-Run Price Elasticities of Demand for Credit: Evidence from a countrywide Field Experiment in Mexico. *Review of Economic Studies*. 86 (4), 1704–1746.

Kim, Moshe, Kliger, Doron and Vale, Bent. (2001). Estimating switching costs: the case of banking. *Journal of Financial Intermediation*. Volume 12, 25-56.

Klemperer, Paul. (1987). The Competitiveness of Markets with Switching Costs. *The RAND Journal of Economics*. 18, 138-150.

Klemperer, Paul. (1995). Competition when consumers have switching costs: An overview with applications to industrial organization macroeconomics, and international trade. *Review of Economic Studies* 62, 515–539.

Lo, Stephanie H. (2017). What is the Microelasticity of Mortgage Demand to Interest Rates? Harvard JCHS. Tilgjengelig: <https://www.jchs.harvard.edu/blog/what-impact-do-changing-interest-rates-have-on-mortgage-demand>

Melgaard Holøien, Aud. Nygård, Geir og Holseter, Anne Marie R.. (2017). Høyest utdanningsnivå i Oslo. Tilgjengelig: <https://www.ssb.no/utdanning/artikler-og-publikasjoner/hoyest-utdanningsniva-i-oslo>. Sist besøkt 19.12.20.

Norges Bank. (2020). Endringer i styringsrenten. Tilgjengelig: <https://www.norges-bank.no/tema/pengepolitikk/Styringsrenten/Styringsrenten-Oversikt-over-rentemoter-og-endringer-i-styringsrenten/>. Sist besøkt 10.11.2020.

Nybygger. (2020). HISTORISK LAV RENTE. Tilgjengelig: nybygger.no. Sist besøkt 11.12.2020.

Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research methods for business students* fifth edition. Harlow: Pearson Education Limited.

Sbanken. (2020). Om Sbanken. Tilgjengelig: sbanken.no. Sist besøkt 12.12.2020.

Sparebanken Øst. (2020). Våre kontorer og åpningstider. Tilgjengelig:
<https://www.oest.no/kontor/>. Sist besøkt 14.12.2020.

Sparebankforeningen (2019). Antall sparebanker 1922 - 2019. Sparebankforeningen.no

Sparebank 1 Nord-Norge. (2019). Visjon og forretningsidé. tilgjengelig:
<https://www.sparebank1.no/nb/nord-norge/om-oss/om-banken/visjon-og-forretningside.html>

Stata. (2020). *append - Append datasets*. Tilgjengelig:
<https://www.stata.com/manuals13/dappend.pdf#dappend>.

Torres-Reyna, Oscar (2011). Merge/Append using Stata. Tilgjengelig:
<https://dss.princeton.edu/training/>. Sist besøkt 19.12.20.

Wooldridge, Jeffrey M. (2014). *Introductory Econometrics, A modern approach*. 5th ed.
Michigan State University: South-Western Cengage Learning.

Appendiks

Tabell 1: Regresjonsresultatene for modell 1

Modell 1 – Sannsynlighet for exit

Avhengig variabel: Sannsynlighet for Exit		Observasjoner=4723		
Variabel	Koeffisient	Standardavvik	t	Signifikans
<i>konstant</i>	.0830044	.0168536	4.93	0.000
<i>p1</i>	-.0197748	.0197737	-1.00	0.317
<i>p2</i>	-.0106259	.0197898	-0.54	0.591
<i>p3</i>	-.0229595	.0196801	-1.17	0.243
<i>p4</i>	-.0279035	.0196586	-1.42	0.156
<i>p6</i>	-.0125074	.0196487	-0.64	0.524
<i>p7</i>	.0182133	.0196435	0.93	0.354
<i>p8</i>	.0091072	.0196643	0.46	0.643
<i>p9</i>	.0102695	.0197331	0.52	0.603
<i>lokal</i>	.0332936	.0110102	3.02	0.003
<i>oslo</i>	.0146797	.0150034	0.98	0.328
<i>a1</i>	.0281747	.0113227	2.49	0.013
<i>a3</i>	-.01915	.0137067	-1.40	0.162
<i>lk1</i>	.0310265	.0127903	2.43	0.015
<i>lk3</i>	.0160793	.0110661	1.45	0.146

Signifikansnivå: $p < 0,05$. p_5 , a_2 og lk_2 er fjernet grunnet multikollinearitet

Note: Tabellen viser sannsynligheten i Sparebanken Øst sin portefølje i løpet av en tremånedersperiode etter renteøkning. Tabellen skiller videre ikke mellom treatment- og kontrollgruppe. Koeffesienten for hver variabel viser mersannsynligheten for exit dersom en legger til variabelen sammenlignet med konstanttermen som i dette tilfellet er kundene som avslutter i måned 1 + måned 2 + måned 3 tilhørende prosentil 50, a_2 og lk_2 .

Tabell 2: Regresjonsresultatene for modell 2.

Sannsynlighet for exit - Treatmentgruppe vs kontrollgruppe: Lånestørrelse i prosentilform			
variabel i kursiv	Sannsynlighet for exit treatmentgruppe	Sannsynlighet for exit kontrollgruppe	Mersannsynlighet for exit treatmentgruppe vs kontrollgruppe
<i>Prosentil 10</i>	1,62 % (0,0173101)	1,46 % (0,0095179)	10,96 %
	0,932**	0,275**	
Observasjoner	27	7	
<i>Prosentil 20</i>	2,38 % (0,0189505)	3,90 % (0,0117252)	-38,97 %
	0,415**	0,481**	
Observasjoner	31	19	
<i>Prosentil 30</i>	2,19 % (0,0176833)	1,66 % (0,0097094)	31,93 %
	0,779**	0,342**	
Observasjoner	30	8	
<i>Prosentil 40</i>	1,42 % (0,017191)	1,45 % (0,0094918)	-2,07 %
	0,988**	0,270**	
Observasjoner	26	7	
<i>Prosentil 50</i>		2,95 % (0,007768)	
		0,02**	
Observasjoner	34	14	
<i>Prosentil 60</i>	3,14 % (0,0187003)	2,71 % (0,01074)	15,87 %
	0,819**	0,860**	
Observasjoner	35	13	
<i>Prosentil 70</i>	4,28 % (0,019804)	3,84 % (0,0117979)	11,46 %
	0,814**	0,514**	
Observasjoner	41	18	
<i>Prosentil 80</i>	4,09 % (0,0192777)	2,97 % (0,0110316)	37,71 %
	0,552**	0,985**	
Observasjoner	40	14	
<i>Prosentil 90</i>	4,80 % (0,0193262)	2,54 % (0,0106274)	88,98 %
	0,224**	0,766**	
Observasjoner	44	12	
Forklaringskraft	0,0109		
Justert forklaringskraft	0,0091		
Observasjoner totalt	308	112	

Signifikansnivå**: $p < 0,05$. p_5 , a_2 og lk_2 er fjernet grunnet multikollinearitet. Robuste standardavvik i parentes.

Note: Tabellen viser regresjonsresultatene fra modell 2 hvor sannsynligheten for exit er den avhengige variabelen og de uavhengige variablene er ni dummyvaribler som representerer kunders lånestørrelse på prosentilnivå der gruppe en er lavest og gruppe ni er størst. Sannsynligheten for exit vises i kolonnen til venstre for treatmentgruppen og i midten for kontrollgruppen. I kolonnen til høyre vises mersannsynligheten for exit for treatmentgruppen sammenlignet med kontrollgruppen. Konstanttermen i denne modellen er kunder i måned 3-gruppen med lånestørrelse til prosentil 50. Disse kundene har en estimert sannsynlighet for exit til 2,95%. Variabler med lavere sannsynlighet enn dette har altså en mindre sannsynlighet for exit sammenlignet med konstanttermen.

Tabell 3: Regresjonsresultatene for modell 3

Sannsynlighet for exit - Treatmentgruppe vs kontrollgruppe: Justert for Alder, bosted og lengde på kundeforhold

	Forklaringskraft	0,0165		
	Justert forklaringskraft	0,0134		
Variablene i kursiv	Sannsynlighet exit treatmentgruppe	Sannsynlighet exit kontrollgruppe	Mersannsynlighet exit, treatmentgruppe vs kontrollgruppe	
<i>Konstantterm</i>			3,00 % (0,0089914) 0,010**	Konstanttermen representerer kunder i kontrollgruppen med prosentil 50, lk2, a2 og som verken tilhører Oslo eller lokal
<i>Prosentil 10</i>	1,57 % (0,0173548) 0,942**	27	1,71 % (0,0096517) 0,345**	-8,19 %
Observasjoner				
<i>Prosentil 20</i>	2,37 % (0,0187783) 0,338**	27	4,18 % (0,0117137) 0,386**	-43,30 %
Observasjoner				
<i>Prosentil 30</i>	2,26 % (0,0176949) 0,820**	31	1,83 % (0,0097537) 0,389**	23,50 %
Observasjoner				
<i>Prosentil 40</i>	1,36 % (0,017207) 0,924**	30	1,54 % (0,0095264) 0,281**	-11,69 %
Observasjoner				
<i>Prosentil 60</i>	3,11 % (0,018706) 0,861**	26	2,78 % (0,0107208) 0,873**	11,87 %
Observasjoner				
<i>Prosentil 70</i>	4,33 % (0,0197631) 0,791**	35	3,83 % (0,0118687) 0,541**	13,05 %
Observasjoner				
<i>Prosentil 80</i>	4,35 % (0,0192348) 0,456**	41	2,95 % (0,0109513) 0,973**	47,46 %
Observasjoner				
<i>Prosentil 90</i>	5,26 % (0,0195525) 0,146**	40	2,51 % (0,010805) 0,721**	109,56 %
Observasjoner				
<i>A1</i>	4,35 % (0,011333) 0,262**	44	3,13 % (0,0063763) 0,868**	38,98 %
Observasjoner				
<i>A3</i>	2,17 % (0,0110108) 0,706**	92	1,68 % (0,0110108) 0,154**	29,17 %
Observasjoner				
<i>Lk1</i>	5,86 % (0,0117733) 0,043**	34	3,40 % (0,007061) 0,650**	72,35 %
Observasjoner				
<i>Lk3</i>	4,98 % (0,0099571) 0,034**	83	2,75 % (0,005459) 0,744**	81,09 %
Observasjoner				
<i>Lokal</i>	6,36 % (0,0097835) 0,002**	153	3,15 % (0,0052281) 0,839**	101,90 %
Observasjoner				
<i>Oslo</i>	3,86 % (0,0134742) 0,560**	196	3,03 % (0,00802) 0,978**	27,39 %
Observasjoner				
		39	16	

Robuste standardavvik i parentes: ** 5-prosent signifikansnivå

Note: Tabellen viser det samme som tabell 2, men er nå kontrollert for sosialdemografiske forhold. Konstanttermen er nå flyttet opp i egen rad der den tidligere fantes i prosentil 50 for kontrollgruppen. Konstanttermen er fremdeles for gruppen i måned 3 med lånestørrelse i prosentil 50, men tilhører også nå gruppe 2 for alder, gruppe 2 for lengden på kundeforhold og bor verken i Oslo eller bankens lokalområde.