

Rapport
5/2023

**Effektanalyse av redusert
minstesats for mottakere
av
arbeidsavklaringspenger
(AAP) under 25 år og
avvikling av ung ufør-
tillegget**

Øystein M. Hernæs
Ole Røgeberg
Helene Berg
Gro Malene Vestøl

Rapport 5/2023

Effektanalyse av redusert minstesats for mottakere av arbeidsavklaringspenger (AAP) under 25 år og avvikling av ung ufør-tillegget

Øystein M. Hernæs
Ole Røgeberg
Helene Berg
Gro Malene Vestøl

Kontakt: www.frisch.uio.no

Rapport fra prosjektet "Effektanalyse av regelendring for AAP" (1773), finansiert av NAV (A137162-SAK18/3634)

ISBN: 978-82-7988-299-2
ISSN: 2704-047X

Effektanalyse av redusert minstesats for mottakere av arbeidsavklaringspenger (AAP) under 25 år og avvikling av ung ufør-tillegget*

Øystein M. Hernæs og Ole Røgeberg, Frischsenteret
Helene Berg og Gro Malene Vestøl, Proba

Sammendrag

Formålet med arbeidsavklaringspenger (AAP) er å gi inntekt til personer som har nedsatt arbeidsevne på grunn av sykdom eller skade, mens de får hjelp fra NAV. Fra 1.2.2020 ble minstesatsen for AAP for mottakere under 25 år redusert, og ung-ufør-tillegget for AAP-mottakere ble fjernet. Disse endringene var ment å gjøre minstenivået på AAP likere andre ordninger og redusere sykdomsfokuset.

Hoveddelen av rapporten omhandler analyse av redusert minstesats for mottakere av arbeidsavklaringspenger (AAP) under 25 år. Vi fokuserer på denne endringen fordi den påvirket alle nye mottakere av AAP under 25 år som kvalifiserte for minstesatsen. Ung ufør-tillegget var det derimot bare en liten andel som fikk, og mange som fikk det, gjorde det først etter AAP-forløpet. I tillegg har vi ikke hatt tilgjengelig individdata om mottak av ung ufør-tillegget.

Arbeidet med evalueringen av endringene ble delt i to faser. Formålet med dette var å sikre at analytiske valg ble gjort på bakgrunn av faglige vurderinger og uavhengig av resultater. Første fase av prosjektet utviklet analysestrategi og -modell på bakgrunn av data fra før regelendringene trådte i kraft. Fase 1 ble avsluttet ved at analytiske valg ble endelig avgjort og loggført. Deretter ble data fra etter endringen gjort tilgjengelig, og fase 2, som besto av estimering av reformeffekter, startet.

Metodevalget i fase 1 ble foretatt ved å teste tre ulike statistiske strategier på data fra før regelendringen trådte i kraft. På bakgrunn av disse analysene ble det besluttet å bruke en statistisk modell som antar at tidsutviklingen i innstrømmingen til AAP blant personer under 25 år med lav tidligere inntekt i utgangspunktet hadde en *alders-spesifikk trend* og en *inntekts-spesifikk trend*. Disse to trendene kan vi anslå fra grupper som ikke ble påvirket av regelendringene. Den alders-spesifikke trenden ble anslått med data for personer med inntekt over 3G, mens den inntekts-spesifikke trenden ble anslått med data for personer som var eldre enn 25 år.

Analysene på pre-reform-data indikerte at den endelige analysen ville ha svak presisjon, og dette viste seg å være tilfellet. De kvalitative intervjuene ga videre et klart inntrykk av at saksbehandlere så det som lite sannsynlig at regelendringene hadde fått store konsekvenser for innstrømmingen til AAP. Vi fokuserer derfor på resultatene fra en modellvariant som legger til grunn at AAP-sannsynligheten neppe ville bli mer enn halvert eller mer enn doblet som følge av regelendringene. Dette innbefatter likevel betydelige potensielle effekter. Det skal også nevnes at det at reformtidspunktet sammenfalt med koronaepidemien og samfunnets respons på denne i 2020 bidrar til ytterligere usikkerhet.

Resultatene fra vår foretrukne modellspekifisering tilsier at regelendringene reduserte antallet nye AAP-tilfeller blant unge med lav inntekt med 208 tilfeller i 2020 og 170 tilfeller i 2021. Ettersom

* Dette prosjektet har vært utført på oppdrag fra Arbeids- og velferdsdirektoratet. Takk til Helle Sundby, Bernhard Weigel, Knut Røed og representanter fra Arbeids- og inkluderingsdepartementet for nyttige kommentarer, Eirik Grønlien for data om ung ufør-tillegget og en lang rekke NAV-veiledere for intervjuer og diskusjoner. Registerdata på utlån fra Statistisk sentralbyrå har vært essensielt for å gjennomføre prosjektet.

det er rundt 3500 nye AAP-tilfeller årlig fra denne gruppen som helhet, tilsier disse «punktestimaterne» en nedgang på rundt 5-6 prosent. Estimaterne er imidlertid innenfor den variasjonen fra år til år som forekom i perioden før regelendringene. Vi kan derfor ikke utelukke at de skyldes tilfeldig variasjon som ikke hadde med regelendringene å gjøre.

I tråd med analyseplanen som ble utarbeidet i prosjektets første fase, ble den statistiske modellen også benyttet på andre utfall. Her var tanken at de som eventuelt ikke begynte å motta AAP fordi minstesatsen ble redusert, i stedet kunne dukke opp som sosialhjelpsmottagere, i utdanningssystemet eller på kvalifiseringsprogrammet. Her var det kun i analysen av kvalifiseringsprogrammet at analysetilnærmingen viste seg egnet. Vi finner ingen tegn til at regelendringene påvirket overgang til kvalifiseringsprogrammet.

Til slutt analyserer vi resultater for faktiske mottakere av arbeidsavklaringspenger. Ettersom vi i de tidligere analysene fant tegn til at innstrømmingen til AAP ble redusert, må vi være varsomme med tolkningen av resultatene for faktiske mottakere, ettersom sammensetningen av denne gruppa kan ha blitt endret. Når vi sammenligner nye mottagere fra før og etter reformen på utfall som tidligere karakterer ved avslutning av grunnskolen og andelen som har fullført videregående skole ser vi imidlertid ingen tegn til slik komposisjonsendring av gruppen.

Analysene av de faktiske mottakerne finner at andelen AAP-mottakere under 25 år med supplerende sosialhjelp har økt tydelig fra og med 2020. Det er naturlig å se dette i sammenhengen med reduksjonen i minstesatsen. Funnet er også konsistent med de bekymringene vi møtte da vi intervjuet veiledere fra tre forskjellige NAV-kontor. De var bekymret for at regelendringene ville føre til en kraftig økning i supplerende sosialhjelp for de yngste mottakerne. Imidlertid må det understrekes at det fortsatt er få AAP-mottakere som mottar supplerende sosialhjelp også etter regelendringene. Blant nye mottakere av AAP under 25 år i 2020 og 2021 var det i en gjennomsnittlig måned 11 prosent som mottok sosialhjelp, mot 7 prosent hos nye mottakere i alderen 25-28 år.

Når det gjelder avgang fra AAP og tilgang til uføretrygd, arbeid eller utdanning, finner vi ingen klare endringer fra og med 2020. Avgangsmønsteret fra AAP er tilnærmet likt for yngre og eldre nye mottakere over tid. Mottakere av AAP i alderen 18-23 ser ut til å bli noe raskere avklart til uføretrygd enn mottakere i alderen 25-28. Dette var tilfellet også før regelendringene, men det er en svak antydning til dette kan ha økt noe for nye AAP-mottakere i alderen 18-23 år fra 2020. Imidlertid er forskjellene små og usikkerheten stor.

Innholdsfortegnelse

| | |
|---|-----------|
| Sammendrag | 1 |
| 1 Oppdraget – formål og problemstillinger | 4 |
| 1.1 Bakgrunn | 4 |
| 1.2 Problemstillinger | 5 |
| 1.3 Koronaepidemien | 6 |
| 1.4 Data | 6 |
| 2 Hva endret seg i praksis: Regelendringene «sett fra innsiden» | 8 |
| 2.1 kvalitative intervjuer med veiledere i NAV | 8 |
| 2.2 Flere regelendringer på kort tid | 8 |
| 2.3 Problematikk blant unge som veilederne har oppfølgingsansvar for | 9 |
| 2.4 Tilstrømming til AAP: Veilederne har ikke merket noen endring | 9 |
| 2.5 Økt bruk av supplerende sosialhjelp | 10 |
| 2.6 Overgang til arbeid/utdanning, eller overgang til uføretrygd? | 11 |
| 2.7 Betydningen av koronaepidemien..... | 11 |
| 2.8 Konklusjoner | 12 |
| 3 Analysestrategi | 13 |
| 3.1 Prosjektet..... | 13 |
| 3.2 Fase 1: Pre-registrering | 13 |
| 3.3 Pre-registrering i vårt prosjekt..... | 13 |
| 3.4 Vår analysestrategi | 15 |
| 3.5 Sekundær-analyse: Endret søketidspunkt grunnet regelinnføring | 18 |
| 4 Resultater fra fase 1: Analyser av pre-reform-data | 23 |
| 4.1 Analyse 1: Finner metodene null-effekt på placebo-reform? | 23 |
| 4.2 Analyse 2: Var det en felles trend i pre-reform perioden? | 24 |
| 4.3 Konklusjon etter analyser av pre-reform-data | 25 |
| 5 Resultater fra fase 2: Analyser av post-reformdata | 26 |
| 5.1 Utfall: innstrømming til AAP | 26 |
| 5.2 Utfall: overgang til kvalifiseringsprogrammet, sosialhjelp, uføretrygd, utdanning | 28 |
| 6 Resultater for faktiske AAP-mottakere | 31 |
| 6.1 Supplerende sosialhjelp blant AAP-mottagere | 31 |
| 6.2 Avgang fra AAP og overgang til uføretrygd, utdanning og arbeid | 34 |
| 7 Konklusjoner | 39 |
| Referanser | 41 |
| Vedlegg | 42 |
| Nærmere om vår implementering av DiD og DiDiD | 42 |
| Implementering av analyse-strategi – den «nye» DiD-litteraturen | 44 |

1 Oppdraget – formål og problemstillinger

1.1 Bakgrunn

Arbeidsavklaringspenger (AAP) skal sikre inntekt til personer som har nedsatt arbeidsevne på grunn av sykdom eller skade, og som får bistand fra NAV til å avklare restarbeidsevne og skaffe eller beholde arbeid. En vedvarende bekymring rundt helserelaterte velferdsordninger er at uheldige *insentiveffekter* kan øke antallet som mottar ytelsene. Generøse ordninger kan redusere de økonomiske insentivene til å tilpasse seg «rundt» sine begrensninger, og i tillegg er det en risiko for at enkelte kan fremstille seg som sykere enn de er for å få tilgang til økonomisk støtte. Utover insentiveffektene er det også en bekymring for at medikalisering og sykdomsfokus kan øke den opplevde avstand til arbeidslivet for en gruppe som allerede er delvis utenfor. Denne problematikken ble også løftet frem av Sysselsettingsutvalget. Utvalget gjennomgikk en omfattende litteratur som viste at oppfølging, arbeidsrettede tiltak og aktivitet kan være positivt for unge både med og uten nedsatt arbeidsevne. Utvalget pekte videre på faren ved utenforskap, passivitet og medikalisering, både generelt og for AAP spesielt (NOU 2019: 7 og NOU 2021: 2 - Tiltak for økt sysselsetting).

På bakgrunn av slike bekymringer ble minstesatsen for AAP-mottagere under 25 år redusert med en tredjedel fra sitt tidligere nivå på 2G, samtidig som ung-ufør-tillegget på 0,44 G, for personer som fikk nedsatt arbeidsevne før fylte 26 år, ble fjernet. Begge endringene trådte i kraft fra 1.2.2020.

Formålet med endringene var å gjøre AAP mindre økonomisk attraktivt relativt til innteksbringende arbeid og andre ytelser og tiltak og å redusere sykdomsfokuset i denne perioden. Endringen i minstesatsen påvirket kun støttenivået til personer under 25 med lav forutgående inntekt (under 3G), men for disse kunne til gjengjeld reduksjonen være betydelig: En potensiell mottager under 25 år som kun kvalifiserte for minsteytelse, ville nå motta 1,33 G i stedet for 2G årlig. Dersom personen i tillegg ville kvalifisert for ung ufør-tillegget under de gamle reglene, var reduksjon på nær 50 % (1,33 G vs. 2,44 G). Siden regelendringene knyttet minstesatsen til løpende alder, var det videre variasjon i hvor *lenge* potensielle mottagere ville være påvirket av regelendringene: En 24-åring ville motta det reduserte beløpet frem til fylte 25, altså under ett år, mens en 22-åring hadde hatt redusert sats gjennom hele sin treårige mottaksperiode.

I tillegg vil fjerning av ung ufør-tillegget påvirke ytelsen i hele AAP-perioden for de personene tillegget ville vært aktuelt for, uavhengig av alder. Informasjon om hvem som mottar ung ufør-tillegget er ikke tilgjengelig i registerdata fra Statistisk sentralbyrå, som har vært vårt datagrunnlag. NAV har derimot opplysninger om dette i sine registre. Ved utgangen av 2016 mottok hhv. 12 og 9 prosent av AAP-mottakere i aldersgruppene 20-24 og 25-29 år ung ufør-tillegget (NOU 2019: 7). Fra Arbeids- og velferdsdirektoratet har vi mottatt tall for antall mottakere av AAP som var registrert som «ung ufør» innen 12 måneder. Denne andelen har vært sterkt fallende over tid og var for nye mottakere av AAP i 2019 under 25 år bare 2 prosent. For eldre mottakere har den tilsvarende andelen vært fallende fra et enda lavere nivå, og var for nye mottakere i 2019 bare 0,3 prosent. Det har også vært tilfellet at «om lag halvparten av de som får innvilget ung ufør-tillegget [får] dette utbetalt først *etter* at perioden på arbeidsavklaringspenger er over, som følge av at retten til tillegget først er blitt avklart underveis i perioden (Prop. 10 L 2019-2020).» Imidlertid var det bare 0,7 prosent av alle nye AAP-mottakere i 2019 som innen 3 år etter at de startet å motta AAP var registrert som ung ufør. Denne andelen er naturlig nok fallende med alder, men til og med blant dem som begynte å motta AAP i alderen 18-24 år var det bare 2,6 prosent som var registrert som ung ufør innen 3 år.

I hovedanalysen i kapittel 5, som sammenlikner personer over og under 25 år, vil effekten av avviklingen av ung ufør-tillegget delvis inngå i den estimerte reformeffekten, siden det var høyere andel under enn over 25 år som mottok ung ufør-tillegget. Vi går ikke nærmere inn på den direkte effekten av fjerningen av ung ufør-tillegget på innstrømmingen til AAP, ettersom både det at det var få som mottok tillegget og det at mange først mottok tillegget først etter AAP-perioden tilsier

at fjerningen skal ha hatt lite å si for innstrømmingen. På tross av det lille omfanget av mottakere, kan ung ufør-tillegget likevel ha spilt en rolle for forløpet for faktiske AAP-mottakere. Vi kommenterer dette i kapittel 6.

Intensjonen bak endringen var at potensielle AAP-mottagere ikke skulle ha økonomiske insentiver til å søke seg til AAP framfor egnet arbeid eller utdanning eller andre mer arbeidsrettede og mindre helseorienterte ordninger, som Kvalifiseringsprogrammet. En tilsiktet virkning av regelendringene vil dermed være redusert innstrømming til AAP. Imidlertid kan ikke dette utfallet bedømmes i isolasjon – det må tas hensyn til om det forekommer andre tilsiktede virkninger, som økt bruk av kvalifiseringsprogrammet eller økt overgang til arbeid eller utdanning, og utilsiktede virkninger, som forverret helse, redusert jobbfastholdelse eller økt bruk av supplerende sosialhjelp.

1.2 Problemstillinger

Målsetningen med prosjektet er å identifisere troverdige kausale effekter av regelendringen. Teoretisk sett kan disse være av ulik slag (figur 1), og en oversikt er verdt å ha i bakhodet, selv om flere av dem trolig vil være av mindre betydning i praksis. Hovedproblemstillingene i prosjektet har vært:

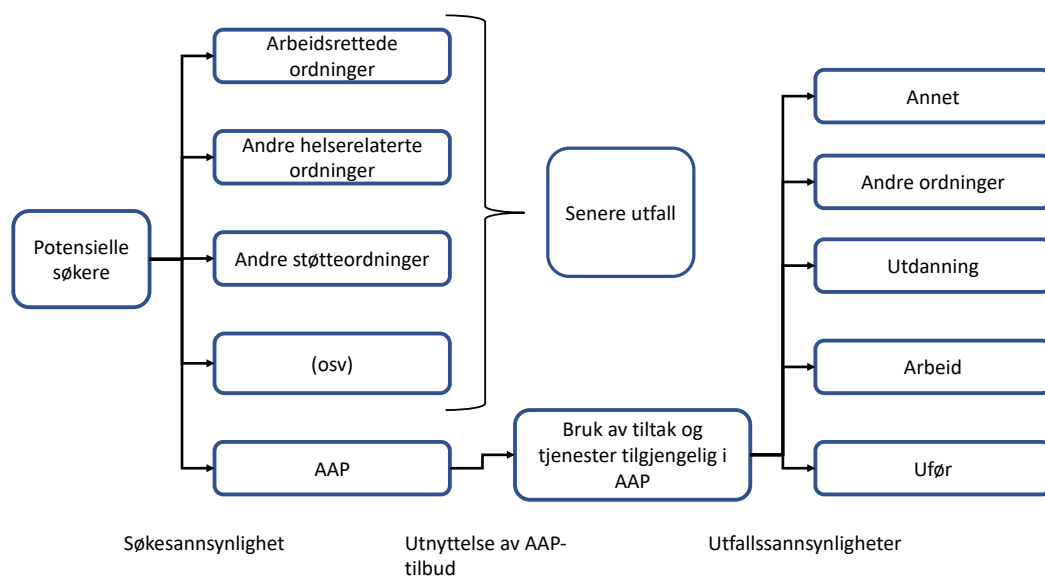
1. *Tidspunkt for søknader rundt reforminnføringen.* Forekom det økt pågang av søkere til AAP og ung-ufør ordningen i forkant av regelendringene?
2. *Innstrømming til AAP.* Har regelendringene påvirket innstrømmingen til AAP blant personer under 25 år?
3. *Innstrømming til andre ordninger.* Har regelendringene påvirket innstrømmingen til arbeid, utdanning eller andre ordninger?
4. *Utfall for mottakere av AAP.* Har regelendringen påvirket mottak av supplerende sosialhjelp eller overgang til arbeid, utdanning eller andre ordninger?

Problemstilling 1, om søkeatferd rundt tidspunktet for reforminnføringen, kan ses på som en forstudie til de etterkommende problemstillingene. Hvis mange AAP-søkere fikk fremskyndet søknaden sin for å komme i forkant av regelendringene, ville dette systematisk endret sammensetningen av søkere på hver side av reformtidspunktet ved å flytte søkere som var mer strategiske, fremadskuende eller informerte til perioden før innføringen. Dette ville i så fall ha stor betydning for tolkningen av resultater.

Hovedproblemstillingen i prosjektet gjaldt innstrømmingen til AAP (problemstilling nr 2): Påvirket innstrømmingen innstrømmingen? Analysestrategien i prosjektet ble bygget rundt dette spørsmålet.

Problemstilling 3 er nært beslektet med problemstilling 2: Dersom innstrømmingen endret seg, ser vi tegn til at denne motsvares av endringer i innstrømmingen til andre ytelser eller ordninger?

Problemstilling 4 dreier seg om utfall for mottakere av AAP. Regelendringen reduserte utbetalingene til nye AAP-mottagere under 25 år som kun kvalifiserte for minsteytelsen. Dette kan ha ført til en økning i supplerende sosialhjelp for å kompensere for det reduserte støttenivået i AAP-ordningen. Endringene kan også ha påvirket utstrømmingen fra AAP. Det at AAP-ordningen ble mindre økonomisk gunstig for dem under 25 år kan ha økt insentivene for aktivitet ved å gjøre utdanning og arbeid relativt mer attraktivt. Samtidig er det en risiko for at en mer presset økonomisk livssituasjon kan være belastende og virke andre veien.



Figur 1 - Kausale effekter: Regelendringen kan ha endret hvilke støtteordninger potensielle AAP-mottagere velger å søke, hvordan de benytter AAP-tiltak, og hvilke utfall de ender opp med i etterkant av AAP. I tillegg vil de som påvirkes til å ikke søke AAP som et resultat av regelendringen kunne ha andre utfall på sikt (ikke vist).

1.3 Koronaepidemien

Analysen er spesielt krevende i det at regelendringene ble innført rett i forkant av koronaepidemien. Pandemien og tiltakene som ble innført, hadde store konsekvenser for en rekke næringer og befolkningsgrupper. Disse konsekvensene kan ha variert systematisk med alder, livssituasjon og geografi. Dette ville potensielt skape problemer for statistiske analyser, ettersom disse analysene identifiserer kausale effekter ved å undersøke om data fra etter regelendringene avviker fra de mønstrene som er observert i tidligere årganger med data. Når en «svart svane»-hendelse som pandemien slår inn samtidig som regelendringene, vil dette potensielt påvirke ulike alders- og inntektsgrupper ulikt med hensyn til jobbmuligheter, utdanningsmuligheter, og psykisk helse. Både prosjektgruppen og oppdragsgiver var på forhånd kjent med at dette kunne gjøre analysene mer usikre.

I tillegg ble NAV direkte påvirket gjennom å måtte håndtere en voldsom mengde permitterte og ledige, noe som kan ha påvirket andre deler av NAVs arbeid. Spesielt kan det ha påvirket kvalifiseringsprogrammet, som er ment å innebære tett oppfølging og som Sysselsettingsutvalget trakk fram som et særlig relevant alternativ for unge AAP-mottakere. Med andre ord: Noe av potensialet ved endringene kan ha blitt svekket av at pandemien reduserte mulighetene til å tett følge opp personer i risikogruppen på mer arbeidsrettede måter.

1.4 Data

De kvantitative analysene er basert på administrative registerdata på utlån fra SSB. Disse gir informasjon om demografi (kjønn, alder), sysselsetting, inntekt og mottak av trygd og andre ytelser og dekker perioden 2010-2022. I første del av prosjektet hadde vi tilgang til data for årene 2010-2018. Til analysene av reformeffektene i andre del av prosjektet fikk vi data oppdatert til og med 2022. Analysepopulasjonen besto av alle registrerte personer i alderen 18-30 i disse årene. Tabell 1 gir en oversikt over utvalgte kjennetegn for denne populasjonen i 2018. Det var nesten 1 prosent

som fikk innvilget AAP for første gang i løpet av året, som er vårt hovedutfall. 3,7 prosent mottok AAP i løpet av året. Færre mottok permanent uføretrygd. 60 og 27 prosent hadde henholdsvis videregående og høyere utdanning som høyeste fullførte utdanning, mens 47 prosent var registrert i utdanning. Gjennomsnittlig yrkesinntekt var 244 561 kroner, 4,7 prosent mottok sosialhjelp, og 0,3 prosent deltok på kvalifiseringsprogrammet.

| | Gjennom- snitt | Standard- avvik | min | maks |
|---|-------------------|--------------------|-------|-------|
| Kvinne, andel | 0,490 | | 0 | 1 |
| Alder | 24,3 | 3.713 | 18 | 30 |
| AAP første gang, andel | 0,007 | | 0 | 1 |
| AAP i løpet av året, andel | 0,037 | | 0 | 1 |
| Uføretrygd første gang, andel | 0,004 | | 0 | 1 |
| Uføretrygd i løpet av året, andel | 0,020 | | 0 | 1 |
| Fullført vgs. høyeste utd., andel | 0,598 | | 0 | 1 |
| Fullført høyere utd. , andel | 0,269 | | 0 | 1 |
| Registrert i utdanning, andel | 0,469 | | 0 | 1 |
| Yrkesinntekt, 1000 kr (2023-kroner) | 244,6 | 267.2 | -5670 | 16658 |
| Sosialhjelp i løpet av året, andel | 0,047 | | 0 | 1 |
| Kvalifiseringsprogrammet i løpet av året, andel | 0,003 | | 0 | 1 |
| N | 1 355 430 | | | |

Tabell 1 Deskriptiv statistikk 2018, analyseutvalg

2 Hva endret seg i praksis: Regelendringene «sett fra innsiden»

2.1 kvalitative intervjuer med veiledere i NAV

I forkant av registerdataanalysen ønsket vi å få mer kunnskap om prosessene hvor unge uten arbeid søker om og eventuelt får AAP, og hvordan reglene kan påvirke dem når det gjelder å søke AAP, søke andre ytelser, og komme i arbeid eller utdanning. Eventuelle effekter av regelendringene fra 2020 må antas å virke gjennom endret atferd blant unge søkere og mottakere av AAP, og/eller endret praksis blant NAV-veiledere. Vi har derfor gjennomført gruppeintervjuer med veiledere ved tre NAV-kontorer, for å undersøke erfaringer med regelendringene og øke forståelsen av mekanismer som kan gjøre seg gjeldende.

De veilederne vi intervjuet jobber i ungdomsteam eller i oppfølgingsavdeling med ansvar for unge brukere. De tre NAV-kontorene er plassert i store kommuner. Ingen av de største byene er representert, men alle tre er blant landets 25 største kommuner. I utgangspunktet planla vi å intervjuve veiledere fra fire kontorer, men det viste seg vanskelig å rekruttere deltakere. Vi så imidlertid at det var stor grad av enighet blant de tre kontorene vi fikk rekruttert. Mot slutten av prosjektperioden gjennomførte vi et fokusgruppeintervju med veiledere fra to NAV-kontor, som begge representerte to av landets ti mest befolkede kommuner. Når det gjaldt de temaene som ble diskutert i intervjuene, sammenfalt meningene blant veilederne i fokusgruppeintervjuet i stor grad med det veilederne fra de mindre kommunene uttrykte.

Ved hvert intervju deltok mellom tre og syv personer. De fleste av dem oppga at de har oppfølgingsansvar for mellom 50 og 70 unge NAV-brukere. De fleste jobbet utelukkende med unge med nedsatt arbeidsevne, hvorav mange, men ikke alle, mottar AAP. Intervjuene ble gjennomført i mai og juni 2022.

Gjennom intervjuene har vi fått kunnskap om hvilke konsekvenser veilederne mener regelendringene har hatt for:

- tilstrømmingen av søkere til AAP
- oppfølging av brukere og bruk av supplerende sosialhjelp
- overgang til arbeid/utdanning
- overgang til uføretrygd

Oppsummeringen nedenfor er basert på de tre gruppeintervjuene som ble gjennomført tidlig i prosjektet, mens resultater fra fokusgruppeintervjuet inngår i drøftingskapittelet 7.

2.2 Flere regelendringer på kort tid

I årene før intervjuene hadde det vært flere regelendringer for AAP som innebærer innstramminger. Veilederne vi intervjuet så i stor grad på disse i sammenheng. Fra januar 2018 ble det innført betydelige innstramminger i varigheten av AAP. Maksimal stønadperiode ble redusert fra fire til tre år, og mulighetene til å bli innvilget forlengelse ble strammet inn. I tillegg ble det innført et karensår i ordningen. Karensåret innebærer at personer som har vært igjennom maksimal stønadperiode, må vente ett år før det eventuelt kan innvilges en ny periode med AAP. På samme måte som størrelsen på ytelsen kan også redusert maksimal varighet medføre at ordningen oppfattes som mindre attraktiv.

I løpet av koronaepidemien ble det i flere omganger vedtatt midlertidige endringer i regelverket for AAP. På grunn av det vanskelige arbeidsmarkedet under pandemien ble det gitt adgang til å forlenge stønadperioden på AAP. Karensregelen ble også opphevet. Da intervjuene ble gjennomført var

karensregelen midlertidig opphevet. Ved regelendringer som trådte i kraft i juli 2022, altså etter at intervjuene, ble opphevingen gjort permanent.

Både fordi det har vært flere endringer i reglene for AAP i løpet av de siste årene, og fordi situasjonen både når det gjelder mulighetene til å få jobb og NAVs ressurser til å følge opp brukerne, var spesiell under koronaepidemien, opplevde vi at NAV-veilederne ikke nødvendigvis skilte skarpt mellom effekter av de ulike regelendringene.

I 2021 offentliggjorde Fafo en rapport hvor de så på bruken av arbeidsavklaringspenger og kvalifiseringsprogram til unge ledige med nedsatt arbeidsevne under 30 år (Strand og Svalund 2021). De gjennomførte blant annet en survey til NAV-veiledere. De undersøkte hvordan veilederne så på nytten av regelendringer og hvorvidt de to ordningene bidrar til overgang til arbeid og aktivitet for unge ledige med nedsatt arbeidsevne, samt tiltaksbruk og oppfølging. Fafo fant at reduksjonen i nivået på stønaden – som er tema i vårt prosjekt – var den endringen lederne og veilederne i NAV var minst opptatt av, og at mange ikke hadde noe å si om dette temaet. Vi opplevde det samme i våre intervjuer.

2.3 Problematikk blant unge som veilederne har oppfølgingsansvar for

Veilederne vi intervjuet beskriver et sammensatt utfordringsbilde blant unge brukere de følger opp. Utfordringer som går igjen er depresjon, angst og rusproblematikk, i tillegg til fysiske helseutfordringer. Mange har droppet ut av videregående skole. Veilederne forteller at mange kommer til NAV når de fyller 18 år. Mange av brukerne bor hjemme hos foreldrene, og veilederne forteller at foreldre ofte er involvert i dialogen med NAV.

Mye psykisk helse og rus. Mange lever et isolert liv, både før og etter korona. Har blitt verre med korona. De færreste har klart å fullført videregående skole, selv om noen har. En del ME, men ikke ellers så mye somatikk.

Mange av de vi følger opp har dårlig nettverk, mange av de vi følger opp har kanskje også foreldre som har vært i systemet, og har en dårlig familiesituasjon.

De mangler det som andre ungdommer har fått med seg hjemmefra, hva det vil si å klare seg selv, hva vil det vil si å betale skatt for eksempel. Det er noe de fleste lærer hjemme, men som disse ikke har fått. Det skaper en del utfordringer for dem og oss.

2.4 Tilstrømming til AAP: Veilederne har ikke merket noen endring

Ingen av veilederne som deltok i gruppeintervjuene hadde lagt merke til noen vesentlig endring i søkertilstrømmingen av unge til AAP.

Hypotetisk kunne man sett for seg en økning i søkertilstrømmingen i månedene før regelendringen trådte i kraft 1.2.2020, i håp om å få innvilget AAP etter de gamle vilkårene. Slik veilederne ser det er det imidlertid få mulige søkere i målgruppen som har den nødvendige innsikten til å spekulere i dette.

Veilederne beskriver generelt en gruppe med nokså lav byråkratisk kompetanse. De ga uttrykk for at unge brukere ofte har nokså liten bevissthet om hvilke ytelser som er mest gunstig økonomisk, i alle fall før de kommer inn på ordningene. Inntrykket er at unge som søker AAP, i liten grad har bestemte assosiasjoner eller forventninger til AAP. Dette gjelder både hvor økonomisk attraktiv ytelsen er, hva som kreves av medisinsk dokumentasjon for å få innvilget AAP, samt hvilke forventninger som stilles til mottakerne. Enkelte av veilederne kjente likevel til noen eksempler på unge søkere som ønsker AAP fordi de forventer at det vil være økonomisk attraktivt. I disse tilfellene kan det være snakk om et familiemedlem som mottar AAP, eller at det har gått rykter i et miljø.

De kommer ikke inn til NAV med den innfallsvinkelen. De fleste med AAP er skikkelig dårlige, jeg tror ikke de har kapasitet til å tenke sånn.

Veilederne har heller ikke merket noen nedgang i tilstrømming til AAP etter at regelendringen trådte i kraft. Slik veilederne fremstiller det er det vanlig at en søknad om AAP skjer etter konsultasjon med en veileder, og veilederne forteller at de i stor grad gir de samme rådene som tidligere. I sakstyper der rådet tidligere ville vært å søke AAP, er fremdeles rådet det samme etter regelendringen. En justering veilederne har gjort i sine råd gjelder muligheten for å søke supplerende sosialhjelp.

2.5 Økt bruk av supplerende sosialhjelp

Den tydeligste konsekvensen veilederne ser av den reduserte minsteytelsen for unge på AAP, er økt bruk av supplerende sosialhjelp. Strand og Svalund (2021) finner også en lignende tendens i sin undersøkelse, der veilederne forteller at dersom unge får for lite penger på AAP blir dette supplert med sosialhjelp.

Enkelte av veilederne vi intervjuet fremstilte det som at nærmere alle AAP-mottakere også får supplerende sosialhjelp. Det samlede inntrykket er at dette vil gjelde mange, men ikke alle. Om den unge bor hjemme og har begrenset med faste utgifter, vil man ikke nødvendigvis få innvilget supplerende sosialhjelp. For denne gruppa mener veilederne at det har blitt vanskeligere å flytte for seg selv etter regelendringen. De ser at mange blir boende hjemme lenger.

En stor konsekvens er at det har skapt en hel haug sosialhjelpsklienter. Har du hybel er du plutselig helt avhengig av supplerende. Det tenker jeg er uheldig.

Det skaper dobbeltarbeid for kolleger.

Det er også tungvint for brukerne, de må da forholde deg til AAP-regelverket og så sette deg inn i sosialhjelp. Det blir mye ytelsesfokus, fokus på penger. Det må på plass for at vi skal hjelpe ut i arbeid.

Veilederne fremhever at praksisen med supplerende sosialhjelp til AAP skaper dobbeltarbeid, både sett fra NAV-ansattes og sett fra brukeres side. Det blir dobbel saksbehandling av søknader, og parallelle systemer for oppfølging. Enkelte informanter pekte også på at utgifter skyves over på kommunen, i og med at sosialhjelp forvaltes på kommunenivå. For brukerne blir det også mer jobb med å søke om sosialhjelp på nytt hver gang for å få den videreført, samt større krav til at NAV skal ha innsyn i privatøkonomien.

Veilederne har inntrykk av at endringen gjør ting mer krevende for brukerne. Brukerne får ikke nødvendigvis dårligere råd, i og med at den reduserte ytelsen på AAP ofte kompenseres med supplerende sosialhjelp. Men overføringene blir delt opp i flere utbetalinger, som kommer på ulike tidspunkter. Mottakere som ikke bor hjemme hos foreldrene, mottar gjerne både AAP, sosialhjelp og bostøtte. Til sammen kan det være 4-5 utbetalinger i måneden. Flere mindre utbetalinger på ulike tidspunkter i løpet av en måned skaper mindre forutsigbarhet og oversiktighet for brukerne. Veilederne fremhever at dette skaper uro blant brukerne, og at økonomien overskygger andre problemer som de unge trenger å ta tak i. Mangel på orden i økonomien gjør det vanskeligere å komme videre med andre prosesser.

Det blir veldig mye økonomifokus. Tilleggssøknad, AAP, sosialhjelp. Hvis du er 18 år og syk, har nedsatt arbeidsevne, er det ikke så lett å manøvrere seg i dette landskapet.

Behandlerne ser at økonomien har mye å si for helse, de kommer ikke videre.

2.6 Overgang til arbeid/utdanning, eller overgang til uføretrygd?

I stor grad tenker veilederne på 2018- og 2020-endringene i regelverket for AAP i sammenheng, begge uttrykk for innstramminger i ordningen. Veilederne er kjent med intensjonene bak innstrammningene, men de aller fleste vi har snakket med er skeptiske til at innstrammningene virker etter hensikten.

Det samlede inntrykket fra våre intervjuer er at veilederne mener innstrammningene har uheldige konsekvenser, ved at de skaper uro, usikkerhet og uforutsigbarhet for brukerne. Økonomiske bekymringer kommer i veien for andre mål det jobbes mot, og det blir et press mot å få avklart status som ufør.

Min magefølelse er at innstrammningene ikke har hatt noen positiv innvirkning for hverken brukerne eller den jobben vi skal gjøre. Har gjort det mer krevende, kortere tid og lavere satser.

Det skaper feil fokus, vi skal fokusere på arbeid og aktivitet, men vi må svare på spørsmål om økonomi.

Det virker kanskje litt motsatt, at de som står langt fra arbeidslivet blir hengende lengre bak. Har ikke kapasitet til å håndtere økonomien og behandling, de gir litt opp.

Dette samsvarer godt med hva Strand og Svalund (2021) fant i sin undersøkelse. De finner en betydelig skepsis blant veiledere til at endringene vil fungere etter hensikten. Respondentene i surveyen ble bedt om å ta stilling til hvorvidt reduksjonen i minstestytelsen på AAP for mottakere under 25 år er nyttig for å få unge raskere i arbeid. 41 prosent svarte at de var helt uenig i påstanden, 24 prosent var delvis uenig og henholdsvis 23 og 12 prosent var delvis enig eller helt enig.

Veilederne vi intervjuet var mer bevisste på virkningene av regelverksendringene i 2018 som omhandlet varigheten av AAP. Den tydeligste konsekvensen de ser av disse endringene, er at det blir mer hastverk og uro i prosessene. Kortere stønadsperiode og mindre adgang til forlengelser gir mindre forutsigbarhet for brukerne. Veilederens inntrykk er at 2018-endringene har bidratt til fortgang i avklaringsprosesser, men ikke nødvendigvis i den intenderte retningen. De tror presset mot raskere avklaring har gitt fortgang i uføreprosesser, heller enn fortgang i prosesser som skal ta ungdommene nærmere arbeid og selvforsørgelse. Enkelte av veilederne tror også at 2020-endringene kan ha bidratt til raskere overgang til uføretrygd.

Foreldre sier at det er for lite på AAP, de får det travelt med å få personen over på uføre. De tenker de vil ha det bedre økonomisk med uføretrygd enn på AAP. Inngangen til AAP er kanskje lik [som tidligere], men utgangen til uføretrygd øker.

I motsetning til med AAP, mener veilederne at mange unge i målgruppa har en bevissthet om uføretrygd. Inntrykket er at en del flere av de unge som kommer i kontakt med NAV, har en bevissthet om at dette er mer økonomisk attraktivt enn andre ytelser.

Inntrykket fra intervjuene med veilederne er at kombinasjonen av kortere stønadsperiode og redusert ytelse har bidratt til et økt trykk fra brukere, og eventuelt foreldre, om å komme seg over på uføretrygd.

2.7 Betydningen av koronaepidemien

I gruppeintervjuene brukte vi også tid på å diskutere mulige konsekvenser av koronapandemien. Veilederne har imidlertid ikke inntrykk av at utbruddet av pandemien, eller ringvirkningene den første perioden, påvirket selve innstrømmingen til AAP på noen merkbar måte.

Veilederne forteller om en annerledes dialog og oppfølging under periodene med nedstenging og restriksjoner, samt om isolasjon og en økt psykisk belastning blant brukerne. Flere veiledere nevnte også at slike belastninger har gjort seg gjeldende i bredere lag av ungdomsbefolkningen, og at

konsekvensene kan gjøre seg gjeldende også på mellomlang og lang sikt, slik at det kan tenkes at flere unge havner i målgruppen for AAP fremover.

Videre var det en konsekvens av pandemien at det ble innført midlertidige forlengelser på AAP-ordningen. Slik veilederne ser det har dette hatt en viss dempende effekt vis-a-vis forkortelsen av stønadsperioden fra 2018.

2.8 Konklusjoner

Ut fra det NAV-veilederne formidlet forventet vi på forhånd følgende resultater av effektanalysen:

- *Tidspunkt for søknader rundt reforminnføringen*: Ingen merkbar virkning
- *Innstrømming til AAP*. Ingen merkbar virkning
- *Utfall for mottakere av AAP*. Økning i bruk av supplerende sosialhjelp. Økt/tidligere overgang til uføretrygd.

Når det gjelder innstrømming til andre ordninger istedenfor AAP, var ikke det tema i intervjuene.

3 Analysestrategi

3.1 Prosjektet

Utarbeidelsen av analysestrategi bygget på to viktige poenger: Behovet for en statistisk tilnærming som tillater oss å anslå kausale effekter på en troverdig måte (identifikasjonsstrategi), og nødvendigheten av å utforme analysemodell og definere variable og datautvalg før man jobber med post-reform-data (statistisk arbeidsprosess).

Prosjektet ble derfor gjennomført i to faser. I første fase ble data fra før regelendringen ble innført (pre-reformdata) brukt til å velge statistisk fremgangsmåte og vurdere forventet power (statistisk utsagnskraft). Etter at analytiske valg var avklart med oppdragsgiver, ble modellspesifikasjon og fremgangsmåte skrevet ned i en pre-analyseplan. Denne ble lastet opp og lagret med loggført lagringstidspunkt på en ekstern open-science-nettside (<https://doi.org/10.17605/OSF.IO/D4PS6>).

I andre fase ble post-reform-data gjort tilgjengelig for prosjektet, og de planlagte analysene ble gjennomført. Der analysene avviker fra pre-analyseplanen er dette omtalt direkte og begrunnet.

I disse analysene fokuserte vi på regelendringen om reduksjonen i minstenivået. For enkelhets skyld omtaler vi denne endringen som «reformen».

3.2 Fase 1: Pre-registrering

Valget om å spesifisere og registrere analyser før post-reform-data ble gjort tilgjengelig var en sentral del av prosjektplanen. Slik pre-registrering er fremdeles uvanlig i ikke-eksperimentell forskning, men har de siste årene blitt stadig vanligere i en rekke fag som et resultat av den fortsatt pågående «replikasjonskrisen», der pre-registrering er ett av flere grep som er foreslått for å gjøre forskningsfunns mer troverdige (Ritchie 2020). Begrunnelsen for en slik praksis i oppdragsforskning er den samme som i forskning forøvrig: At det øker troverdigheten til de funnene som rapporteres ved å gjøre det tydelig hvilke analytiske valg som ble tatt før det var mulig å vite hvordan disse valgene påvirket det endelige estimatet. Se Boks 1 for nærmere om dette.

I en oppdragsforskningskontekst vil insentivene være noe ulike fra forskningen ellers, men også ha stort overlapp. Mye oppdragsforskning blir omarbeidet til vitenskapelige publikasjoner, og i tillegg vil det ofte være en forhåpning eller et ønske fra oppdragsgiver om å finne (eller ikke finne, avhengig av kontekst) bestemte resultater, eller i det minste klare svar. For forskere kan det også ligge en stolthet i å klare å finne et svar og slik sett «lykkes» med oppdraget, noe som kan føre til at man prøver ulike strategier på samme datasett til man «får det til», eller at man justerer og «skjerper» analysen for å «tydeligere få frem» de teoretisk plausible funnene man tror man kan finne.

3.3 Pre-registrering i vårt prosjekt

Konkret valgte vi å dele analysearbeidet i to. I første fase brukte vi data frem til og med 2018, altså før regelendringen hadde funnet sted, for å velge en egnet analysemodell og spesifisere analysene. Først når disse analysene var ferdige og analytiske valg var avklart med oppdragsgiver og logget med tids-stempel på en ekstern tjener (<https://doi.org/10.17605/OSF.IO/D4PS6>), fikk vi tilgang på faktiske post-reform-utfallsdata.

Formålet med disse analysene av pre-reform-data var å prøve ut og sammenligne analysestrategier. Målgruppen (de under 25 med inntekt under 3G) opplevde ingen særlige AAP-relaterte reformer i perioden 2012-2018. Dersom vi «later som» det fant sted en reform i 2016 og analyserte konsekvensene av denne placebo-reformen på AAP-innstrømmingen i 2017 og 2018 burde en egnet metode vise nettopp dette: En mangel på effekt.

Boks 1: Bakgrunn for pre-registrering – replikasjonskrisen

Kort fortalt går replikasjonskrisen ut på at mange tidligere publiserte arbeider beskriver funn som ikke lar seg gjenfinne når andre forskere forsøker. En rekke tilsynelatende troverdige og statistisk signifikante resultater i psykologifaget, for eksempel, viste seg å krympe eller forsvinne når eksperimentene ble gjennomført på nytt med større antall deltagere og mer presist spesifiserte forskningsprotokoller (OPEN SCIENCE COLLABORATION 2015). Dette har skapt en økende bevissthet rundt problematiske sider ved både den tradisjonelle arbeidsprosessen til forskere og den tradisjonelle publikasjonsprosessen.

Tradisjonelt har forskere arbeidet ved å iterativt raffinere data, modell og hypoteser. Dette innebærer at forskeren tar en rekke store og små valg som bestemmer datautvalg, konstruksjon av utfallsvariable, og spesifisering av analyseoppsett, og ofte vil det være *flere rimelige valg* som kan tas. Slike «frihetsgrader» i forskningen har vist seg å påvirke forskningsresultater mer enn man tidligere var klar over. Når ulike forskergrupper gis samme datasett og samme forskningsspørsmål er det likefullt stor variasjon i anslagene som gruppene finner (Huntington-Klein et al. 2021; Silberzahn et al. 2018). Siden disse valgene tas under arbeidet med de faktiske analysedataene, er det videre en risiko for at disse valgene påvirkes av tilfeldige og systematiske mønstre som finnes i dette spesifikke datautvalget. Hadde forskerne fått et annet uavhengig datautvalg ville de dermed ha gjennomført en annen analyse enn den de faktisk gjorde. Med andre ord: analysen man gjennomfører blir påvirket av tilfeldigheter i data, mens statistiske tester legger til grunn at man ville gjennomført nøyaktig samme analyse uavhengig av hvordan data så ut. Denne forskjellen gjør at de statistiske testene som er gjennomført ikke lenger er gyldige. Et «statistisk signifikant» estimat betyr bare at vi sjelden ville fått så ekstreme funn dersom den sanne effekten var null og analysen vår ble gjennomført *helt identisk* på et stort antall uavhengige datasett. Dersom vi faktisk ville gjort *ulike* statistiske analyser på hvert datasett, er det rimelig å tro at vi ville funnet signifikante resultater langt oftere (Gelman 2007), særlig når det er i forskernes egeninteresse å finne statistisk signifikante funn som øker sjansene for å publisere og publisere godt.

De sterke insentivene for å publisere godt vil også kunne påvirke sannsynligheten for at et forskningsarbeide «overlever» hele veien frem til publisering. Forskere som forfølger en ide og ser at den ikke vil føre frem til en «god» publikasjon med signifikante og spennende funn, vil trolig ha høyere sannsynlighet for å legge arbeidet i skuffen. Selv om de fullfører arbeidet, blir det typisk vanskeligere å publisere godt, og høyere sannsynlighet for at forskeren gir opp og lar det ligge i en skuff. De arbeidene som til slutt blir publisert er dermed et selektert utvalg av den forskningen som er gjort. Å gjennomføre statistisk støyfulle studier er litt som å kaste en terning, og dersom bare de som kaster seksere rapporterer om dette i fagtidsskrift får vi et villedende inntrykk av hva terningkast typisk gir. Dette kan være en del av forklaringen på et snodig funn man ser i forskningslitteraturen: Statistiske testresultat har en unaturlig høy opphopning på «riktig side» av vanlige signifikansverdier (Brodeur et al. 2016; Brodeur, Cook, and Heyes 2020). Det er også konsistent med et vanlig mønster der mindre presise studier typisk rapporterer større effekter: Når studiet har lav presisjon er det bare studiene som finner veldig store effekter som er statistisk signifikante (Ioannidis, Stanley, and Doucouliagos 2017).

Teoretisk sett virker det også sannsynlig at disse problemene rundt arbeidspraksis og publikasjonsprosess forsterker hverandre: Miljøer som lykkes med å få mange gode publikasjoner gjennom mer eller mindre bevisst utnyttelse av forskernes frihetsgrader, vil fremstå som attraktive og vellykkede og bli nettopp de miljøene som tiltrekker seg og lærer opp nye ambisiøse og flinke forskere (McElreath and Smaldino 2015).

Pre-reformdata ble ytterligere brukt for å undersøke presisjonen vi kan forvente å få på anslag av faktiske effekter, ved at tallene ble justert slik at vi bakte inn «syntetiske» reformeffekter (av en kjent størrelse) i datamaterialet. Som en illustrasjon, ant at det er 1000 personer som potensielt kunne bli AAP-mottagere i en bestemt alders- og inntekts-celle, og at 10 av disse faktisk ble AAP-mottagere. Det vil si at én av hundre i denne cellen ble AAP-mottagere. I det syntetiske datasettet antok vi da at sannsynligheten for å bli AAP-mottager var 1 prosent i denne cellen, og ba datamaskinen trekke et nytt antall AAP-mottagere fra denne cellen med denne sannsynligheten. De syntetiske effektene ble lagt inn ved å justere sannsynlighetene vi brukte på cellene med inntekt under 3G og alder under 25 opp eller ned med den kausaleffekten vi ønsket å legge inn. Bruker vi for eksempel en trekningssannsynlighet på 2 prosent, betyr dette at den syntetiske reformen doblet sannsynligheten for at noen i denne cellen skulle innvilges AAP (fra 1 til 2 prosent).

Det fine med en slik metodikk er at vi vet hvordan vi har justert sannsynlighetene og kan regne ut hvor mye høyere eller lavere det forventede antallet AAP-mottagere var. Dette gir oss en fasit som vi kan sammenholde med anslagene fra statistiske modeller. Disse resultatene gir nyttig veiledning i valg av statistiske modell for hovedanalysen av faktiske reformeffekter – når vi analyserer de faktiske post-reform-dataene har vi ingen fasit og er nødt til å stole på den statistiske modellen som er valgt.

3.4 Vår analysestrategi

Boks 2 beskriver alternative analysestrategier i prosjektet. I hovedteksten videre beskriver vår valgte analysestrategi. «Difference-in-differences»-strategien (DiD) kan anvendes der vi har data på et sett med enheter (f.eks. kommuner) over tid, og der noen gjennomfører en reform før andre. Hvis det er *tilfeldig* hvilken rekkefølge kommunene innfører reformen kunne vi i prinsippet nøydt oss med å sammenligne gjennomsnittlig utfall i reform- og kontroll-kommuner etter at noen hadde innført reformen, men dette ville gitt et upresist anslag fordi kommuner også varierer mye av andre grunner. Kommuner varierer mye i utgangspunktet, og antallet behandlede kommuner kan være lite. DiD-strategien tar sikte på å rense ut slike stabile forskjeller ved å sammenligne *tidsutviklingen* i behandlede og kontrollkommuner, for dermed å få et mer presist anslag på behandlingseffekten.

Den sentrale antagelsen i DiD-strategien er antagelsen om en *felles trend* som både de behandlede og kontroll-enhetene ville fulgt dersom ingen av dem var blitt behandlet. Med denne antagelsen kan vi hente ut en tidsutvikling fra «kontrollenhetene» og anta at også de som ble behandlet ville fulgt denne tidsutviklingen dersom de ikke hadde blitt påvirket av behandlingen. Med andre ord, hadde ikke kommune X innført reformen, antar vi at de ville hatt samme utvikling som den vi ser i kommuner som ikke innførte noen reform.

Om en slik felles trend antagelse er troverdig eller ikke avhenger av den spesifikke konteksten. Noen kommuner vil alltid ha en bedre utvikling over tid enn andre, og dess mer data vi har dess mer finmaskete forskjeller i trender kan vi påvise. Men dersom kommunene ble tilfeldig innrullert i reformen, så vil forventet «ubehandlet tidstrend» likefullt være den samme for de som ble behandlet tidlig, sent eller aldri. I virkeligheten er det nok sjelden slik at f.eks. alle kommuner hadde samme sannsynlighet for å innføre en reform tidlig eller sent. I slike situasjoner vil DiD-strategien egentlig være en variant av den eldre strategien «kontroller for så mye som mulig og håp at det gjenværende er kausalt». Fra et slikt perspektiv er styrken til DiD at den automatisk kontrollerer for alle *persistente* forskjeller mellom enhetene, selv der vi ikke har data på disse. I tillegg kan en slik DiD analyse virke tillitsvekkende dersom vi ser at behandlings- og kontroll-enhetene faktisk fulgte samme utvikling frem til det tidspunktet der den ene ble innlemmet i behandlingsgruppen. Som en illustrasjon: Anta at forskere kunne vise at kvinner og menn hadde tilnærmet identisk lønns- og karriereutvikling frem til det tidspunktet de får barn, men at de deretter hadde svært ulike karrierer. Selv om det ikke er tilfeldig trukket hvem som skal få barn og når, så vil mange være enig i at dette viser at barn påvirker karrieren ulikt for menn og kvinner.

Boks 2: Vurderte, men forkastede strategier for kausal inferens

Den kausale effekten av en intervensjon defineres som forskjellen mellom det som faktisk skjedde etter intervensjonen og det som ellers ville skjedd (Rubin 1974). I vårt tilfelle er det *kontrafaktiske* det som ville inntruffet dersom regelendringen ikke hadde funnet sted. Det grunnleggende problemet i kausal inferens følger direkte fra dette: Bare den ene siden av denne sammenligningen kan observeres og måles. Vi ser hva som skjedde, men det som *ville* skjedd dersom tiltaket ikke hadde blitt gjennomført kan bare anslås. Troverdigheten til anslaget på den kausale effekten hviler derfor på troverdigheten til anslaget på dette kontrafaktiske utfallet.

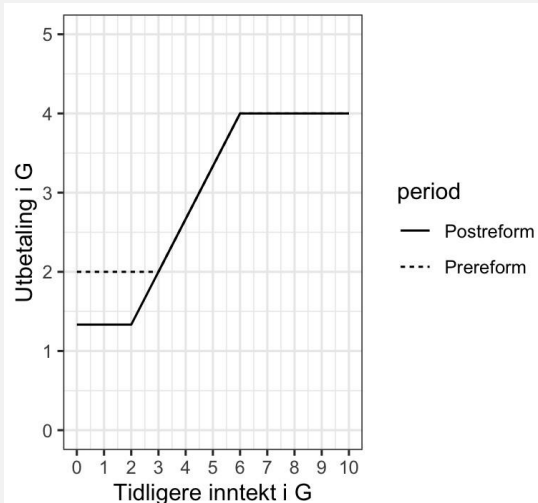
I de aller fleste tilfeller vil forskere anslå det kontrafaktiske utfallet ved å måle hvordan det går med en gruppe som ikke selv ble behandlet. Her er utfordringen å finne den rette sammenligningsgruppen. De som «behandles» er typisk *selektert* inn i behandlingen på en eller annen måte gjennom en kombinasjon av egne valg og interesser og «systemet» de står overfor (saksbehandlere/regler/tilgjengelighet). De behandlede skiller seg dermed allerede i utgangspunktet ut fra andre, og den relevante sammenligningsgruppen må følgelig skille seg ut på samme måte for at sammenligningen skal bli gyldig.

Den sikreste måten å oppnå dette på er gjennom et randomisert eksperiment. Her kan vi være sikre på at det *bare* er tilfeldigheter som skiller de med og uten behandling, og vi kan måle kausaleffekten ved å sammenligne de behandlede med kontrollgruppen.

Typisk vil forskere kun ha tilgang på ikke-eksperimentelle observasjonsdata. Tradisjonelt forsøkte man å gjøre behandlede og kontroller like ved å justere statistisk for *observerbare* forskjeller som alder, utdanning, familiesituasjon og lignende. Dette anses ikke lenger som en troverdig måte å anslå kausale effekter. Det er potensielt mange subtile, men betydningsfulle forskjeller som ikke kan observeres, og disse kan godt tenkes å ligge bak de «justerte» forskjellene vi sitter igjen med. I stedet leter man nå etter «eksperimentlignende» variasjon i data – det vil si forskjeller i behandling som med rimelighet kan anses som «så godt som tilfeldige» (Angrist and Pischke 2009; Morgan and Winship 2007).

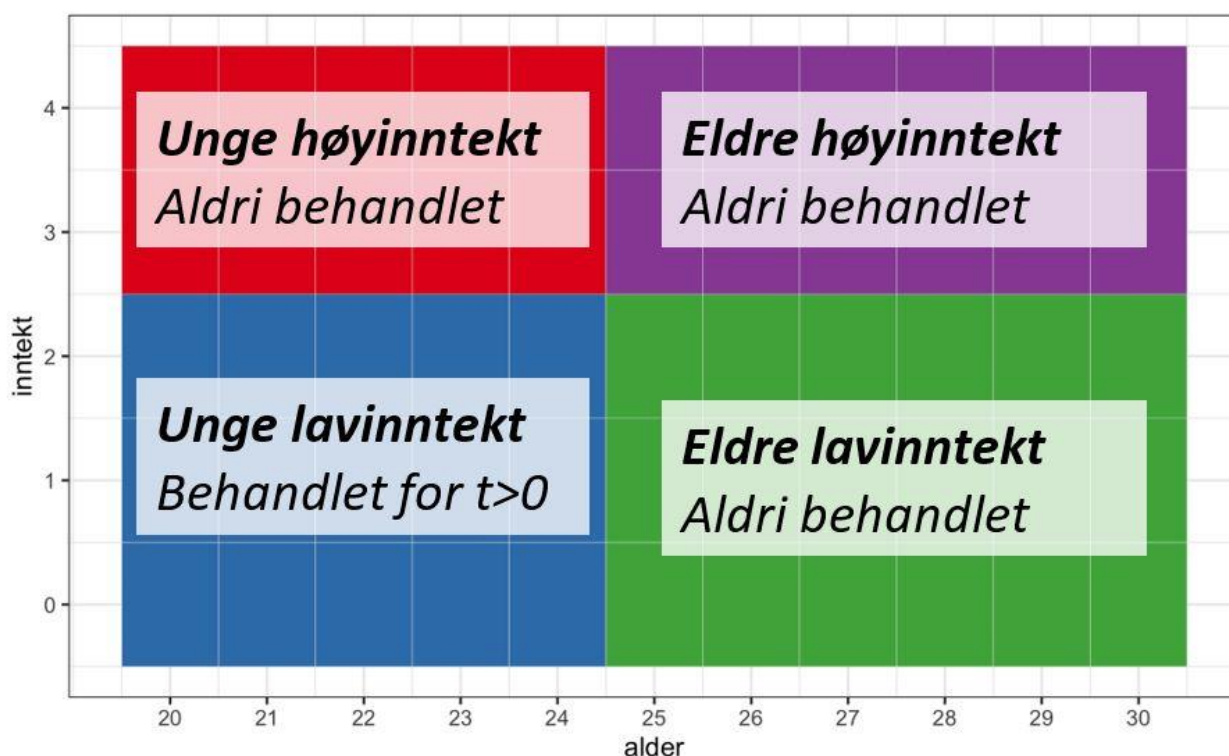
En hyppig anvendt strategi i slike situasjoner er «regression discontinuity». Her sammenlignes folk på hver sin side av en skarp grenseverdi der det skjer et plutselig og markant hopp i andelen som behandles. Gitt at bestemte antagelser er oppfylt, som at folk ikke har tilpasset seg i forkant ved å flytte til den «attraktive» siden av grenseverdien, kan vi bruke regresjonsmodeller på hver side av terskelen til å anslå hvordan forventede utfall skiller seg for de som er «inntil» grensen på hver side av denne. Hvis behandlingen har en effekt, vil forskjellen i behandling på hver side av terskelen vise seg som en forskjell i utfall.

I vårt tilfelle var reduksjonen i minstesatsen for AAP avgrenset av både en aldersgrense og en inntektsgrense, men ikke på en måte som egnet seg for denne metoden. Støttebeløpet var bare redusert frem til mottageren fylte 25, så en som fikk innvilget AAP dagen før de fylte 25, ville kun ha redusert støttebeløp i en enkelt dag. Og siden reformen endret minstesatsen, fikk reformen minimal betydning for de med en inntekt rett under 3G (se avviket mellom stiplet og heltrukket linje rundt 3G i Figur 2). Det var med andre ord ingen skarp diskontinuitet ved grensene.



Figur 2 Relasjon mellom inntektsgrunnlag og faktisk utbetaling før og etter regelendringen

I vår situasjon kunne vi brukt en DiD som sammenlignet de under 25 år med de over (heretter kalt alders-DiD), eller de med inntekt under 3G med de over (heretter kalt inntekts-DiD). Samtidig er det åpenbart langt fra tilfeldig hvilke inntekt-alder-celler som ble behandlet. Tegner vi et tverrsnitt av gruppene for et enkelt kalenderår er alle de behandlede cellene klumpet sammen i et hjørne av figuren (Figur 3). Spørsmålet blir da om det er plausibelt at f.eks. utviklingen for *eldre* potensielle mottagere sammenfaller med den utviklingen yngre ville opplevd uten reform – eller om det allerede i utgangspunktet var aldersspesifikke trender. Og analogt: Om utviklingen for dem med inntekt over 3G sammenfaller med den de under 3G ville opplevd uten reform, eller om det er inntektsspesifikke trender i AAP-tilstrømming.



Figur 3 Illustrasjon av Difference-in-Differences-tilnærmingen

Dersom det skulle være inntekts- eller aldersspesifikke trender, har vi en tredje mulig strategi som kan utnytte det at de behandlede skiller seg fra andre *både* langs alder og inntekt. Dette er en såkalt «difference-in-difference-in-differences» strategi (ofte omtalt som trippel-diff). Her er antagelsen at det kan være både alders-spesifikke trender og inntekts-spesifikke trender. Siden reformen ikke påvirket personer med inntekt over 3G, kan vi bruke disse til å anslå hvordan ulike aldersgrupper utviklet seg i forhold til hverandre. Siden reformen heller ikke påvirket personer over 25 år, kan vi bruke personer over 25 til å anslå hvordan ulike inntektsgrupper utviklet seg i forhold til hverandre. Deretter antar vi at de unge med lav inntekt – i fravær av reformen – ville opplevd en tidsutvikling som kombinerte både den alders- og inntektsspesifikke utviklingen vi har anslått.

Det positive med en trippel-diff modell er at den er mer generell enn en DiD, ved at den tillater aldersspesifikke trender. Ulempen er at den statistiske presisjonen ofte faller betydelig, siden vi nå sammenligner DiD-anslag som allerede i utgangspunktet er upresise. I tillegg er antagelsene bak modellen mer innfløkte, og kan være vanskelig å gjennomskue og vurdere med intuisjon: Vi antar nå at trendene for AAP-innvilgelse kan variere med alder, men at *forskjellene* mellom disse trendene ikke påvirkes av inntekt. Dess mer subtile slike antagelser blir, dess vanskeligere er det å vurdere dem ved å tenke på mulige moteksempler eller unntak. Som med vanlig DiD, kan det være nyttig å undersøke om antagelsen var oppfylt i forkant av reformen.

I dette prosjektet valgte vi å implementere DiD- og DiDiD-strategiene på celler definert etter kalenderår, alder, og inntekt foregående år. I hver celle teller vi antallet personer i registerdata som potensielt kunne motta AAP («at risk», det vil si ikke uføretrygdet eller tidligere mottager av AAP), samt antallet personer fra denne cellen som mottok AAP for første gang dette året. Personer som tidligere har mottatt AAP kan ha et annet forløp inn i ytelsen på nytt. Blant personer under 25 år vil det også være få som vil rekke å motta AAP på nytt. Vi ser derfor bare på førstegangsmottakere av AAP. Dette utgjør datamaterialet i hovedanalysen.

I vedlegget beskriver vi implementeringen mer detaljert. Når modellen er estimert bruker vi denne til å anslå (med usikkerhetsbånd) forventet antall nye AAP-tilfeller i hver celle. Disse anslagene blir deretter lagt til grunn for DiD og DiDiD-anslagene:

- **Alders-DiD** bruker kun celler med inntekt under 3G, siden de over ikke ble påvirket av reformen. Deretter slår vi sammen alle aldre under 25 og alle aldre over 25, og standardiserer seriene slik at forskjellen i pre-reform året blir null. Dette gir oss en event-studie, dvs. en analyse av utvikling akkurat rundt innføringen av reformen, som viser oss om de som var over og under 25 hadde samme tidsutvikling både før og etter pre-reform året.
- **Inntekts-DiD** bruker kun cellene med alder under 25, slår sammen alle inntektsnivå over 3G og alle under, og standardiserer seriene slik at forskjellen var null i pre-reform året.
- **Trippel-diff** bruker alle cellene. Hvis vi tenker fra inntekts-DiD siden er det som om vi estimerer inntekts-DiD separat for de med inntekt over og under 3G (begge standardisert slik at differansen er satt til null i pre-reform året), og deretter anslår forskjellen mellom disse. På denne måten får vi også her muligheten til en event-studieaktig analyse, ved at vi får sett om forskjellen mellom disse to DiD-anslagene var den samme over tid i pre-reform perioden og først avvek i etterkant

3.5 Sekundær-analyse: Endret søketidspunkt grunnet regelinnføring

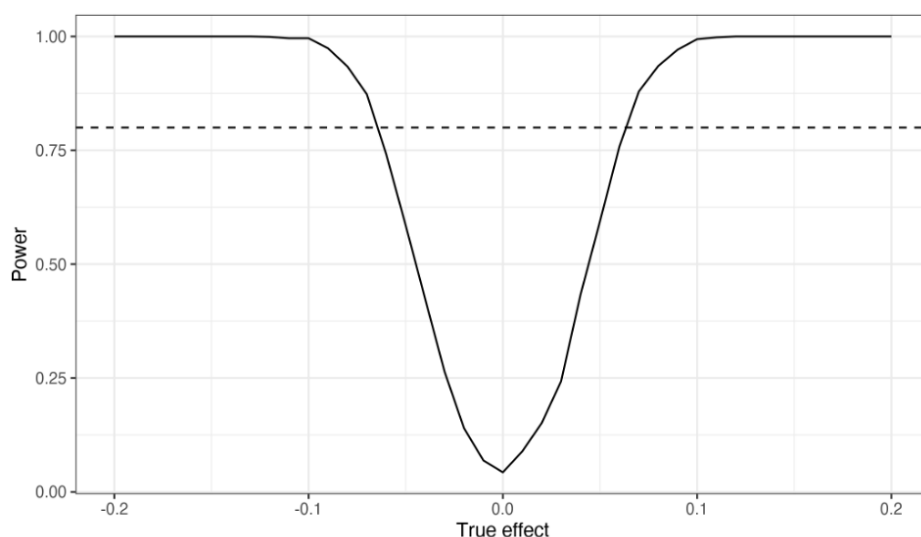
Oppsummering: Den planlagte analysen viste seg å være uegnet og ble erstattet, men vi finner ingen tegn til fremskynding av søknader innenfor et to måneders vindu på hver side av reformtidspunktet.

Et annet spørsmål vi ønsket å besvare var hvorvidt søkere til AAP fremskyndet søknaden sin rundt reforminnføringen for å falle inn under det gamle regelverket. Her skisserte vi en enkel analyse i pre-analyseplanen som i etterkant viste seg uegnet.

Analysen vi skisserte tok utgangspunkt i en enkel ide, der vi antok at søknadsvolumet vil variere over året, men at de *relative* forskjellene mellom måneder typisk vil være mer stabile. I så fall kan vi beregne et forholdstall (ratio) for antallet søknader i løpet av åtte uker i etterkant av 1 februar relativt til antallet søknader åtte ukene i forkant. Deretter kan vi undersøke om denne ratioen var annerledes i reformåret enn i tidligere år for de under 25. Mer spesifikt:

- Tell antallet søknader (“#”) i løpet av to måneder etter og forut for reformtidspunktet (bruk toppskrift “-“ for «før» og “+” for «etter»), og gjør dette både i reform året T året før (T-1)
- Beregn forholdstallet $\frac{\#_T^- / \#_T^+}{\#_{T-1}^- / \#_{T-1}^+}$
- Bruk den lognormale tilnærmingen for å ettergå om forholdstallet hadde endret seg signifikant med 5% som signifikansterskel

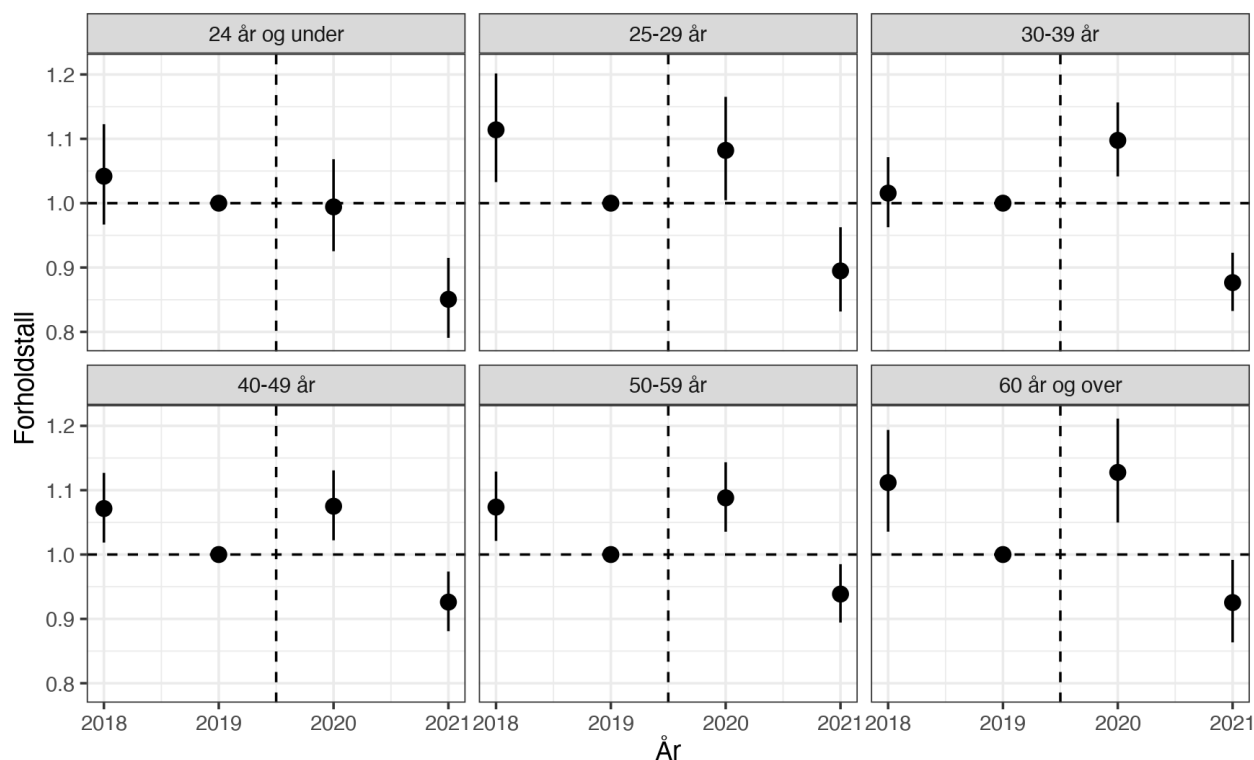
Basert på tall for antallet søknader totalt per år for den relevante aldersgruppen anslo vi statistisk styrke (power) som tilfredsstillende, jf. figur 4.



Figur 4 – Anslått statistisk power for odds-rate sammenligning.

Note: Figuren viser sannsynligheten for at en trekning av data skal gi signifikante resultater som en funksjon av hvor stor den underliggende reformeffekten på søknadsfremskynding antas å være.

Når analysen i etterkant fremstår lite egnet for å analysere problemstillingen skyldes dette at den underliggende antagelsen om at forholdstallet vil være stabilt over år ikke finner dekning i data. Sagt med andre ord: "Sesongmønstre" i søknadstilfanget er mindre stabile fra år til år enn vi forventet. Dette kan f.eks. skyldes at søknadstilfanget i ulike måneder er betydelig påvirket av enkeltfaktorer. Dessverre ble det ikke gjort placebo-analyser for denne analysen på pre-reform-data, men disse ville avslørt denne svikten: Bruker vi 2019 som referanseår og beregner ratioen i hvert år relativt til denne for ulike aldersgrupper ser vi til dels stor variasjon, jf. Figur 5. Ratioen i 2018 avvek signifikant fra ratioen i 2019 for fire av seks aldersgrupper. Dette er tall for søknader, men også AAP-innstrømmingen dette året var uvanlig lavt - trolig som et resultat av innskjerpinger i AAP-ordningen (Danielsen, Kalstø, and Grønlien 2023). Resultatene fra post-reform perioden er heller ikke tillitvekkende: Vi ser store og signifikante forskjeller fra 2019-ratioen i etterkant av den faktiske reformen for alle de uberørte aldersgruppene.

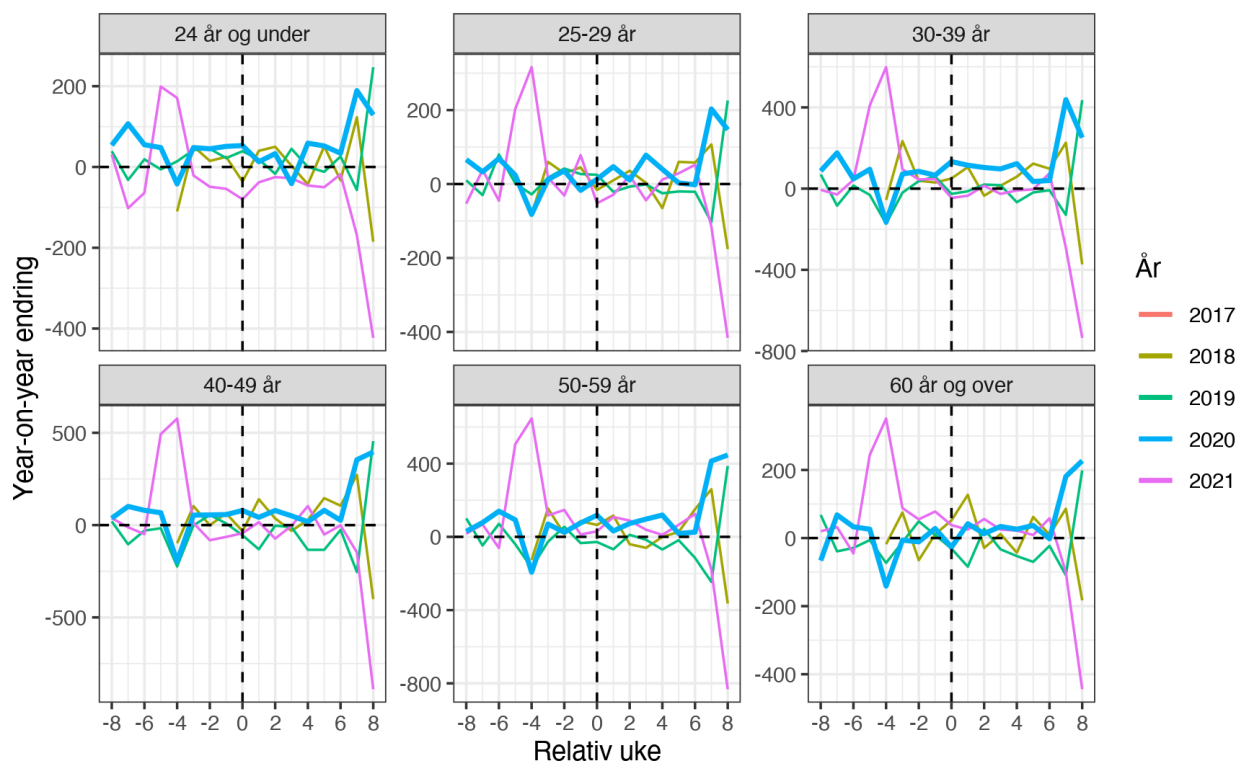


Figur 5 Resultater fra odds-ratioanalyser av søkertall i to ulike åtte-ukers perioder på tvers av år.

Positive tall vil si at et år – sammenlignet med 2019 for samme aldersgruppe - hadde relativt flere søknader i uker 5-12 relativt til uke 1-4 + uke 49-52 foregående år.

Et annet problem med analysen er at en reformeffekt uansett ville forventes å forrykke denne ratioen, siden antallet søkere i etterkant av reformen potensielt ville kunne bli redusert av at AAP ordningen ble mindre gunstig for en undergruppe av de potensielle søkerne.

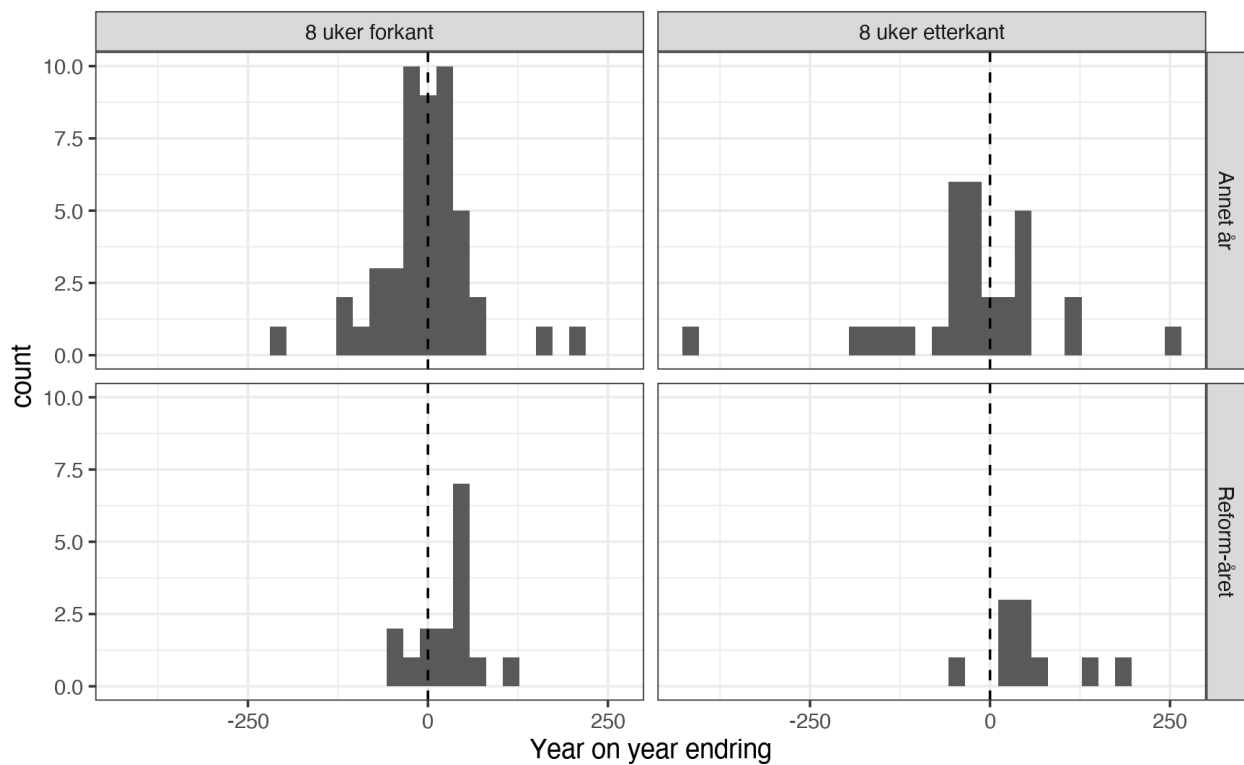
Spørsmålet vi lurer på er om det var en *fremskynding* av søknader, dvs. at antallet søknader var høyere enn det ellers ville vært i ukene før reformen ble innført og lavere enn det ellers ville vært i ukene etter at reformen ble innført. Ser vi på antallet søkere uke for uke i ulike år (Figur 6) er det lite som på øyemål tyder på at dette var tilfellet. Kurven for reformåret 2020 viser ikke tegn til et stort hopp blant de yngste i ukene før regelendringene.



Figur 6 Antallet søknader per uke sammenlignet med året før.

Note: Ukene dekker en periode på 17 uker sentrert rundt månedsskiftet januar/februar – for ulike år for ulike aldersgrupper. Reformåret 2020 er uthøvet. Uke 0 er årets femte uke og angir omtrentlig månedsskifte januar/februar, tidspunktet reformen ble innført i 2020.

For å gjøre sammenligningen enklere, beregner vi «år-til-år»-endring for hver kalenderuke, dvs. endring relativt til året før, og undersøker hvor store endringer det var vanlig å se i ukene på ulike sider av månedsskiftet januar/februar. Vi gjør dette både for reformåret og for de andre årene i data (Figur 7). Vi kan deretter sammenligne dette med størrelsen på endringene i reformåret. Analysen viser at år-til-år-veksten om noe lå litt høyere enn vanlig i ukene før reformen, som er konsistent med en forflytningshypotese. Samtidig viser analysen om noe høyere vekst enn normalt også i ukene *etter* reformen – altså det *motsatte* av hva vi skulle sett ved en fremskynding. Antallet søknader økte etter nedstengningen 12. mars 2020. Ettersom de to siste ukene i mars er inkludert i 8-ukersvinduet forklarer dette de to største ukentlige søknadsøkningene, men det er likevel en vekst også etter regelendringene. I tillegg vil reformen i seg selv potensielt bidra til å redusere søkevolumet – som ville trukket i samme retning og predikert lavere vekst i etterkant av reformen. Gitt det lave kunnskapsnivået om regler og støttebeløp som NAV-veilederne oppgir at potensielle AAP-søkere har, er vår konklusjon at det er liten grunn til å tro at reformen utløste betydelige forflytninger i når folk søkte. Det bør likefullt bemerkes at en eventuell nedgang i søknadsvolumet i månedene rett etter reformen kan ha blitt maskert av den kraftige økningen i AAP-tilfang man så på tvers av aldersgrupper i begynnelsen av korona-pandemien (Danielsen, Kalstø, and Grønlien 2023).



Figur 7 Fordelingen av år-til-år-endring i ukentlige AAP-søknader for de under 25 for ulike sett med uker

Note: Vi skiller mellom ukene 8 uker i forkant og etterkant av månedsskiftet januar/februar, og sammenligner ukenes år-for-år-endringer i årene med og uten reform.

4 Resultater fra fase 1: Analyser av pre-reform-data

Som beskrevet i kapittel 3 ble den kvantitative analysen gjennomført i to faser. I en første fase brukte vi kun data fra pre-reform perioden. Her var formålet å teste ut og sammenligne ulike statistiske modeller. I dette kapitlet gjennomgår vi resultatet av analysene vi gjorde på pre-reform-data og redegjør for valget av endelig analysemodell.

4.1 Analyse 1: Finner metodene null-effekt på placebo-reform?

Oppsummering: Tre ulike estimeringsstrategier ble hver implementert i fire ulike modeller som varierte i hvor skeptisk de var til store reform-effekter. Modellene ble brukt for å estimere reformeffekt på AAP-innstrømming i etterkant av en fiktiv placebo-reform som ikke fant sted, og for å ettergå antagelsene om at den faktiske utviklingen til unge voksne med svak inntektshistorikk fulgte trender estimert fra grupper som var upåvirket av reformen. Basert på resultatene ble det besluttet i samråd med oppdragsgiver å gå for difference-in-difference-in-differences modellen, som antar at unge med svak inntektshistorie har samme aldersspesifikke trend som unge med høyere inntekt og samme inntekts-spesifikke trend som de med lav inntekt og alder over 25.

Som beskrevet i kapittel 3 undersøkte vi tre strategier for kausal inferens:

- Alders-DiD: En difference-in-differences (DiD) som antar at utviklingen i antallet nye AAP saker i fravær av reformen ville vært den samme for de under 25 (med inntekt under 3G) som den faktisk var for de over 25 (med inntekt under 3G).
- Inntekts-DiD: En DiD som antar at utviklingen i antallet ny AAP-saker i fravær av reformen ville vært den samme for unge med lav inntekt (<3G) som det den faktisk var for unge med høy inntekt (>3G).
- Trippel-diff (DiDiD): En difference-in-difference-in-differences analyse som antar at unge med svak inntekstshistorie ville hatt de samme alders-spesifikke trendene som de med sterkere inntekstshistorie og de samme inntekts-spesifikke trendene som de over 25.

Hver av disse strategiene ble implementert i samme statistiske modell, som anslo sannsynligheten for at en tilfeldig trukket person i en gitt alder x kalenderår x inntekt celle ville få innvilget AAP i løpet av dette kalenderåret. Disse anslagene på cellenivå ble deretter brukt til å utarbeide DiD og DiDiD anslag på reformeffekter.

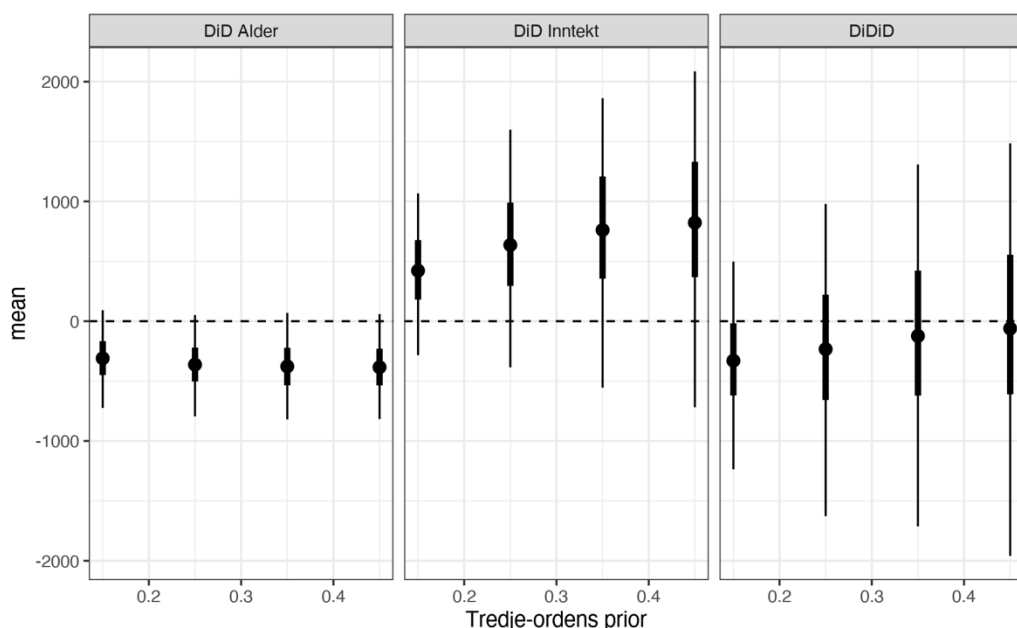
I analysene av pre-reform-data ble den underliggende modellen estimert fire ganger, der forskjellen lå i hva slags effektstørrelser modellen anser som plausible før den mates med data, også kalt priorfordeling. Den mest konservative av disse modellvariantene tar som utgangspunkt at det er usannsynlig med en reformeffekt som øker AAP-sannsynligheten til mer enn det dobbelte eller reduserer den til mer enn halvparten av hva den ville vært uten reformen. Vi anser også dette som relativt store potensielle effekter. I den andre enden av skalaen har vi en modellvariant som legger til grunn at også større endringer, opp til en tidobling/desimering, er innenfor hva som er plausibelt. Dersom data er informative og tilstrekkelig til å identifisere en presis kausaleffekt vil modellene komme til lignende konklusjoner, men dersom data er støyfulle og vi leter etter et svakt signal kan vi få høyere presisjon med de mer konservative modellene.

Vi har nå tre ulike strategier (2 x DiD + DiDiD) som alle implementeres innenfor hver av fire modellvarianter. Når vi estimerer modellene på pre-reform-data later vi som det kom en reform i 2016 som påvirket insentivene til unge under 25 med lav inntekt. Denne reformen fant ikke egentlig sted før 2020, så i pre-reform-data er dette en *placebo-reform* – en «fantom-reform» som analysestrategien bør finne en null-effekt av.

Resultatet av analysene er vist i Figur 8. Vi ser at estimerer basert på alders-DiD tyder på en nedgang i antallet AAP-tilfeller som et resultat av placebo-reformen, mens inntekts-DiD tyder på en vekst.

Trippel-diffen kombinerer de to andre og peker i retning nedgang – men betydelig mindre bastant enn alders-DiD gjør. Med alders-DiD anslår modellen at det er 94-96% sannsynlighet for at spøkelsesreformen førte til en nedgang, mens trippel-diffen gir en sannsynlighet et sted mellom 50 og 77 prosent for at det var en nedgang. Inntekts-DiD, på sin side, gir en sannsynlighet på 87-89% for at placeboreformen økte antallet AAP-tilfeller.

Figuren viser også at betydningen av de mer og mindre konservative modellene. For alders-DiD analysen er det nok data til at valg av modellvariant har lite å si. Derimot er det få personer med høy inntekt ved ung alder, som gjør at vi får lite presisjon når vi skal anslå sannsynligheten for at disse skal motta AAP. Anslagene fra inntekts-DiD og trippel-diffen som kombinerer de to DiD'ene blir derfor både mer usikre, og mer avhengig av hvor skeptiske antagelser vi legger til grunn. Sagt på en annen måte, vil de mer konservative modellene ta svake indikasjoner i data med en større klype salt.



Figur 8 – Anslag på kausal effekt av placeboreform i 2016

Note: Utfall: Total endring i antall nye AAP-tilfeller i 2017-2018). Anslag angis med tykk strek for 50% kredibilitetsintervall (KI) og tynn strek for 95% KI.

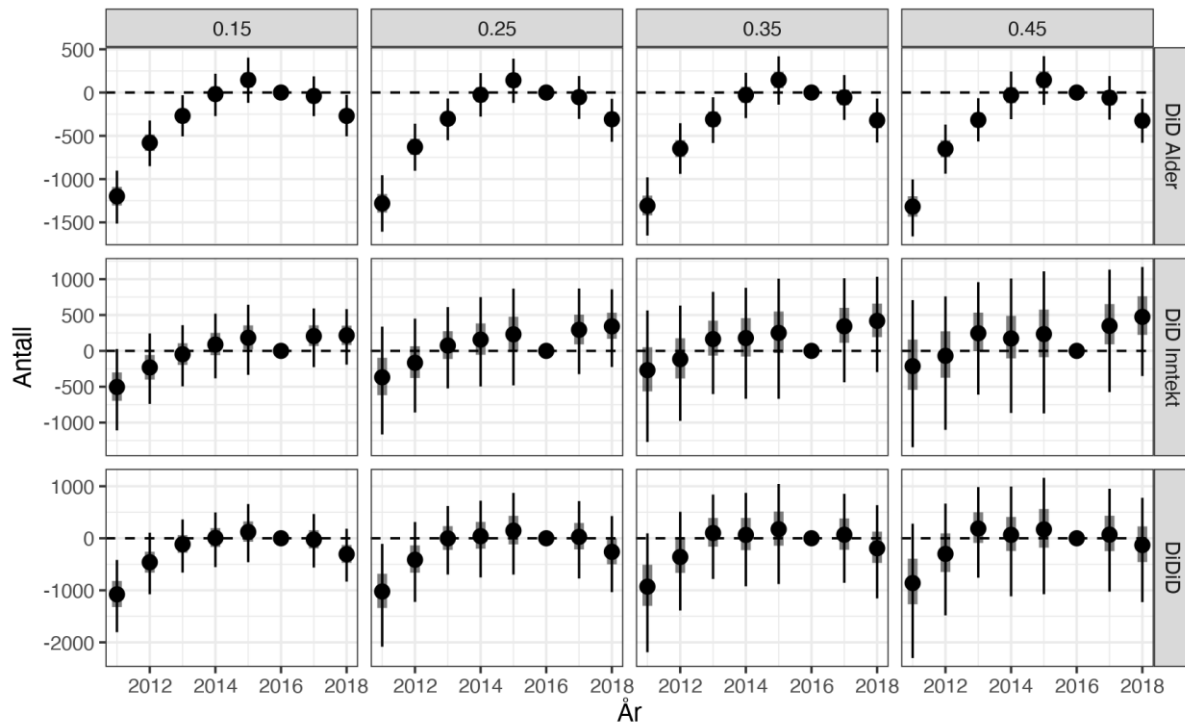
Ingen av de tre metodene kan utelukke at placeboreformen var uten effekt, men anslagene som bruker DiD-strategi peker i retning av at det mest sannsynlig var en effekt av den ikke-eksisterende reformen. Avhengig av prior anslår modellen at sannsynligheten for en negativ effekt på rundt 94-96% ved bruk av alders-DiD strategien, mens sannsynligheten for at effekten er positiv er 87-89% med inntekst-DiD strategien. DiDiD-strategien tilsier at effekten trolig er negativ (77% sannsynlighet) dersom den er liten (konservativ prior), mens sannsynligheten er omtrent 50-50 dersom vi har den mer liberale tredjeordens-prioren.

4.2 Analyse 2: Var det en felles trend i pre-reform perioden?

Som diskutert i kapittel 3 er DiD-analyser betinget på en antagelse om at den behandlede gruppen ville utviklet seg på samme måte som en sammenligningsbane dersom de selv ikke hadde mottatt noen behandling. Denne antagelsen kan ikke direkte ettergås i data – de behandlede enhetene ble behandlet, og vi kan ikke kjøre historien på nytt. Derimot kan vi sjekke om de to utviklingsbanene fulgte hverandre frem til det tidspunktet noen av dem ble (potensielt) påvirket av behandlingen. Dette undersøker vi gjennom såkalte «event-plots» som viser forskjellen mellom den faktiske

utviklingen i AAP-tilfeller over tid og sammenligningsbanen. Sammenligningsbanen er, som forklart tidligere, kun basert på utviklingen vi ser i grupper som ikke var omfattet av reformen.

I figur 9 viser vi event-plot for alle tre strategiene basert på modell-anslag med hver av de fire tredjeordens-prior verdiene. Event-plottet til alders-DiD analysen tyder på klare aldersrelaterte forskjeller i trend gjennom pre-perioden, og trendforskjellen er også tydelig i post-placeboreform perioden der vi ser et klart fall to år etter reformen (Figur 9). Event-plottet til inntekts-DiD har rent visuelt sett en antydning til trendforskjeller ved at anslagene blir systematisk høyere dataperioden gjennom – men samtidig er null dekket av 50% KI i alle pre-reform år. DiDiD-modellen ligger lavt de to første årene, men er stabilt rundt null i de gjenværende årene (noe lavt siste året i modellen med den mest konservative prior).



Figur 9 – Event-plott for placebo-reform i 2016.

Note: Event-plot for hver strategi er gjentatt for analyser basert på ulike tredjeordens-prior fra 0.15 (mest konservativ) til 0.45. Anslag angis med tykk strek for 50% kredibilitetsintervall (KI) og tynn strek for 95% KI.

4.3 Konklusjon etter analyser av pre-reform-data

Basert på analysene av pre-reformdata ble det besluttet å bruke DiDiD-analysene for post-reformdata. DiD-analysene som brukte enten alders- eller inntekstvariasjon alene utviste brudd på fellestrend antagelsen både i post-reform periode (der den faktisk trengs for å anslå en kausaleffekt) og i pre-reform perioden, noe som særlig var tydelig for alders-DiD. Valget mellom inntekts-DiD og DiDiD er mindre klart, men de klareste indikasjonene på avvik vi ser i event-plottet for DiDiD er de to første årene da AAP-ordningen først ble innført. Det er mange grunner til at dynamikken de første årene etter innføringen kan ha skilt seg fra senere. I de senere analysene har vi derfor konsentrert oss om tidsperioden fra og med 2013.

5 Resultater fra fase 2: Analyser av post-reformdata

I dette kapittelet gjennomgår vi resultatet av de pre-registrerte analysene av regelendringenes effekt på innstrømming til AAP. Som oppsummert i kapittel 2 er saksbehandlernes inntrykk at endringene i liten grad hadde noen betydning på innstrømmingen til AAP – kunnskapen om støttebeløp er lav blant potensielle søkere, saksbehandlere har ikke opplevd noen merkbar endring i sammensetningen av søkere, og deres vurdering var at det mest sannsynlige ville være en økt bruk av supplerende sosialhjelp. Vi fokuserer derfor på resultatene fra den konservative modellen som legger til grunn at en eventuell kausal effekt av reformen ligger et sted mellom en dobling og halvering av risikoen for AAP.

5.1 Utfall: innstrømming til AAP

Oppsummering: Som forventet var de statistiske analysene beheftet med stor usikkerhet, og statistiske usikkerhetsbånd kunne ikke utelukke at reformen var uten effekt. Punkttestimatene tilsier en reformeffekt som reduserte sannsynligheten for AAP-opptak med ~5-6 prosent. Usikkerheten rundt konklusjonen er forsterket av covid-pandemien, som kan ha forstyrret analysene i den grad pandemien fikk utslag som varierte med alder, arbeidsmarkeds- og utdanningstilknytning.

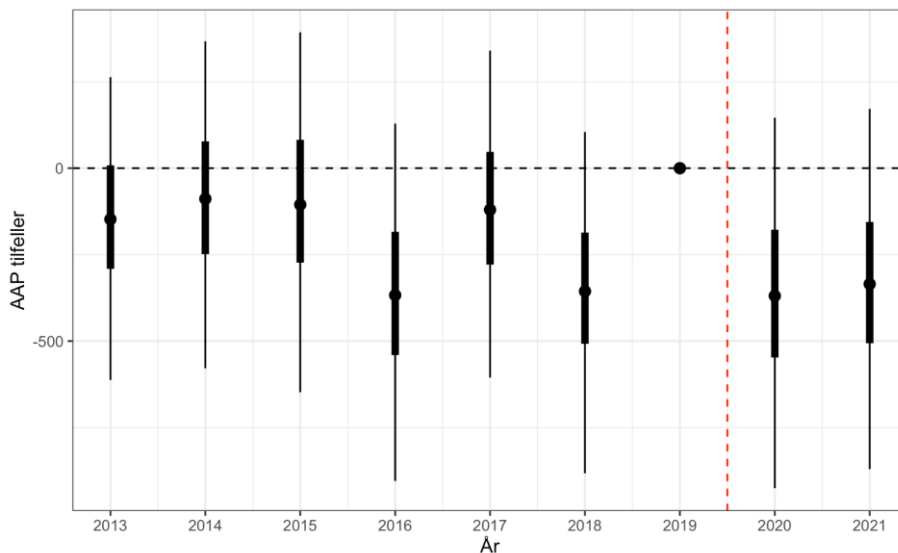
Reformen reduserte støtten til unge under 25 med svak inntektshistorie på AAP. Vi sammenligner den faktiske utviklingen i antallet AAP-tilfeller i denne gruppen med et anslag på hvordan utviklingen deres ville vært i fravær av reformen («sammenligningsbanen»). Forskjellen mellom den faktiske utviklingen og denne sammenligningsbanen gir et anslag på hvordan reformen kausalt endret innstrømmingen til AAP.

Sammenligningsbanen er bygget opp av to typer informasjon. Siden yngre kan ha en annen tidsutvikling enn eldre henter vi ut *alders-spesifikke trender* fra de som var upåvirket av reformen på grunn av sin inntekt (over 3G). Siden de med sterk inntektshistorie kan ha en annen tidsutvikling enn de med svak inntektshistorie henter vi også ut *inntekts-spesifikke trender* fra de som var upåvirket av reformen på grunn av sin alder (over 25).

Sammenligningsbanen for unge med lav inntektshistorie kan nå bygges opp som summen av disse to relative trendene. Med andre ord antar vi at unge med svak inntektshistorie har den samme alders-spesifikke utviklingen som unge med høyt inntektsgrunnlag, og den samme inntekts-spesifikke utviklingen som de over 25.

I Figur 10 viser vi forskjellen mellom den faktiske utviklingen og sammenligningsbanen. De to kurvene er begge satt til null i 2019, året før reformen fant sted. Vi omtaler derfor 2019 som referanseår. Vi ser at forskjellen er negativ i årene 2013-2018, det vil si at faktisk antall AAP-tilfeller er lavere enn det sammenligningsbanen angir. Forskjellen er av varierende størrelse mellom år, men 95 % usikkerhetsbånd, vist av de horisontale linjene for hvert årlige punkttestimat, kan ikke utelukke statistisk at forskjellen er 0. I de to årene etter reformen ser vi at den faktiske utviklingen også har ligget under sammenligningsbanen, men at 95% usikkerhetsbånd heller ikke her kan utelukke null. Gjennomsnittsanslaget fra modellen er en nedgang på 369 tilfeller i 2020 (95% bånd: -926 til +146) og 334 i 2021 (95% bånd: -870 til +172). Tar vi utgangspunkt i at det er rundt 3500 nye AAP tilfeller årlig fra gruppen som helhet er dette en nedgang på ~9-10%.

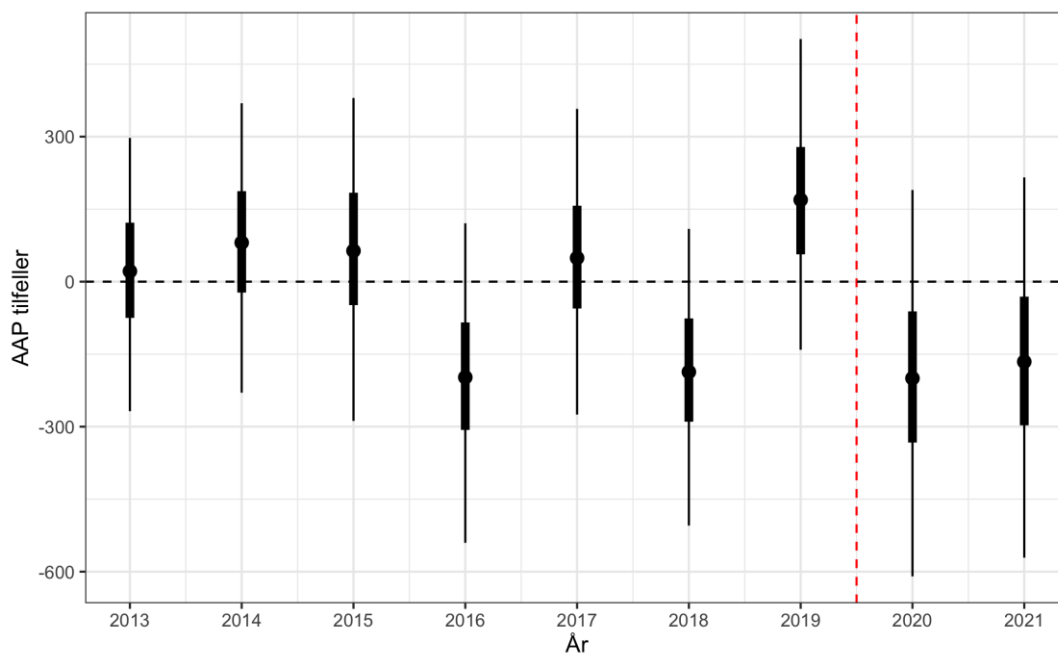
Vi ser at estimatene for 2016 og 2018 også er negative og på om lag samme størrelse som de estimerte reformeffektene i 2020 og 2021. Det gir en tydelig indikasjon på at endringer av den størrelsesordenen vi har estimert forekommer også i år uten reformer eller regelendringer som forventes å påvirke unge potensielle AAP-mottakere med lav tidligere inntekt. Intuitivt forklarer dette hvorfor de estimerte reformeffektene ikke er statistisk signifikante – vi ser endringer av denne størrelsen også ved andre tidspunkt, det er ikke systematisk slik at endringene oppstår der vi skulle forvente hvis det var reformen som var årsaken.



Figur 10 – Event-studie plott med referanseår 2019.

Note: Figuren viser forskjellen mellom den faktiske utviklingen og en sammenligningsbane fra en difference-in-difference-in-differences i nye AAP-tilfeller blant unge under 25 med inntekt under 3G. Anslag angis med tykk strek for 50% kredibilitetsintervall (KI) og tynn strek for 95% KI.

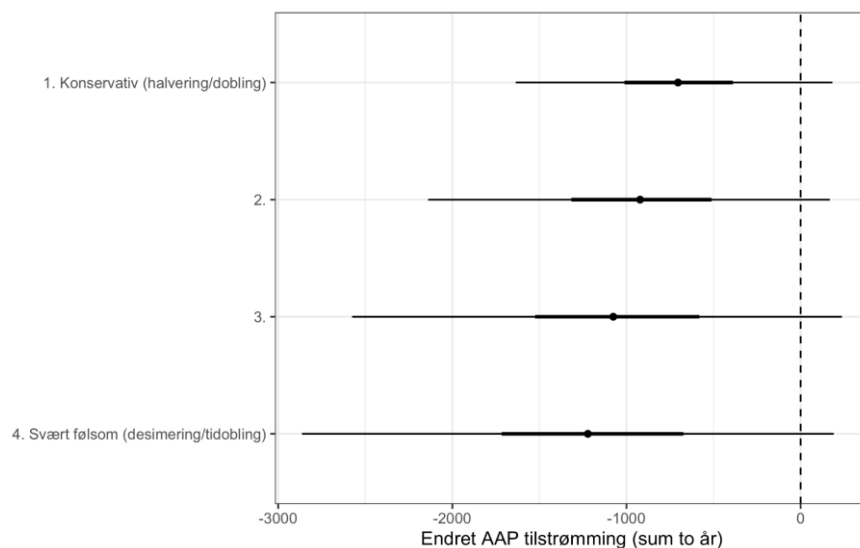
En potensiell bekymring ut fra Figur 10 er at valget av referanseår kan drive resultatene i for stor grad. Punktestimaterne i alle år 2013 til og med 2018 ligger lavere enn referanseåret 2019, om enn ikke i like stor grad som punktene etter. 2019 var også året da AAP-tilgangen begynte å øke igjen etter mange års nedgang. Vi definerer derfor et referansenivå slik at *gjennomsnittet* i pre-reform årene 2013 til og med 2019 er null (Figur 11), i motsetning til å se endringer i forhold til bare 2019. Med dette referansenivået blir den anslåtte reformeffekten nå -200 i 2020 (95% bånd: -610 til +190) og -166 i 2021 (95% bånd fra -571 til +216). Tar vi utgangspunkt i at det er rundt 3500 nye AAP tilfeller årlig fra gruppen som helhet er dette en nedgang på ~5-6%. Heller ikke disse estimatene er statistisk signifikant forskjellig fra 0.



Figur 11 – Event-studie plott med referansenivå satt til snitt av avvik 2013-2019.

Note: Figuren viser forskjellen mellom den faktiske utviklingen og en sammenligningsbane fra en difference-in-difference-in-differences i nye AAP-tilfeller blant unge under 25 med inntekt under 3G. Anslag angis med tykk strek for 50% kredibilitetsintervall (KI) og tynn strek for 95% KI.

Resultatene så langt har brukt den konservative modellen som legger til grunn at en plausibel reformeffekt vil ligge et sted mellom en halvering og dobling av sannsynligheten for å få AAP. Modellene som legger til grunn at reformen kan ha betydelig større effekter enn dette gir betydelig mer usikre anslag som i snitt peker mot større reformeffekter, se figur 12. Vi anser disse som lite troverdige.



Figur 12 Samlet effekt over to år, referanseår 2019

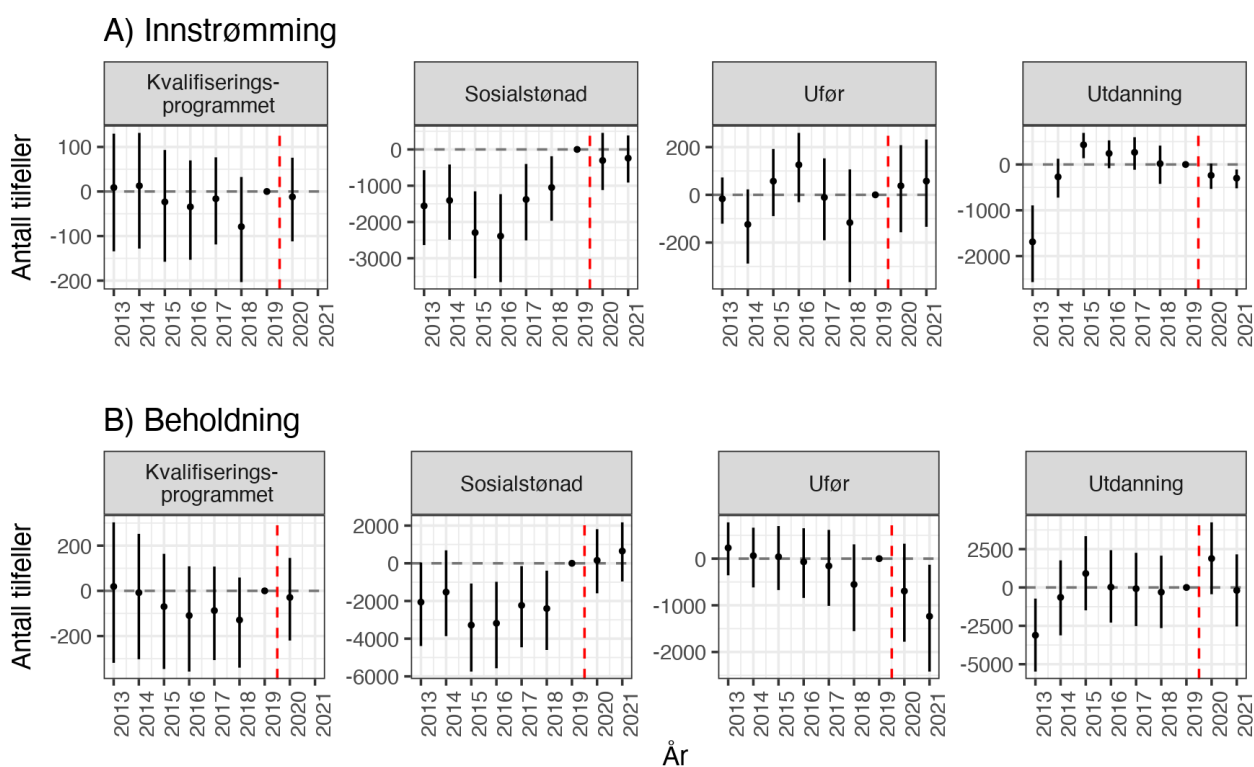
Note: Snitt, 50% og 95% kredibilitetsintervall på modellens anslåtte reformeffekt på AAP tilstrømmingen over to post-reformår for ulike verdier på tredjeordens-prior.

Reformens sammenfall i tid med covid-pandemien gir ytterligere grunn til å være varsom med å trekke bastante konklusjoner. Pandemi-tiltakene førte til at hele bransjer ble satt på hold, og mange utdanningsinstitusjoner stengte eller la om undervisningsform. Siden yrke, sektor og utdanningstilknytning varierer systematisk med alder og individuell inntektshistorie tilsier dette at de *inntektsspesifikke* trendene kan ha endret seg ulikt for de med høy og lav alder, mens de *aldersspesifikke* trendene kan ha endret seg ulikt for de med høy og lav inntekt. Hvis det er tilfellet, kan pandemien og pandemiresponsen ha skapt spuriøse reformeffekter i våre analyser.

5.2 Utfall: overgang til kvalifiseringsprogrammet, sosialhjelp, uføretrygd, utdanning

Sammendrag: For kvalifiseringsprogrammet følger sammenligningsbanen de faktiske tallene gjennom pre-reform perioden, men det ene året med post-reform-data viser ingen antydning til endret innstrømming eller beholdning. De statistiske modellene som var utarbeidet for innstrømming til AAP-programmet viste seg uegnet til å analysere sosialhjelp og utdanning. For uførhet finner analysen av innstrømming ingen reformeffekter.

Det sentrale utfallet vi var interessert i var AAP-tilstrømming, men dersom det er betydelige effekter på AAP-tilstrømming ville det være nyttig å vite hvor disse endrede strømmene kommer fra eller går til. Vi kjørte derfor estimeringsmodellen vi lagde for AAP-tilstrømming med den samme konservative modellen for fire utfall: Kvalifiseringsprogrammet, sosialhjelp, mottak av uføretrygd og utdanning. Vi presenterer her DiDiD-resultater for to ulike mål for hvert utfall: Innstrømming (sannsynlighet for å gå inn i tilstanden gitt at du er *at risk* for AAP (potensiell mottaker AAP) og ikke har vært i den angjeldende tilstanden tidligere i dataperioden), og beholdning (sannsynlighet for å være i tilstanden gitt at du er *at risk* for AAP).



Figur 13 – Analyser av andre utfall.

Note: Forventning og 95% KI når estimeringsmodellen kjøres på innstrømming og beholdning for kvalifiseringsprogrammet, sosialhjelp, mottak av uføretrygd og utdanning.

Modellen virker rimelig på kvalifiseringsprogrammet, men resultatene viser ingen antydning til noen reform-utløst endring: Anslaget på reform-effekt på innstrømming er -13 personer (95% KI fra -110 til +76) mens den på beholdning er -29 personer (95% KI fra -238 til +153). For kvalifiseringsprogrammet har vi kun ett post-reformår med data.

Den direkte innstrømmingen til uføretrygd er også relevant å analysere. Reduksjonen i minsteytelsen for AAP gjorde forskjellen mot uføretrygdens minsteytelse enda større enn tidligere. Dette kan forventes å gi økt press for å komme direkte inn i uføreordningen. Avviklingen av ung-ufør-tillegget for AAP-mottakere vil trekke i samme retning. Vi analyserer derfor mottak av uføretrygd på samme måte som mottak av AAP. For uførhet er modellen kun rimelig å anvende på innstrømming, ettersom det ikke er noen grunn til å tro at varighet eller utstrømming påvirkes. Vi ser at estimatene for innstrømming til uføretrygd i pre-reformårene svinger rundt null, som tyder på at sammenligningsbanen matchet den faktiske utviklingen her. I de to årene etter reformen er det ingen tegn til endring. Punkttestimatene er svært nært 0. Vi konkluderer med at vi ikke ser noen tegn til at regelendringene førte til økt mottak av uføretrygd. For konsistens i figuren viser vi også resultatene når samme modell brukes på beholdningen.

Pre-reformestimatene i eventplottet for sosialhjelp og utdanning tyder på at det er klare brudd på modell-antagelsene for disse utfallene. De klare avvikene i pre-reform perioden når vi bruker modellen vår på sosialhjelp og utdanning er ikke i og for seg pussig, da antagelsene som er rimelige når vi ser på en helserelatert «forsikringsordning» som AAP kan være mindre passende for andre utfall. Her er vi igjen litt tilbake til det problemet vi drøftet i kapittel 3 rundt DiDiD-modeller: Med en trippel-diff beregner vi effekter fra subtile «høyere-ordens» sammenligninger som kan være krevende å følge med intuisjonen. Vi antar f.eks. at aldersrelaterte forskjeller i trender er de samme

for personer med høy og lav inntekt, noe som virker lite plausibelt for sosialhjelp. Sosialhjelp er et siste sikkerhetsnett i det offentlige støttesystemet som gis til personer uten oppsparte verdier å tære på når andre ytelser ikke er tilgjengelig eller tilstrekkelig. Det vil være få som mottar denne blant de med høyere inntekt, og endringer i økonomiske forhold vil trolig slå ut ulikt på denne gruppen sammenlignet med måten de samme endringene i renter, priser og arbeidsmarked slår inn på de med lavere inntekt.

Modellen fremstår også plausibel i pre-reform årene 2013-2019 for beholdningen i utdanning, men her er punktestimatet +1915 (-368 til +4315) i 2020 og tett på null med -148 (-2445 til +2140) i 2021. Gitt at en reell reformeffekt ville vedvare utover 2020, at tallet er betydelig større enn det vi finner av rimelige responser på AAP-innstrømming, og gitt sammenfallet mellom reformtidspunkt og koronapandemien ser vi liten grunn til å tolke dette som en reformeffekt. Utdanning skiller seg også fra de andre utfallene ved at en mye større andel av personer som er i denne tilstanden ikke er sannsynlige mottakere av AAP. Forbindelsen mellom AAP-regelverket og endringer i dette og utfallet blir dermed svakere med relativt større rom for påvirkning fra andre faktorer.

I sum tilsier disse resultatene at modellen fremstår som uegnet for å analysere hvordan reformen påvirket innstrømming og beholdning til utdanning og sosialhjelp. For kvalifiseringsprogrammet ser modellen mer egnet ut, men her er det ingen tegn til effekter av reformen. Som nevnt i innledningen kan dette henge sammen med pandemien og responsen på den: AAP-reformen var ment å være del av en dreining mot økt bruk av arbeidsrettede tiltak for unge personer med svak tilknytning til arbeidsmarkedet, men kapasitetsproblemer i NAV og økonomisk usikkerhet kan ha gjort at dette i liten grad ble realiteten.

6 Resultater for faktiske AAP-mottakere

6.1 Supplerende sosialhjelp blant AAP-mottagere

I dette kapittelet undersøker vi utfall for faktiske mottakere av AAP. Regelendringene reduserte utbetalingene til nye AAP-mottagere under 25 år som kun kvalifiserte for minsteytelsen og avvirket ung ufør-tillegget for nye AAP-mottakere. Begge endringene kan ha ført til en økning i supplerende sosialhjelp for å kompensere for det reduserte støttenivået i AAP-ordningen. Imidlertid tror vi at bidraget fra avviklingen av ung ufør-tillegget hadde vært lite å si, ettersom svært få fikk dette. Fra Arbeids- og velferdsdirektoratet har vi mottatt tall for antall mottakere av AAP som var registrert som «ung ufør» innen 12 måneder. Denne andelen har vært sterkt fallende over tid og var for nye mottakere av AAP i 2019 under 25 år bare 2 prosent.

Vi fokuserer på personenes første AAP-forløp for å unngå komplikasjoner knyttet til å definere hva som er ulike forløp og tolke resultater for perioder mellom to forløp. Vi grupperer personene etter hvilket år de startet sitt første AAP-forløp og fokuserer på årene 2018-2021 – årene rundt regelendringene. Vi betinger også på at inntektsgrunnlaget for AAP var 3G eller mindre, så alle som er inkludert her ville ha kvalifisert for minstenivået på AAP før regelendringene. Disse personene følger vi måned for måned i 12 måneder før de begynte å motta AAP og 24 måneder etter. For hver måned kan vi da beregne hvor mange som for eksempel mottok sosialhjelp, hadde arbeid, osv.

Figur 14 viser, for personer som startet et AAP-forløp i løpet av årene 2018-2021, andelen som mottok sosialhjelp måned for måned før og etter de startet AAP-forløpet. Punktet 0 på den horisontale akse betegner måneden før AAP-forløpet startet, 1 er første måned personen mottar AAP, 2 andre måned, osv. Vi følger personene opp til 24 måneder fra de begynte å motta AAP. De som slutter å motta AAP før dette er bare med i beregningen av tallene mens de faktisk mottar AAP. Tilsvarende følger vi personer i tiden før AAP-forløpet. Siden 0 er måneden før AAP-forløpet startet betegner -1 måneden før dette, osv. Vi følger AAP-mottakerne bakover til og med måned -11 slik at 12 måneder før AAP-forløpet er inkludert.

Vi viser tall for to aldersgrupper, dem som var hhv. 18-23 og 25-28 år ved oppstarten av forløpet. 24-åringer er utelatt fordi de relativt snart blir 25 år og dermed mottar det samme som de eldre og da ikke egner seg som sammenligningsgruppe. Ettersom sosialhjelp er en ytelse for personer over 18 år, teller vi ikke med observasjoner før AAP-forløpet for personer som begynte å motta AAP det året de fylte 18 år.

Først ser vi på nivået på andelen som mottar sosialhjelp blant personer som startet et AAP-forløp i 2018 eller 2019, altså før regelendringene. Det er tydelig at i månedene som ledet fram til første AAP-mottak var det flere og flere som mottok sosialhjelp – fra 10-20 prosent ett år før AAP-forløpet (tidspunkt -11), til i underkant av 30 prosent ved tidspunkt 0, dvs. måneden før AAP-forløpet. Det er noen prosentpoeng færre blant de yngste som mottar sosialhjelp, dette kan henge sammen med at det er flere i denne gruppa som bor hos foreldre.

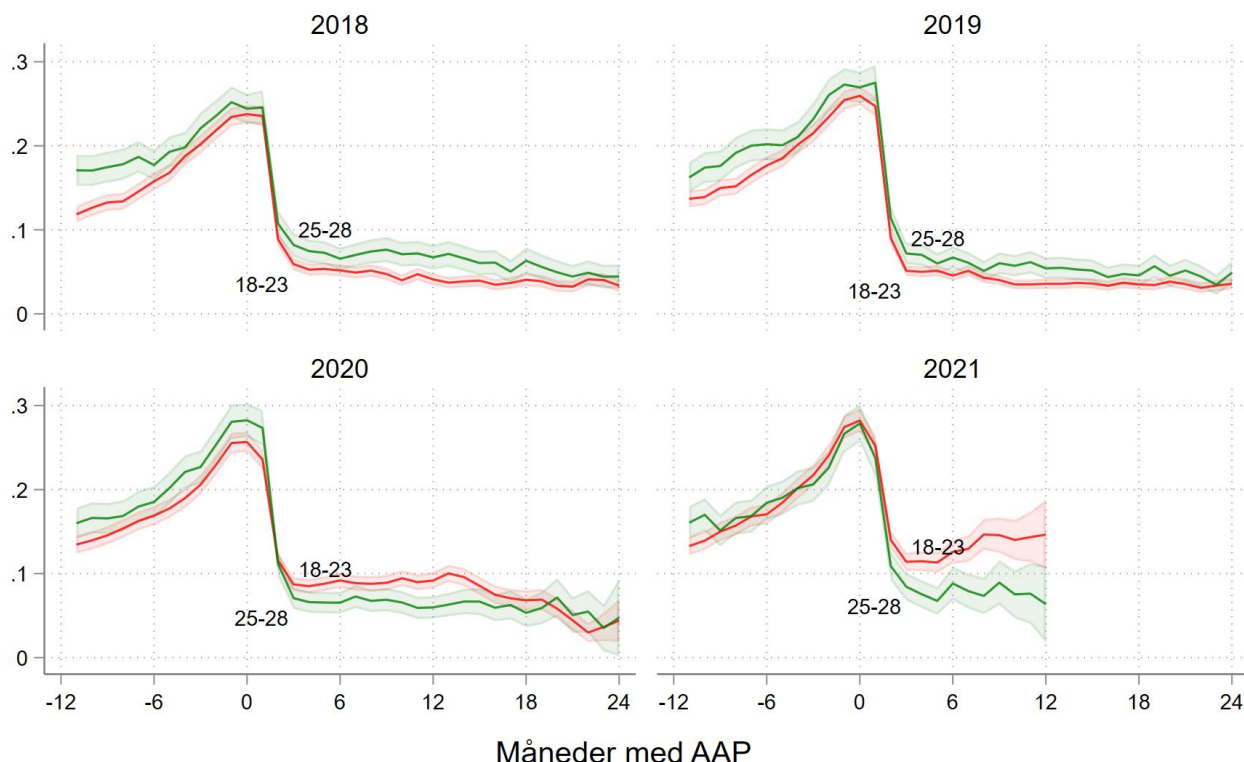
Fra og med innvilgelse av AAP, dvs. fra måned 1 på den horisontale akse, er det naturlig nok færre som har behov for sosialhjelp, og vi ser at andelen med sosialhjelp faller raskt før den flater ut. Blant dem som startet sitt AAP-forløp i 2018 og 2019 ser vi at andelen med supplerende sosialhjelp falt til rundt 5 prosent fra og med måned 3 for de yngste, og et par prosentpoeng høyere blant de eldre AAP-mottakerne.

Så går vi videre til å se på sosialhjelpsmottaket for personer som startet sitt AAP-forløp i 2020 eller 2021, og altså var omfattet av regelendringene. Fram til tidspunkt 0 er mønsteret svært likt for personer med oppstart i 2020 som tidligere. I 2021 er det samme stigning i andelen over tid fram til tidspunkt 0, men vi ser at de yngste ligger mer på nivå med de eldste enn det som var tilfellet de tidligere årene. For supplerende sosialhjelp, dvs. mottak av sosialhjelp samtidig med AAP, skjer det en tydelig endring i forholdet mellom de yngste og de eldste. Vi fokuserer på perioden fra og

med tidspunkt 3, dvs. tredje måned av AAP-forløpet, da andelene med supplerende sosialhjelp i stor grad har stabilisert seg. For personer som startet sitt AAP-forløp i 2020 og 2021 er det i motsetning til tidligere år de yngste mottakerne som i størst grad mottar supplerende sosialhjelp. Fra mottaks måned 3 og utover er det for personer med oppstart i 2020 eller 2021 11 prosent blant de yngste som mottar supplerende sosialhjelp, mot 7 prosent hos de eldste.

Det er naturlig å tolke det at mønsteret mellom de to aldersgruppene endret seg akkurat fra og med 2020 som en konsekvens av regelendringene. Hvis vi tar utgangspunkt i gjennomsnittlig nivå på supplerende sosialhjelp i måned 3-12 av AAP-forløpet, kan vi beregne et forskjell-i-forskjeller-estimat på reformeffekten, basert på en antakelse om at nivåene ville ha utviklet seg likt på samme måte i fravær av regelendringene. Dette ville gitt et estimat på ca. 6 prosentpoeng i supplerende sosialhjelp blant de yngste, ettersom forskjellen mellom gruppene gikk fra -2 (i årene 2018 og 2019) til +4 (i årene 2020 og 2021). Vi skal imidlertid være forsiktige med tolkningen her, ettersom regelendringen også kan ha medført sammensetningsendringer i gruppene. Punkttestimatet på innstrømmingen blant personer under 25 år var negativt, jf. kapittel 5.1. Hvis innstrømmingen til AAP faktisk ble redusert og reduksjonen var drevet av personer med bedre andre alternativer og relativt lav tilbøyelighet til å motta supplerende sosialhjelp, kan det være en grunn til at andelen blant dem som er igjen skulle være høyere. De faktiske AAP-mottakernes mottak av sosialhjelp før AAP-forløpets start gir noe informasjon om seleksjonen inn kan ha blitt påvirket. For personer som startet sitt AAP-forløp i 2018-2020 ser det i liten grad ut til å være noen forskjell i andelen som mottar sosialhjelp før de begynte å motta AAP. For personer med oppstart i 2021 er det imidlertid annerledes, her har de yngste mottakerne en liten, men tydelig høyere sosialhjelpsandel før AAP-forløpet enn det de yngste hadde i tidligere år. Det kan være en indikasjon på at de yngre som etter hvert begynte på AAP i gjennomsnitt hadde større behov enn tidligere også før de begynte å motta AAP. Imidlertid skulle vi da forventet det samme også i 2020. Det er også viktig å være klar over at forskjellen i 2021 er liten – i 2018-2020 var det i gjennomsnitt to prosentpoeng færre blant de yngste enn de eldste som mottok sosialhjelp i året før AAP-forløpet, og det er denne forskjellen som er borte for personer med AAP-oppstart i 2021. Ettersom vi kun ser på første AAP-mottak er det ikke knyttet til endringer i hvor mange som begynte et nytt forløp.

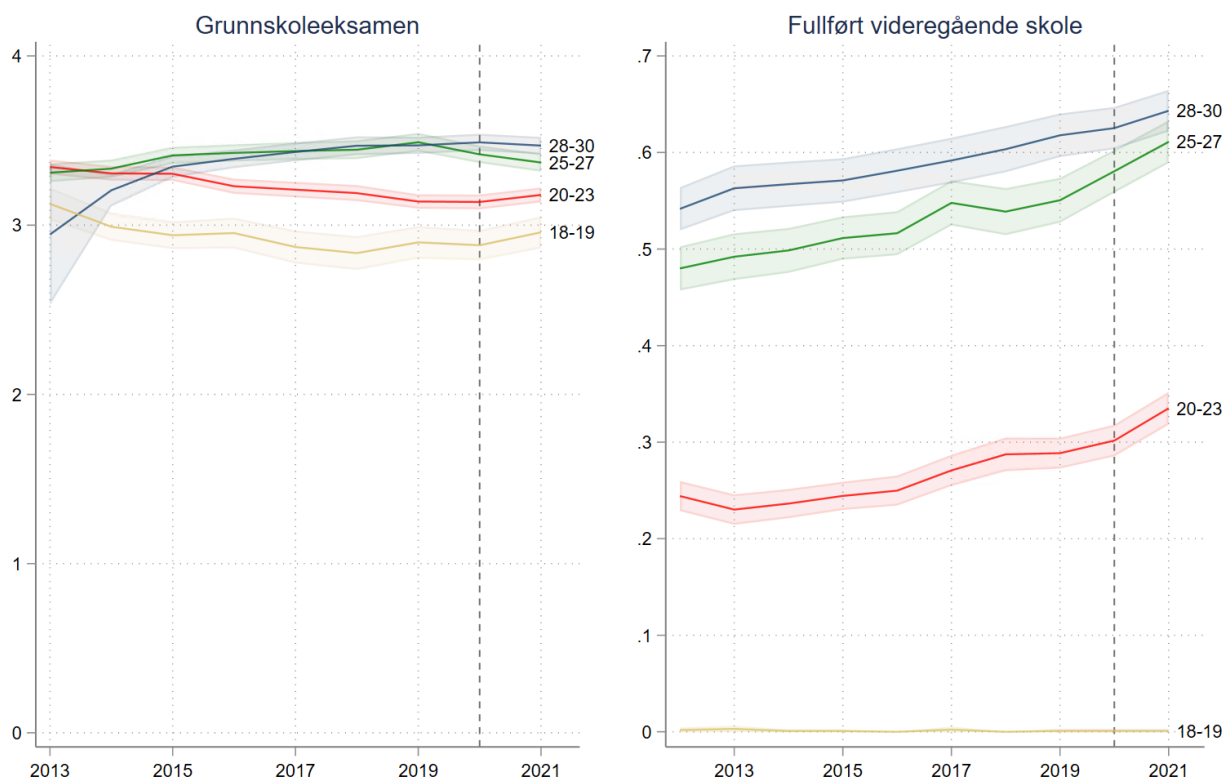
Mottak av sosialhjelp etter startår for AAP-forløp og alder ved oppstart



Figur 14 Mottak av sosialhjelp etter startår for AAP-forløp og alder ved oppstart

Note: Hvert punkt viser andel som mottok sosialhjelp per måned, der månedene er definert relativt til AAP-forløpet. Skraverete områder angir 95 % konfidensintervall.

I figur 15 går vi ytterligere inn i spørsmålet om endringer i sammensetningen av mottakergruppen ved å se på eksamenskarakterer fra grunnskolen og fullført utdanning. Vi ser her kun på nye mottakere i hvert år. Venstre panel viser gjennomsnittlig eksamenskarakter ved avslutningen av grunnskolen. Dette er altså mange år før AAP-forløpet begynte, men det kan likevel fortelle noe om utviklingen i sammensetningen av de nye AAP-mottakerne. Selv om det er forskjell i nivået på de ulike aldersgruppene, ser det ut til at det gjennomsnittlige karakternivået i hovedsak er stabilt. Spesielt ser vi ikke tegn til endring i retning av mer eller mindre skolesterke personer i den yngste aldersgruppen av nye AAP-mottakere. Høyre panel viser andel nye mottakere som hadde fullført videregående skole året før. Vi ser at det fra 2013 har vært en økende andel nye AAP-mottakere med fullført videregående skole og at denne trenden ikke ser ut til å være brutt fra 2020. Oppsummert finner vi ikke tegn til at sammensetningen av nye AAP-mottakere endret seg i 2020. Dette er i tråd med Danielsen, Kalstø og Grønlien (2023), som så på utviklingen i andel med ulike diagnoser på aggregert nivå.



Figur 15 Eksamenskarakter og fullføring av videregående skole hos nye AAP-mottakere

Note: Hvert punkt viser gjennomsnitt for nye mottakere per år. Skraverte områder angir 95 % konfidensintervall.

6.2 Avgang fra AAP og overgang til uføretrygd, utdanning og arbeid

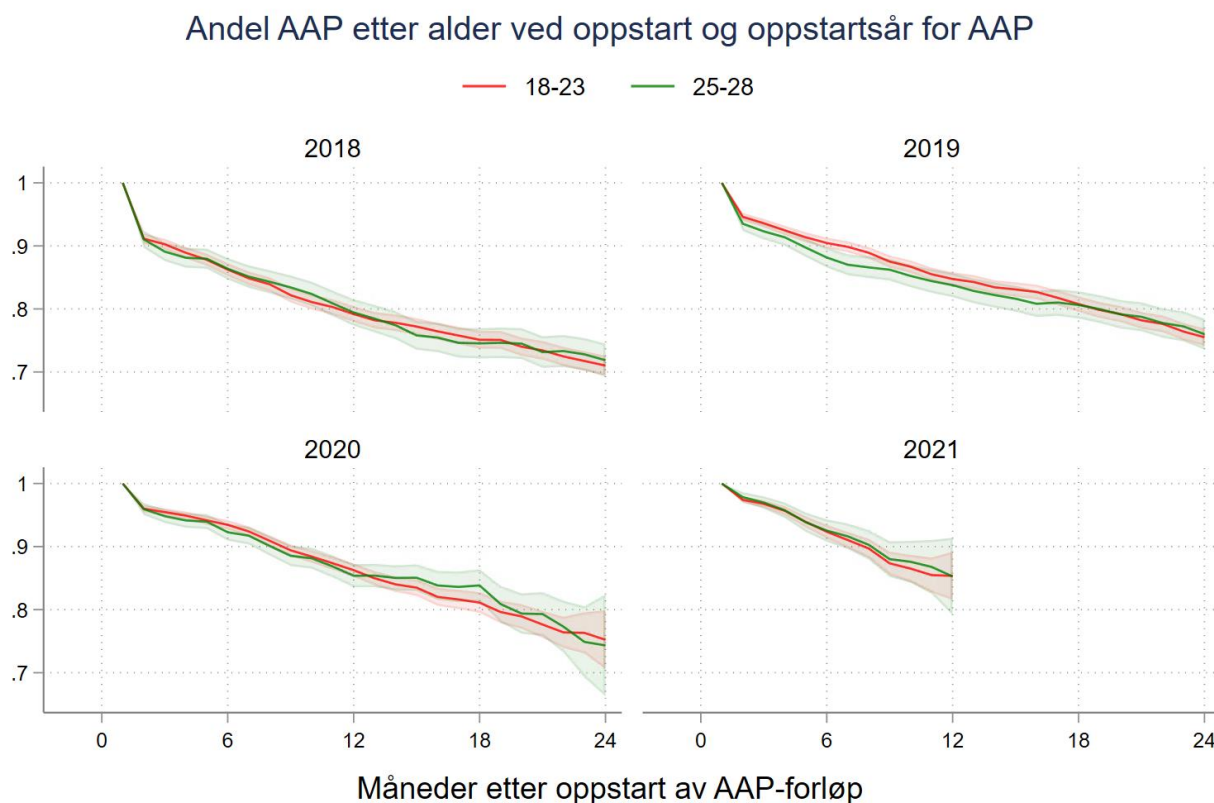
Regelendringene kan også ha påvirket utstrømmingen fra AAP. Det at AAP-ordningen ble mindre økonomisk gunstig for mottakere under 25 år kan ha økt insentivene for aktivitet ved å gjøre utdanning og arbeid relativt mer attraktivt. Samtidig er det en risiko for at en mer presset økonomisk livssituasjon kan være belastende og virke i den andre retningen, noe mange veiledere var opptatt av i intervjuene og fokusgruppemøtet.

En enkel måte å undersøke avgang fra AAP og overgang til andre tilstander er å sammenlikne mønsteret for personer under 25 år, som ble påvirket av endringene, med personer over 25 år, som ikke ble påvirket. Det gjør vi i dette delkapittelet. Til orientering var disse analysene ikke gjenstand for utprøving og forhåndsregistrering på samme måte som hovedanalysene av virkning på innstrømmingen til arbeidsavklaringspenger. Som i delkapittel 6.1 fokuserer vi her på personer som startet sitt første AAP-forløp i årene 2018-2021 og skiller mellom personer som da var i alderen 18-23 eller 25-28. Siden vi her er interessert i overganger til andre tilstander, fokuserer vi på perioden fra og med den måneden AAP-forløpet startet.

Figur 16 viser andel registrert med AAP i løpet av hver måned. Definisjonsmessig er denne andelen lik 1 for alle i den første måneden av et AAP-forløp. Nivået på denne andelen blir dermed et mål på avgangen fra AAP. Vi ser at avgangen ser ut til å være tilnærmet uendret etter regelendringene. For personer med oppstart av AAP-forløpet i 2018 er avgangsmønsteret nærmest identisk for de yngste og de eldste. I 2019 er det en tendens til at de eldste forsvinner noe raskere ut av AAP i de første 18 månedene. I 2020 er kurvene igjen svært like, men det er mulig å ane en forskjell i avgangen fra og med måned 12, der kurven for de yngste ligger noe under kurven for de eldste. Det kan skyldes raskere avgang for de yngste, men det ser faktisk ut til at det er kurven for de eldste

som skiller seg fra tidligere år, med en større andel som fortsatt er på AAP mellom 12 og 24 måneder. Dette kan ha flere årsaker. Spesielt kan pandemien ha hatt noe å si – store deler av arbeidsmarkedet og flere tiltak fungerte ikke som normalt og det tok lenger tid å få avklart mottakere. Dette kan potensielt ha påvirket de yngste annerledes enn de eldste. Det kan også skyldes tilfeldig variasjon. Forskjellene er uansett små.

For personer med oppstart i 2021 ser mønsteret igjen likt ut i den perioden vi kan følge dem.

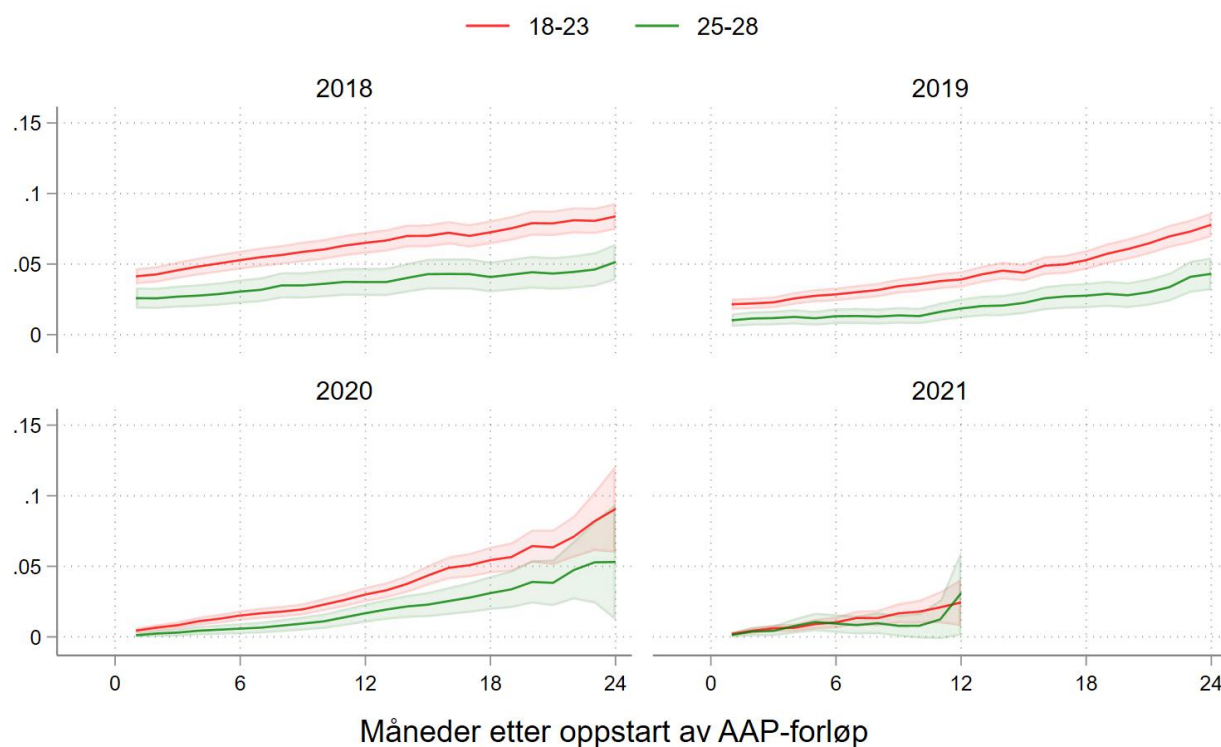


Figur 16 Andel AAP etter alder ved oppstart og oppstartsår for AAP

Note: Hvert punkt viser andel som mottok AAP per måned, der månedene er definert relativt til AAP-forløpet. Skraverte områder angir 95 % konfidensintervall.

Figur 17 viser hvor mange som er registrert med uføretrygd måned for måned etter oppstart av AAP-forløpet. Her er det noen endringer over tid. I 2018 og 2019 var det helt fra starten av AAP-forløpet en liten andel i begge grupper som mottok uføretrygd. Det er uklart hva dette skyldes. Avgangsmønsteret over tid er ganske likt for begge gruppene disse årene, men det er en tendens til at andelen med uføretrygd øker raskere blant de yngste AAP-mottakerne enn blant de eldre. For forløp startet i 2018 og 2019 var forskjellen i andel med uføretrygd mellom de yngste og de eldste 1-1,5 prosentpoeng ved oppstarten av AAP-forløpet (tidspunkt 1), men hadde vokst til nærmere 3,5 prosentpoeng ved etter 24 måneder. Det tyder på at de yngste ble raskere avklart til uføretrygd. For personer med oppstart i 2020 er den raskere overgangen hos de yngste noe mer tydelig, siden begge kurver starter nærmere 0. For disse personene er forskjellen etter 24 måneder også rundt 3,5 prosentpoeng. Dette kan tolkes som at regelendringene førte til raskere avklaring til uføretrygd hos de yngste, men forskjellen er liten og i 2021 er det i praksis ingen forskjell mellom gruppene. Det er også betydelig usikkerhet, jf. stort overlapp mellom konfidensintervallene.

Andel uføretrygd etter alder ved oppstart og oppstartsår for AAP

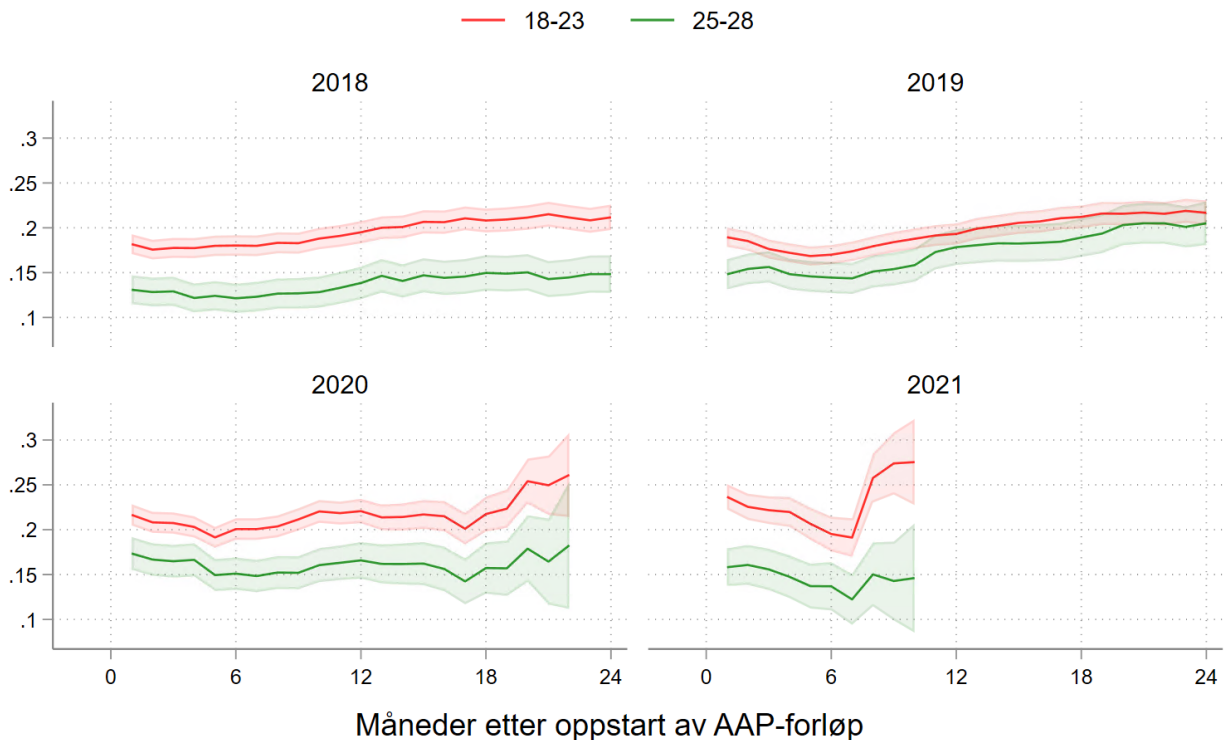


Figur 17 Andel uføretrygd etter alder ved oppstart og oppstartsår for AAP

Note: Hvert punkt viser andel som mottok uføretrygd per måned, der månedene er definert relativt til AAP-forløpet. Skraverte områder angir 95 % konfidensintervall.

Figur 18 viser andel registrert i utdanning. Både videregående opplæring og høyere utdanning er her regnet med. En klart høyere andel hos de yngste enn hos de eldste er registrert i en eller annen form for utdanning. Forskjellen mellom aldersgruppene varierer mye fra år til år. Det gjør det vanskeligere å konkludere. For personer med AAP-oppstart i 2018 var forskjellen i andelen i utdanning stor, men konstant. I 2019 var forskjellen mindre og minkende over tid. I 2020 var forskjellen stor i utgangspunktet og økende etter hvert. Det samme var tilfellet i enda sterkere grad i 2021. Andelene mot slutten av perioden i panelene for 2020 og 2021 er basert på færre og færre personer, ettersom det er færre vi ha mulighet til å følge så lenge etter AAP-oppstarten. Det gjør at disse tallene er mer usikre, noe som indikeres av de økende konfidensintervallene. I tillegg medfører det enkelte endringer i sammensetningen av mottakerne gjennom at det bare er de som begynte AAP-forløpet sitt helt i starten av 2020 eller 2021 vi kan følge til hhv. måned 22 og 10. (Utdanningsregisteret er oppdatert per oktober 2021, vi kan derfor ikke følge noen fra 2021 i opp til 12 måneder.) Oppsummert finner vi ikke noen systematisk endring i overgang til utdanning mellom yngre og eldre AAP-mottakere etter regelendringene.

Andel i utdanning etter alder ved oppstart og oppstartsår for AAP

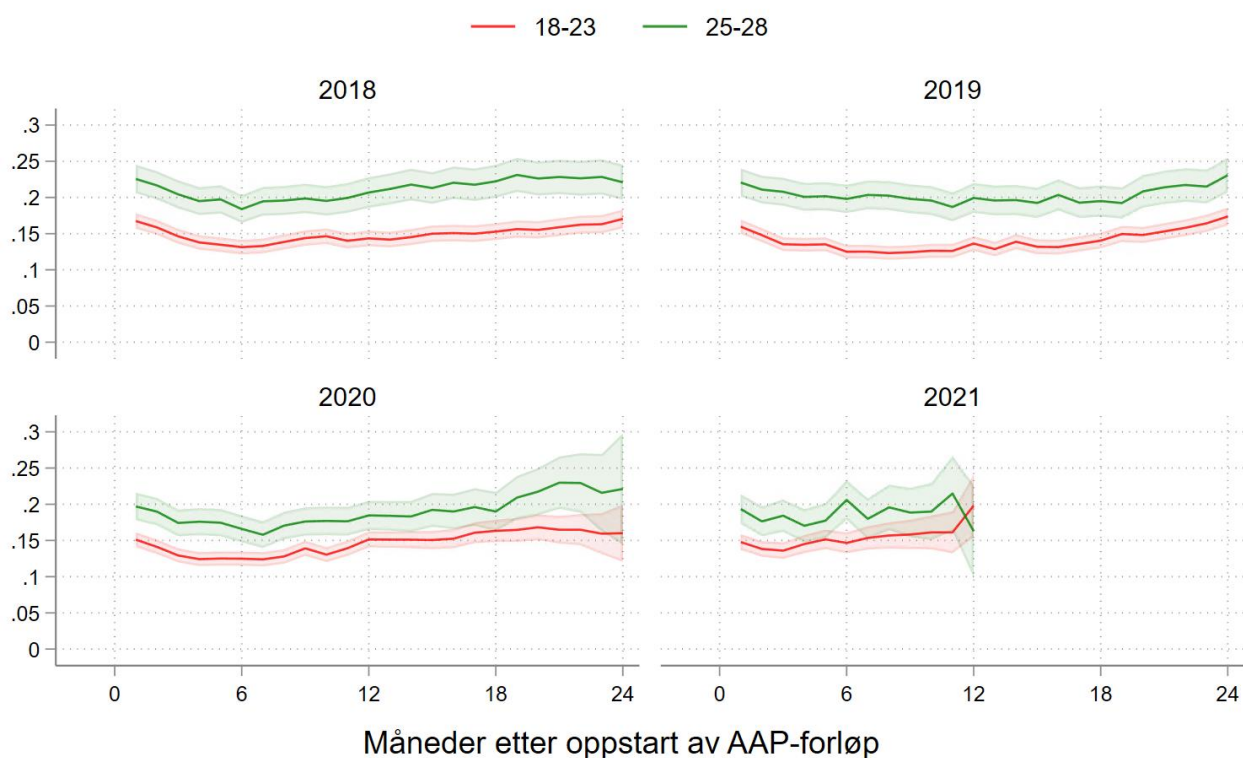


Figur 18 Andel i utdanning etter alder ved oppstart og oppstartsår for AAP

Note: Hvert punkt viser andel som var registrert i utdanning per måned, der månedene er definert relativt til AAP-forløpet. Skraverete områder angir 95 % konfidensintervall.

Til slutt går vi inn på andel som er registrert med et arbeidsforhold. Vi teller bare med arbeidsforhold hvor det er registrert utbetalt lønn. For hver måned beregner vi andelen som er registrert med et arbeidsforhold og har fått utbetalt minst 1 krone i lønn. Figur 19 viser denne andelen over tid for de ulike gruppene på samme måte som tidligere figurer. Vi ser at det blant de eldste er en klart høyere andel med lønnsinntekt for personer som startet AAP-forløpet i 2018 og 2019. Over AAP-forløpet beveger begge grupper seg relativt likt, uten store endringer i andel med lønnsinntekt. For forløp som startet i 2020 og 2021 er forskjellen mellom de to aldersgruppene noe mindre. At forskjellen er mindre er konsistent med en hypotese om at en mindre generøs AAP-ytelse kan ha gjort arbeid mer attraktivt. Forskjellen er imidlertid svært liten, og ved nærmere ettersyn kan vi se endringen først og fremst skyldes at det var noe færre med lønn blant 25-28-årsgruppa for personer med AAP-oppstart i 2020 og 2021, noe det er naturlig å sette i sammenheng med pandemien. Oppsummert finner vi ikke klare tegn til at regelendringene førte til økt overgang til lønnet arbeid for AAP-mottakere. Hvis endringene hadde en slik effekt, må den ha vært liten.

Andel med lønn etter alder ved oppstart og oppstartsår for AAP



Figur 19 Andel med lønn etter alder ved oppstart og oppstartsår for AAP

Note: Hvert punkt viser andel som mottok lønn per måned, der månedene er definert relativt til AAP-forløpet. Skraverte områder angir 95 % konfidensintervall.

7 Konklusjoner

I denne rapporten har vi evaluert virkninger av redusert minstesats for mottakere av arbeidsavklaringspenger (AAP) under 25 år og avvikling av ung ufør-tillegget for AAP-mottakere. Vi har hovedsakelig fokusert på endringen om reduksjon i minstenivået, ettersom svært få AAP-mottakere har mottatt ung ufør-tillegget. Fra Arbeids- og velferdsdirektoratet har vi mottatt tall for antall mottakere av AAP som var registrert som «ung ufør» innen 12 måneder. Denne andelen har vært sterkt fallende over tid og var for nye mottakere av AAP i 2019 under 25 år bare 2 prosent. Det har også vært tilfellet at «om lag halvparten av de som får innvilget ung ufør-tillegget [får] dette utbetalt først *etter* at perioden på arbeidsavklaringspenger er over, som følge av at retten til tillegget først er blitt avklart underveis i perioden (Prop. 10 L 2019-2020).»

Vi har to sett av kvantitative hovedresultater. Det første er knyttet til virkninger på innstrømmingen til AAP, det andre knyttet til ulike utfall for faktiske mottakere av AAP. I tillegg har vi to sett av kvalitative resultater. Det første bygger på intervjuene med NAV-veiledere, jf. kapittel 2. Det andre bygger på et fokusgruppeintervju med veiledere fra to NAV-kontor som ble gjennomført mot slutten av prosjektperioden. Formålet med dette gruppeintervjuet var å tolke, nyansere og utdype funn fra den kvantitative analysen. Fokusgruppeintervjuet var lagt opp slik at deltakerne fikk presentert funn og diskuterte deretter disse med utgangspunkt i sine erfaringer. Det var til sammen syv deltakere. Blant disse var det representert både avdelingsledere, teamkoordinator og fagutvikler for AAP-unge, veiledere som fulgte opp unge AAP-mottakere og økonomisk rådgiver. I dette avslutningskapittelet oppsummerer vi og setter det kvantitative og kvalitative materialet i sammenheng.

1. februar 2020 ble minstesatsen for AAP-mottagere under 25 år redusert med en tredjedel fra sitt tidligere nivå på 2G, samtidig som ung-ufør-tillegget på 0,44 G, for personer som fikk nedsatt arbeidsevne før fylte 26 år, ble fjernet. Formålet med endringene var å gjøre AAP mindre økonomisk attraktivt relativt til innteksbringende arbeid og andre ytelser og tiltak og å redusere sykdomsfokuset i denne perioden. Endringen i minstesatsen påvirket kun støttenivået til personer under 25 med lav forutgående inntekt (under 3G). En potensiell mottager under 25 år som kun kvalifiserte for minsteytelse, ville etter endringene motta 1,33 G i stedet for 2G årlig. Dersom personen i tillegg ville kvalifisert for ung ufør-tillegget under de gamle reglene, var reduksjon på nær 50 % (1,33 G vs. 2,44 G).

For å få mer kunnskap om prosessene hvor unge uten arbeid søker om og eventuelt får AAP, og hvordan reglene kan påvirke dem når det gjelder å søke AAP, søke andre ytelser, og komme i arbeid eller utdanning, gjennomførte vi gruppeintervjuer med veiledere ved tre NAV-kontorer. Ingen av veilederne som deltok i gruppeintervjuene hadde lagt merke til noen vesentlig endring i søkertilstrømmingen av unge til AAP. Veilederne beskriver generelt en gruppe med nokså lav byråkratisk kompetanse og har heller ikke merket noen nedgang i tilstrømming til AAP etter at regelendringen trådte i kraft. Slik veilederne fremstiller det er det vanlig at en søknad om AAP skjer etter konsultasjon med en veileder, og veilederne forteller at de i stor grad gir de samme rådene som tidligere. I sakstyper der rådet tidligere ville vært å søke AAP, er fremdeles rådet det samme etter regelendringen.

Resultatene fra den kvantitative analysen av innstrømmingen til AAP, jf. delkapittel 5.1, kan tyde på noe større effekt på AAP-tilgangen enn det veilederne trodde. Resultatene derfra tilsier at regelendringene reduserte antallet nye AAP-tilfeller blant unge med lav inntekt med 208 tilfeller i 2020 og 170 tilfeller i 2021. Ettersom det er rundt 3500 nye AAP-tilfeller årlig fra denne gruppen som helhet, tilsier disse «punktestimaterne» en nedgang på rundt 5-6 prosent. Selv om en reduksjon på 5-6 prosent er substansielt, vil det på ett enkelt NAV-kontor kanskje dreie seg om 1-2 personer. Det er derfor naturlig at veiledere kan oppfatte dette som ingen forskjell. Det er også viktig å være klar over at estimatene er innenfor den variasjonen fra år til år som forekom i årene før 2020. Vi kan derfor ikke utelukke at de skyldes tilfeldig variasjon som ikke hadde med regelendringene å

gjøre. Reformtidspunktet sammenfalt videre tett med Covid-pandemien og samfunnets respons på denne, noe som også bidrar til å gjøre analysen mer usikker.

Veilederne i fokusgruppeintervjuet mente også at de aller fleste aktuelle AAP-mottakere i liten grad var i stand til å shoppe ytelse. Dette er også i tråd med resultatene fra Strand og Svalund (2021). Potensielle mottakeres familie kunne derimot spille en viktig rolle, men hovedinntrykket var uansett at de aller fleste søkerne søkte AAP på grunn av dårlig helse og i liten grad lot seg påvirke av nivået på ytelsen. Imidlertid var ikke dette helt entydig, og det ble nevnt at også etter endringene var AAP økonomisk mer gunstig enn tiltakspenger. At satsen på tiltakspenger var redusert ble trukket fram som negativt, ettersom det gjorde tiltak mindre attraktivt sammenliknet med andre aktiviteter, som AAP. Flere var positive til mer bruk av kvalifiseringsprogrammet, men trakk fram både trange kommunale budsjetter og behov for kompetanse som forklaringer på at dette ikke ble brukt mer. Angående direkte overgang til uføretrygd trakk de fram at dette var svært uvanlig.

Når det gjaldt bruk av supplerende sosialhjelp var veilederne i intervjuene samstemte i at regelendringene kom til å ha stor effekt. Enkelte av veilederne vi intervjuet fremstilte det som at nærmere alle AAP-mottakere også får supplerende sosialhjelp, og selv om ikke det gjaldt alle veilederne, var det samlede inntrykket at regelendringene kom til å føre til en kraftig økning i supplerende sosialhjelp.

I tråd med veiledernes forventinger fant vi en økning i bruk av supplerende sosialhjelp hos den nye mottakere av AAP under 25 år etter regelendringene. Imidlertid var både nivået i utgangspunktet lavere enn vi fikk inntrykk av, og økningen mindre, jf. kapittel 6.1. Vi finner at sosialhjelp ser ut til å spille en viktig rolle som inntektssikring i tiden før en person starter et AAP-forløp – da mottar opp mot 30 prosent av framtidige AAP-mottakere sosialhjelp. Etter at AAP-forløpet er i gang er det imidlertid mange færre som mottar sosialhjelp. Blant dem som startet sitt AAP-forløp i 2018 og 2019, altså før regelendringene, finner vi at andelen med supplerende sosialhjelp var rundt 5 prosent for de nye AAP-mottakere i alderen 18-24, og et par prosentpoeng høyere blant nye mottakere i alderen 25-28. Blant nye mottakere av AAP under 24 år i 2020 og 2021 var det i en gjennomsnittlig måned 11 prosent som mottok sosialhjelp, mot 7 prosent hos nye mottakere i alderen 25-28 år. Altså endret forholdet mellom de yngre og de eldre seg fra og med regelendringene. I stor grad tolker vi dette som en virkning av regelendringene, men det er verdt å være klar over at omfanget også etter regelendringene er relativt lite. Veilederne i fokusgruppeintervjuet trakk fram at det at mange antakelig bodde hjemme hos foreldrene og muligheten for bostøtte som mulige forklaringer på at økningen ikke var større. Resultatene er på linje med Lima (2023), som ser på supplerende sosialhjelp for AAP-mottakere i 4. kvartal 2017-2022, og finner en sterk økning blant mottakere i aldersgruppen 18-24 år i 2021 og 2022. Lima (2023) finner at andelen med supplerende sosialhjelp i denne aldersgruppa er om lag doblet i 2022 vs. 2017-2019, som er omtrent det samme vi finner.

Når det gjaldt overgang fra AAP til arbeid eller utdanning, trodde ikke veilederne i intervjuene at endringene skulle ha noe å si. De trakk fram at innstramminger har uheldige konsekvenser, ved at de skaper uro, usikkerhet og uforutsigbarhet for brukerne, og at økonomiske bekymringer kommer i veien for andre mål det jobbes mot. Derimot trodde de at presset mot uføretrygd kom til å øke. Vi finner imidlertid få klare tegn til dette. Mottakere av AAP i alderen 18-23 ser generelt ut til å bli noe raskere avklart til uføretrygd enn mottakere i alderen 25-28. Dette var tilfellet også før regelendringene i 2020, men det er en svak antydning til dette kan ha økt noe for nye AAP-mottakere i alderen 18-23 år i 2020. Imidlertid er forskjellene små og usikkerheten stor. Veilederne i fokusgruppeintervjuet hadde ikke noen spesielle bemerkninger til dette.

Referanser

- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics* 119 (1): 249–75.
- Caetano, Carolina, Brantly Callaway, Stroud Payne, and Hugo Sant'Anna Rodrigues. 2022. "Difference in Differences with Time-Varying Covariates." *arXiv Preprint arXiv:2202.02903*.
- Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with a Continuous Treatment." *arXiv:2107.02637 [Econ]*, July. <http://arxiv.org/abs/2107.02637>.
- Carpenter, Bob, Andrew Gelman, Matt Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Michael A. Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2016. "Stan: A Probabilistic Programming Language." *J Stat Softw.* <http://www.uvm.edu/~bbeckage/Teaching/DataAnalysis/Manuals/stan-resubmit-JSS1293.pdf>.
- Danielsen, Mia, Åshild Male Kalstø, and Eirik Grønlien. 2023. "Innstramminger i Regelverk Og Pandemi - Nye AAP-Mottagere Fra 2016-2022." *Arbeid Og Velferd*, no. 1: 33–47.
- Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics*, Themed Issue: Treatment Effect 1, 225 (2): 254–77. <https://doi.org/10.1016/j.jeconom.2021.03.014>.
- Lima, Ivar. 2023. «Få mottakere av sosialhjelp i 2022 – men økning mot slutten av året». *Arbeid og velferd*, 2/2023.
- NOU 2019: 7. (2019). *Arbeid og inntektssikring — Tiltak for økt sysselsetting*. Arbeids- og sosialdepartementet. <https://www.regjeringen.no/no/dokumenter/nou-2019-7/id2637967/>
- NOU 2021: 2. (2019). *Kompetanse, aktivitet og inntektssikring — Tiltak for økt sysselsetting*. Arbeids- og sosialdepartementet. <https://www.regjeringen.no/no/dokumenter/nou-2021-2/id2832582/>
- Ritchie, Stuart. 2020. *Science Fictions: Exposing Fraud, Bias, Negligence and Hype in Science*. Random House.
- Roth, Jonathan. 2022. "Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends." *American Economic Review: Insights* 4 (3): 305–22. <https://doi.org/10.1257/aeri.20210236>.
- Roth, Jonathan, Pedro H. C. Sant'Anna, Alyssa Bilinski, and John Poe. 2022. "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." *arXiv*. <https://doi.org/10.48550/arXiv.2201.01194>.
- Roth, Jonathan, and Pedro HC Sant'Anna. 2023. "When Is Parallel Trends Sensitive to Functional Form?" *Econometrica* 91 (2): 737–47.
- Prop. 10 L (2019-2020) Endringer i folketrygdloven og enkelte andre lover. Arbeids- og sosialdepartementet. <https://www.regjeringen.no/no/dokumenter/prop.-10-l-20192020/id2676058/>
- Strand, Anne Hege, og Jørgen Svalund. 2021. "Velferdsordninger til unge." *Fafo-rapport 2021:35*

Vedlegg

Nærmere om vår implementering av DiD og DiDiD

Dataene analyseres med en hierarkisk Bayesiansk modell som anslår sannsynligheten for at en tilfeldig trukket person med en gitt alder og inntektshistorikk får innvilget AAP i løpet av et kalenderår. Denne tilnærmingen ble valgt fordi den gjør det mulig å pålegge «myke» restriksjoner av en type som i mange situasjoner har vist seg å føre til mer presise prediksjoner når det er begrenset data (Feller og Gelman, 2015). Intuisjonen bak dette er at kunnskap om en parameter ofte lærer oss noe om andre parametere i samme modell. I vårt tilfelle er det endringer som vi ville forvente skjer gradvis, slik at forskjellene mellom to grupper typisk er mindre dess likere gruppene er i alder, inntekt og observasjonstidspunkt. Gevinsten kommer av at modellen bruker informasjon «på tvers» av celler: Modellen anslår hvor mye genuin variasjon mellom celle-sannsynligheter som trengs for å forklare observerte data, og antar deretter at de individuelle celle-parametere ligger innenfor dette variasjonsområdet.

Teknisk sett er modellen en logit-modell som spesifiserer sannsynligheten π_i for at en tilfeldig valgt person med inntektsnivå k , alder a vil oppleve hendelsen (AAP-innvilgelse) i kalenderår t :

$$\text{logit}(\pi_i|a_i, t_i, k_i) = \mu + \tau^t + \tau^k + \tau^{t \times a} + \tau^{t \times k} + \tau^{k \times a} + \tau^{t \times a \times k}$$

Her angir τ de ulike førsteordens («direkte») parameterne¹, samt parameterne for alle toveis interaksjoner og tredjeordens-interaksjoner. Hver enkelt av disse parametersettene modelleres som en egen fordeling av parametere med forventning null, og modellen anslår i hvilken grad data viser tegn til at det er genuin parametervariasjon over den dimensjonen disse parameterne skiller langs. For de to parametersettene med toveisinteraksjoner som varierer over alder ($\tau^{t \times a}$, $\tau^{k \times a}$) ble parametrene estimert ved å anta at det var *forskjellen* over alder fra et år til et annet (innad i hhv kalenderår og inntektsgrupper) som var trukket fra en fordeling

En slik hierarkisk modell kan virke rart hvis man er vant til en vanlig OLS. I en vanlig regresjonsmodell ville en slik spesifikkasjon være «mettet» og resultatene ville være en ren beskrivelse av datamateriale med alle sine særegenheter. Faktisk ville den være overmettet: En OLS med en parameter for sannsynligheten i hver $t \times a \times k$ celle ville anslå en sannsynlighet i hver av disse cellene som var lik den observerte andelen i denne cellen, og den beskrevne modellen ville ikke la seg estimere fordi den ville være ha flere parametere enn det var antall celler vi skulle beskrive.

Årsaken til at dette fungerer i vår setting, er at vi ser på de ulike parametersettene som trukket fra ulike fordelinger, og når vi anslår disse fordelingene bruker vi ytterligere informasjon som gjør at de ulike cellene bidrar til å estimere hverandre. For å se dette kan vi tenke oss to ytterpunkter. Det ene ytterpunktet er at modellen antar at alle cellesannsynligheter er like. Det andre ytterpunktet er at alle cellesannsynligheter kan variere helt fritt fra hverandre – det vi ser for 21, 22, 23 og 25-åringer forteller oss ingenting om hva som er sannsynlig for 24-åringer (det vil si en fast effekt modell). En hierarkisk modell ligger mellom disse, og bruker data til å anslå hvor mye vi kan lære «på tvers». Hvis data viser at 21, 22, 23 og 25 åringer har svært like sannsynligheter for å begynne med AAP, så antar modellen at dette trolig også vil være tilfelle for 24 åringer. Dersom data tilsier at aldersgruppene har svært ulike sannsynligheter, derimot, vil modellen trekke den slutningen at det er lite vi kan si om 24-åringene uten mye data. Den hierarkiske modellen bruker slik sett data til å finne ut hvor den skal plassere seg i spennet mellom «alle er like»-antagelsen og «ingen er like på noen måte»-ytterpunktet på den andre siden.

¹ Direkte parametere for alder ble tatt ut under testing på pre-analysedata, da det var for stor variasjon i alders-trender over de ulike andre dimensjonene til at en felles delt alderstrend hjalp modellen med estimeringer.

Som en illustrasjon: Hvis mange personer kaster ulike terninger gjentatte ganger, kunne vi bedt modellen undersøke om det er systematisk tegn til at kvinner kaster høyere tall enn menn eller om noen terningfarger gir høyere tall når de kastes. Dette ville være det vi her kalte «direkte effekter». Vi kunne også se om det er tegn til større forskjell i hva terninger med ulike farger gir når det er kvinner som kaster dem, som ville være en toveis-interaksjon mellom kjønn og terningfarge. Hvis terningene ikke er tuklet med ville modellen finne at forskjellene i hva man får når terninger med ulik farge kastes av ulike kjønn varierer – men bare på det nivået man ville forvente hvis det var tilfeldigheter som lå bak. Anta nå at vi har bare fire observasjoner fra menn som kastet brune terninger, og at alle disse fire kastene ga seks. En mettet OLS ville da anslå at menn som hiver brune terninger forventningsmessig vil få seks (punktestimert). Den hierarkiske modellen ville tatt inn over seg at det ikke var noen tegn til genuin parametervariasjon langs kjønn og fargekombinasjoner, og ville predikere at menn som kastet brune terninger har samme sannsynligheter som andre når de kaster terningen sin.

Merk at i vår modell behandler vi tredje-ordens effektene ulikt: For de direkte og annen-ordens parameterne anslår vi hvor mye genuin variasjon det ser ut til å være langs disse dimensjonene, men for tredjeordenseffektene pålegger vi en fast prior som ikke påvirkes av data. Grunnen til det er at disse tredjeordenseffektene skal være null i forventning overalt bortsett fra i de behandlede cellene der behandlingen slår ut. Hvis vi anslår hvor mye genuin variasjon det er på dette nivået ville det være mange ubehandlede celler der det var rundt null, og vi ønsker ikke at modellen skal bruke mangelen på reformeffekter i pre-reform perioden til å dempe (potensielt svake) signaler på behandlingseffekt i det mindretallet av celler som faktisk var behandlet. En alternativ løsning ville være å «fortelle» modellen at de behandlede cellene er annerledes og tatt fra en annen fordeling. Problemet med dette er at vi da ikke lenger behandler alle cellene likt: Vi ville informert modellen om at vi forventet noe annet i den ene delen av datamaterialet enn i de andre, og dette ville forstyrre tolkningen av event-plot som sammenlikner utfall før («pre») og etter («post») «behandlingen». Modellen ville nå være mer skeptisk til indikasjoner på tredjeordens-interaksjoner i kontroll-celler enn i behandlede celler, og det ville være en risiko for at pre-trendene ser uproblematisk ut (og annerledes enn post-trender) fordi vi har bedt modellen vise mer skepsis til signaler på avvik i pre-reform perioden og ubehandlede alders/inntektsgrupper. Dette er uønsket – siden vi ønsker oss en statistisk analyse av data som er «blindet» - i den forstand at alle celler behandles symmetrisk og likt. Modellen som predikerer celle-sannsynligheter får derfor ingen informasjon om hvilke celler som ble behandlet – den informasjonen trekkes bare inn når vi skal beregne DiD og DiDiD anslagene på bakgrunn av de estimerte celle-sannsynlighetene.

Siden signalet i data er forventet å være svakt, vil anslagene forventes å variere med hvor «skeptisk» vi ber modellen være. Vi estimerte derfor modellen separat i fire ulike varianter som varierer etter hvor skeptiske de er mot svake indikasjoner på store effekter. Den mest skeptiske (laveste) tilsier at vi tror plausible reformeffekter vil ligge innenfor en halvering/dobling av risikoen for AAP-tilstrømming, mens den minst konservative (mest følsomme) legger til grunn at reduksjonen i minstebeløp er et sted mellom en desimering/tidobling.

Modellen er Bayesiansk, som innebærer at hver parameter gis en «prior-fordeling» - en antagelse om hva slags verdier som anses som plausible eller mulige selv før vi ser data. Modellen oppdaterer disse i lys av observasjonene, slik at parameterverdier som er mer forenlige med observasjonene får større sannsynlighetsvekt mens de som er uforenlige med observasjonene tones ned. Dette gir oss posteriorfordelingen, som sier oss hvor sannsynlig ulike parameterverdier er gitt det vi ser i data, et oppdatert anslag som har lært fra de observerte dataene.

Strengt tatt har ikke en slik Bayesiansk modell et punktestimert, siden den anslår en sannsynlighetsfordeling over mulige parameterverdier. Vi bruker gjennomsnittsanslaget som punktestimert, og oppgir «kredibilitetsintervall» (KI) som dekker 95% av posteriorfordelingen. Tolkningen av et slikt intervall er at det dekker den sanne verdien med 95% sannsynlighet, betinget på modellens antagelser.

I praksis vil slike kredibilitetsintervall brukes på lignende måte som konfidensintervall fra ikke-Bayesianske modeller, selv om de strengt tatt har ulike tolkninger. Et konfidensintervall er et intervall som beregnes basert på data, og dersom vi trekker mange uavhengige datasett, vil vi få mange uavhengige konfidensintervall. I det lange løp vil 95% av disse dekke den sanne verdien, men dette betyr likevel noe annet enn at det er «95% sannsynlighet for at den sanne verdien er i dette intervallet vi beregnet nå». Et eksempel vil være at du kjøper en ny terning og får seks på hvert av de tre første kastene. Under H_0 (rettferdig terning) vil dette kun skje rundt en halv prosent av de gangene du kaster tre kast. Vi kan dermed forkaste H_0 , og konfidensintervallet for sekser-sannsynligheten vil utelukke $1/6$. En Bayesiansk analyse ville legge til grunn at vi har en rimelig forventning om at terningens sekser-sannsynlighet trolig er i nærheten av $1/6$, og vil oppdatere denne forventningen i lys av data. Hvis vi i utgangspunktet mente det var rundt $2/3$ sannsynlighet for at seksersjansen var omtrent $1/6$ vil modellen si det fortsatt er rundt 60% sannsynlighet for dette etter tre kast².

Modellen ble spesifisert og estimert med det statistiske programmeringsspråket Stan (Carpenter et al. 2016). Estimeringsresultatet kommer i form av (forventningsmessig) representative trekninger fra posterior-fordelingen. For hver slik trekning kan vi beregne hver av de tre anslagene våre (alders-DiD, inntekts-DiD og DiDiD), og i sum over alle trekningen får vi dermed «automatisk» en fordeling av den statistiske usikkerheten rundt disse anslagene.

Vår fremgangsmåte er slik sett en slags «to-steps» variant der vi estimerer sannsynligheter på cellednivå og deretter beregner DiD og DiDiD anslagene basert på modell-estimatene.

Implementering av analyse-strategi – den «nye» DiD-litteraturen

Strategier for kausal identifikasjon kan implementeres på ulike måter, og den praktiske implementeringen har noen ganger løpt i forkant av hva den statistiske litteraturen ga grunnlag for.

Et eksempel på dette er standardfeil, som må ta hensyn til at analysen sammenligner *enheter* (f.eks. kommuner) som består av mange *individer*. Dette betyr at ulike observasjoner bidrar på ulike måter: Nye observasjoner fra kommuner som allerede er dekket i data *reduserer målestøy* ved å la oss anslå utfall i disse kommunene mer presist. Nye observasjoner fra nye kommuner *styrker inferensgrunnlaget* ved å øke antallet enheter i kontroll- og treatment-gruppa, siden det er på kommunenivå vi egentlig analyserer reformeffekter. I tidlige DiD-arbeider ble ikke denne forskjellen tatt hensyn til, og mange tilsynelatende robuste funn viste seg å ikke holde (Bertrand, Duflo, and Mullainathan 2004).

Et annet eksempel fra senere år gjelder antagelsene om hvordan behandlingseffekten slår ut: Den vanlige implementeringen av DiD-modeller la implisitt til grunn at reformeffekten var umiddelbar og permanent, slik at man bare avvek fra fellestrenden i en enkelt periode før man igjen fulgte samme utvikling som før. Metoden brukte derfor også tidsutviklingen i behandlede enheter for å anslå felles-trenden. Dette blir feil dersom behandlingen påvirker *tidsutviklingen*, for eksempel ved at behandlingen gradvis får større effekt ettersom et tiltak får satt seg, eller gradvis får lavere effekt ettersom given og entusiasmen fra innføringstidspunktet falmer. (Goodman-Bacon 2021). Når vi har data på en gradvis utrullet reform, vil modellen i slutten av dataperioden estimere den kontrafaktiske trenden i stor grad basert på *allerede behandlede enheter*.

Slike problem reflekterer ikke DiD-strategien som sådan, men implementeringen av den, og i etterkant av at disse problemene ble kjent har en rekke alternative implementeringer blitt utviklet og tatt i bruk (Roth et al. 2022).

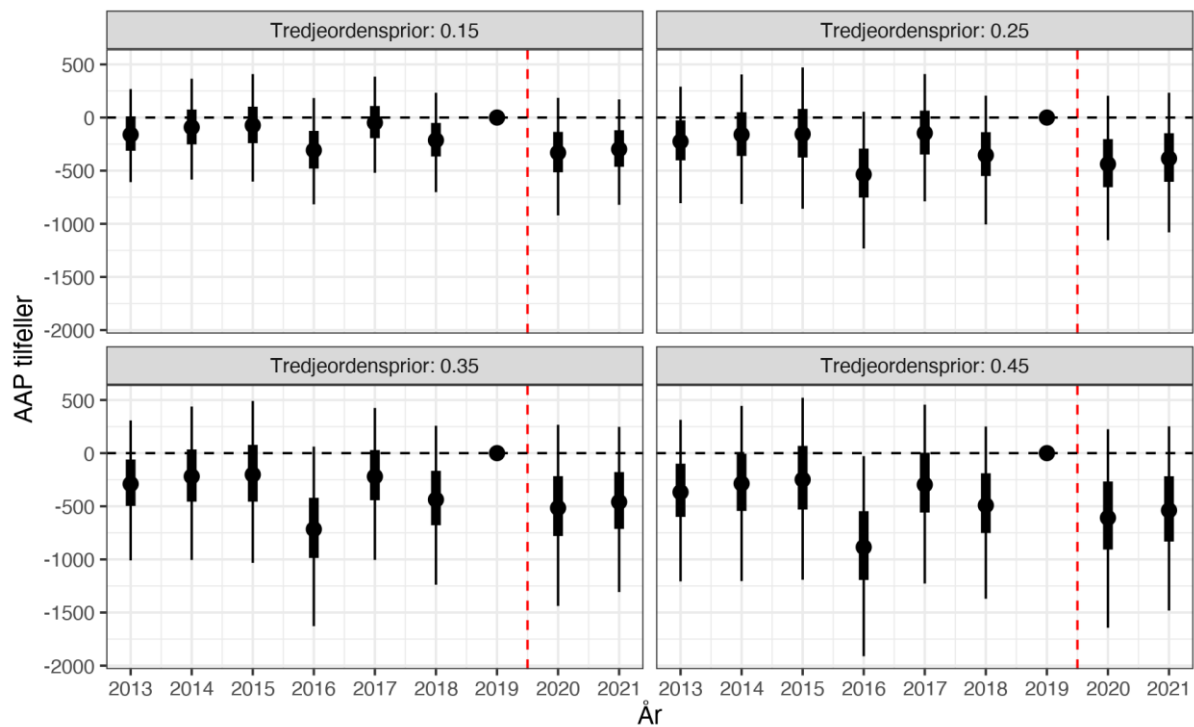
² Bayesiansk analyse modellerte terningkast som binomial-trekning og satte på en $\text{beta}(10,50)$ prior som tilsier rundt 73% sannsynlighet for at seksersjansen er pluss/minus fem prosentpoeng av $1/6$.

Litteraturen har også klargjort kompleksiteten i antagelsene som ligger bak DiD-analyser der kontroll- og behandlings-enheter skiller seg kun i hvilken *grad* de er behandlet (Callaway, Goodman-Bacon, and Sant’Anna 2021), og problematisert bruken av ytterligere kontrollvariable i DiD-analyser (Caetano et al. 2022). Den har også tydeligere løftet frem at antagelsen om en felles trend vil være spesifikt til en skala: Hvis behandlede og ubehandlede enheter har ulikt nivå på utfallet i utgangspunktet, for eksempel, så får vi ulik kontrafaktisk estimat for de behandlede enhetene avhengig av skalaen vi bruker: Vokser alle enhetene med 10% i året har vi en felles trend på log-skala, men da vil en enhet med 100 vokse med 10 og en enhet med 50 vokse med 5 så trenden blir ikke lik med absolutte tall (Roth and Sant’Anna 2023).

En del av disse utfordringene er knyttet til situasjonen med *gradvis innføring* av behandling over enhetene. I vårt prosjekt er regelendringene i all hovedsak innført på ett og samme tidspunkt for alle behandlingsgrupper. Det problemet som derimot vedvarer, er problemene knyttet til test av felles-trend antagelsen i pre-reform-data.

Som tidligere nevnt hviler DiD-analyser på en antagelse om at behandlede og kontroll-enheter har en felles «ubehandlet trend» *i den perioden de behandlede enhetene ble behandlet*. Dette kan ikke ettergås i data, siden de behandlede nødvendigvis var behandlet. Derimot vil denne antagelsen fremstå mer plausibel og sannsynlig dersom kontroll- og treatment-enheter utviklet seg likt frem til reformtidspunktet: Selv om det ikke er noen logisk nødvendighet virker det intuitivt rimelig å forvente at to enheter som lenge har fulgt hverandre i utvikling trolig vil fortsette å gjøre det også fremover.

Slike «pre-trend»-analyser er viktige og kan være nyttige, men i praksis er de ofte upresise fordi forskere bruker færre årganger med data i pre-perioden enn i post-reform perioden (Roth 2022). Samme arbeid påpeker også at hvis vi forkaster alle DiD-analyser som feiler denne testen vil vi få et selektert utvalg av analyser som blir publisert. Fra de analyse-situasjonene der det *var* forskjeller i trender, vil vi nå kun beholde de analysene der tilfeldighetene tilslørte disse trendforskjellene i pre-perioden. I den lengre post-perioden vil disse trendene etterhvert vise seg, og vi får dermed en skjevhet (bias) og misvisende resultater i den publiserte litteraturen.



Figur 20 – Event-plott for DiDiD modell med 2019 som referanseår.

Note: Anslag angis med tykk strek for 50% kredibilitetsintervall (KI) og tynn strek for 95% KI.

Publikasjoner fra Frischsenteret

Alle publikasjoner er tilgjengelig i Pdf-format på : <https://www.frisch.uio.no/publikasjoner/>

Rapporter

| | | |
|--------|---|--|
| 1/2011 | Yrkesdeltaking på lang sikt blant ulike innvandregrupper i Norge | Bernt Bratsberg, Knut Røed, Oddbjørn Raaum |
| 1/2012 | NAV-refomen: Flere i arbeid – færre på trygd? | Ragnhild Schreiner |
| 2/2012 | Privatization of the absenteeism scheme: Experiences from the Netherlands | Julia van den Bemd, Wolter Hassink |
| 1/2013 | Til, fra og mellom inntektssikringsordninger – før og etter NAV | Elisabeth Fevang, Simen Markussen, Knut Røed |
| 2/2013 | Sluttrapport fra strategisk instituttprogram om pensjonsforskning 2007-2012 | Erik Hernæs |
| 1/2014 | Produktivitetsutviklingen etter NAV-reformen | Sverre A.C. Kittelsen, Finn R. Førund |
| 2/2014 | Sysselsetting blant funksjonshemmede | Ragnhild C. Schreiner, Simen Markussen, Knut Røed |
| 3/2014 | Produktivitetsanalyse av Universitets- og Høgskolesektoren 2004 – 2013. | Dag Fjeld Edvardsen, Finn R. Førund, Sverre A. C. Kittelsen |
| 1/2015 | Kan kjønnsforskjellen i sykefravær forklares av holdninger, normer og preferanser? | Karen Hauge, Simen Markussen, Oddbjørn Raaum, Marte Ulvestad |
| 2/2015 | Effekter av arbeidspraksis i ordinær virksomhet: Multiple og sekvensielle tiltak | Tao Zhang |
| 1/2016 | Kompensasjonsgrader i inntektssikringssystemet for personer med svak tilknytning til arbeidsmarkedet | Øystein Hernæs, Simen Markussen, Knut Røed |
| 2/2016 | Bevegelser inn i, mellom og ut av NAVs ytelser | Elisabeth Fevang, Simen Markussen, Knut Røed, Trond Christian Vigtel |
| 1/2017 | Yrkesaktivitet og pensjonsuttak etter pensjonsreformen | Erik Hernæs |
| 2/2017 | Pensjonsordninger og mobilitet | Erik Hernæs |
| 1/2018 | Virkninger av endringer i permitteringsregelverket Delrapport 1, utarbeidet av Oslo Economics og Frischsenteret på oppdrag fra Arbeids- og sosialdepartementet | Ragnhild Haugli Bråten, Aleksander Bråthen, Nina Skrove Falch, Knut Røed |

| | | |
|--------|---|---|
| 2/2018 | Virkninger av endringer i permitteringsregelverket Delrapport 2, utarbeidet av Oslo Economics og Frischsenteret på oppdrag fra Arbeids- og sosialdepartementet | Nina Skrove Falch, Knut Røed, Tao Zhang |
| 3/2018 | New Insights from the Canonical Fisheries Model | Eric Nævdal and Anders Skonhoft |
| 4/2018 | Arbeids- og velferdsetatens arbeid med langtidsledige | Nina Skrove Falch, Ragnhild Cecilie Haugen, Magnus Våge Knutsen, Knut Røed |
| 1/2019 | Effektivitets- og produktivitetsanalyse av norske tingretter | Finn R. Førsum, Sverre A.C. Kittelsen |
| 1/2020 | Gråsoner i arbeidsmarkedet og størrelsen på arbeidskraftreserven | Elisabeth Fevang, Simen Markussen, Knut Røed |
| 2/2020 | Framework for Optimal Production and Transmission of Electricity | Finn R. Førsum |
| 3/2020 | Kvantitativ beskrivelse av institusjonspopulasjonen | Nina Drange, Øystein M. Hernæs |
| 1/2021 | Early intervention in temporary DI: A randomized natural field experiment reducing waiting time in vocational rehabilitation | Karen Evelyn Hauge, Simen Markussen |
| 2/2021 | Supported Employment eller «vanlig» oppfølging? Resultater fra et stort randomisert forsøk i NAV | Helene Berg, Karen E. Hauge, Simen Markussen, Tao Zhang |
| 3/2021 | Delrapport 2: Kvantitativ evaluering av innføring av plikt for kommunene til å stille vilkår om aktivitet til sosialhjelpsmottakere under 30 år | Øystein M. Hernæs |
| 4/2021 | Rapport Delprosjekt 1: Beskrivende analyser – Barn og familier i barnevernet | Nina Drange, Øystein M. Hernæs, Simen Markussen, Inger Oterholm, Oddbjørn Raaum, Tor Slettebø |
| 5/2021 | Underholdskrav, familieinnvandring og integrering | Bernt Bratsberg, Oddbjørn Raaum |
| 1/2022 | The Social Gradient in Employment Loss during COVID-19 | Annette Alstadsæter, Bernt Bratsberg, Simen Markussen, Oddbjørn Raaum, Knut Røed |
| 2/2022 | Barn, unge og familier i barnevernet – En longitudinell registerstudie. Delprosjekt 2: Hvordan går det med barna? | Nina Drange, Øystein M. Hernæs, Simen Markussen, Inger Oterholm, Oddbjørn Raaum, Tor Slettebø |
| 3/2022 | Indikatorer for å måle omstillingstempo i norsk økonomi | Elisabeth T. Isaksen, Knut Røed, Tao Zhang |

| | | |
|--------|---|--|
| 1/2023 | Yrkesdeltakelse, inntekt og trygd blant ansatte i stillinger med særaldersgrense | Asbjørn Goul Andersen, Simen Markussen |
| 2/2023 | Utvikling av et forsøk med stipend til fagarbeidere. Rapport fra modellutviklings-fasen (fase 1) | Karen Hauge, Nina Drange, Simen Markussen, Ester Bøckmann, Bjorn Dapi, Anna Hagen Tønder |
| 3/2023 | Forsøk med stipend til fagarbeidere. Rapport fra pilot-fasen (fase 2) | Karen Hauge, Nina Drange, Simen Markussen, Ester Bøckmann, Bjorn Dapi, Anna Hagen Tønder |
| 4/2023 | Forsøk med stipend til fagarbeidere. Rapport fra hovedforsøket (fase 3) | Nina Drange, Karen Hauge, Tao Zhang, Markus Roos Breines, Bjorn Dapi, Anna Hagen Tønder |
| 5/2023 | Effektanalyse av redusert minstesats for mottakere av arbeidsavklaringspenger (AAP) under 25 år og avvikling av ung ufør-tillegget | Hernæs, Øystein M., Ole Røgeberg, Helene Berg, Gro Malene Vestøl |

Arbeidsnotater

| | | |
|--------|---|--|
| 1/2011 | Job changes, wage changes, and pension portability | Erik Hernæs, John Piggott, Ola L. Vestad, Tao Zhang |
| 2/2011 | Sickness and the Labour Market | John Treble |
| 1/2012 | Dummy-encoding Inherently Collinear Variables | Simen Gaure |
| 2/2012 | A Faster Algorithm for Computing the Conditional Logit Likelihood | Simen Gaure |
| 3/2012 | Do medical doctors respond to economic Incentives? | Leif Andreassen, Maria Laura Di Tommaso, Steinar Strøm |
| 1/2013 | Pension systems and labour supply – review of the recent economic literature | Erik Hernæs |
| 1/2016 | Occupational crosswalk, data and language requirements | Maria B. Hoen |
| 1/2022 | Dokumentasjon av data fra undersøkelse om turnusordninger i kommunenes helse- og omsorgstjenester Spørreundersøkelse sendt til ledere ved virksomheter i helse- og omsorgssektoren | Vilde Hoff Bernstrøm, Elisabeth Fevang, Heidi Gautun, Mari Holm Ingelsrud, Andreas Lillebråten, Otto Lillebø |



Frischsenteret

Stiftelsen Frischsenteret for samfunnsøkonomisk forskning er en uavhengig stiftelse opprettet av Universitetet i Oslo. Frischsenteret utfører samfunnsøkonomisk forskning i samarbeid med Økonomisk institutt ved Universitetet i Oslo. Forskningsprosjektene er i hovedsak finansiert av Norges forskningsråd, departementer og internasjonale organisasjoner. De fleste prosjektene utføres i samarbeid mellom Frischsenteret og forskere ved andre norske og utenlandske forskningsinstitusjoner.

**Frischsenteret
Gaustadalléen 21
0349 Oslo
frisch@frisch.uio.no
www.frisch.uio.no**