

Forsøk med Arbeidstidsreformer/Sekstimers-dag

Den mest hensiktsmessige utformingen av forsøk med arbeidstidsreformer avhenger av hva man ønsker å finne ut og av hvordan eventuelle funn tenkes brukt i framtida. Forsøk innrettet på å gi kunnskap om individers egen verdsetting av økt fritid vil ha en annen utforming enn forsøk innrettet på å gi kunnskap om hvordan redusert arbeidstid påvirker produktivitet og sykefravær. Dersom man ønsker generaliserbar kunnskap må utvelgelsen av forskningsenheter være randomisert (fra den populasjon man ønsker å trekke generelle slutninger om). Forsøk vil i liten grad kunne bringe fram ny kunnskap om såkalte generelle likevektseffekter, dvs. hvordan redusert normalarbeidsdag vil påvirke økonomien som helhet (i form av sysselsetting, arbeidsledighet, lønnsnivå etc.)

1. Innledning

Frischsenteret er bedt av Moderniseringsdepartementet om ”..innspill ..som beskriver hva som kan være viktige elementer vi bør være oppmerksomme på knyttet til design, metode mm knyttet til forsøkene [med arbeidstidsreformer, vårt tillegg] og til evalueringen av disse”.

Notatet er ført i pennen av seniorforskerne Oddbjørn Raaum og Knut Røed, men er basert på innspill og synspunkter fra en rekke personer ved Frischsenteret og Økonomisk institutt, Universitetet i Oslo. Dersom departementet har behov for utdypninger eller klargjøring på enkelte punkter, ber vi om det tas kontakt.

Det er en omfattende internasjonal litteratur om sosiale forsøk. Likevel er det gjennomført forbausende få slike forsøk, sett på bakgrunn av den massive satsingen på samfunnsforskning vi har hatt i Norge og andre land. På flere områder er det grunn til å tro at forsøksbaserte data vil kunne gi oss et godt utgangspunkt for identifikasjon av effekter av politikk-endringer. De fleste sosiale forsøk (eksperimenter) har imidlertid en langt enklere struktur enn hva tilfellet vil være på arbeidstidsområdet. Oftest er fokuset på dosering av offentlige tiltak (”kvalifisering av arbeidsledige”), reguleringer (”fartsgrense eller skilting i trafikken”) eller ressurser (”klassestørrelse og lærertetthet”). I sosiale eksperimenter av en viss skala blir enkeltindivider tilordnet ulike miljø eller rammebetingelser på en tilfeldig (randomisert) måte. Kunnskap om effekter kan deretter avledes fra variasjon i utfall mellom individer eksponert for ulike tiltak, reguleringer eller ressurser. Også i vårt tilfellet vil forsøk eller eksperimenter endre rammebetingelsene for arbeidstakere og virksomheter, men de vil dessuten gripe inn i individuelle og kollektive avtaler (etablerte rettigheter) mellom arbeidstakere og arbeidsgivere. Dessuten er ambisjonene for hvilke typer effekter en skal identifisere antakelig betydelige høyere.

Det er liten tvil om arbeidstidsforsøk med god design og gjennomføring kan gi verdifull kunnskap som grunnlag for endringer i arbeidstidsregimet i Norge. Gitt de praktiske, institusjonelle og økonomiske begrensninger som trolig vil finnes, varierer troen på både gjennomførbarhet og kunnskapsutbyttet fra forsøk/eksperimenter på arbeidstidsområdet mellom personer som har bidratt til dette notatet. Personer i arbeidsgruppen har ulike erfaringer med forsøk/eksperimenter, herunder i arbeidsmarkedspolitikken og ved arbeidstidsreformer fra noen år tilbake. I vurderingen av mulige forsøk er det viktig å både

tenke prinsipielt/forskningsmessig når det gjelder alternative design, samtidig som hensynet til gjennomførbarhet vil måtte veie tungt.

Eventuelle forsøk må avspeile arbeidstidsreformer (politikkalternativer) som kan være relevante å gjennomføre. Generelt vil reformer innebære at

- a. valgmulighetene for den enkelte arbeidstaker (i form av realiserbare kombinasjoner av arbeid, fritid og forbruk) endres, samtidig som
- b. rammebetingelsene for arbeidsgiverne i form av sammenheng mellom avtalt arbeidstid og lønnsutgifter endres og
- c. kontrakter som regulerer forholdet mellom arbeidstakere og arbeidsgivere reforhandles.

Før forsøkene designes bør målsettingene spesifiseres nøye. Målsettingene bør tilfredsstillende to typer krav: For det første må de være empirisk etterprøvbare, slik at graden av målsoppnåelse kan kvantifiseres. For det andre bør de være politikk-relevante, på den måten at det man eventuelt kan finne ut ved hjelp av forsøkene forventes å ha konsekvenser for framtidige utforminger av arbeidstidsreguleringer i Norge. Det er viktig at man på forhånd tenker igjennom hvilke konsekvenser man ser for seg at eventuelle forsøksresultater kan få for den alminnelige utforming av arbeidstidbestemmelser eller lønnsavtaler i det norske arbeidslivet. De arbeidstidsmodeller som testes ut bør med andre ord enten være aktuelle for gjennomføring i større skala (dersom målsettingene ved forsøkene skulle bli innfridd), eller være designet for å skaffe til veie kunnskap (f.eks. knyttet til folks verdsetting av fritid) som anses avgjørende for politikikutformingen på dette feltet.

Departementet bør vurdere en oppfølging med forskningsinstitusjoner i arbeidet med vurdering av alternative forsøksordninger. Etter at målsettinger er spesifisert og aktuelle reformer avklart, bør forskere bli presentert konkrete enkeltforslag og bidra med sitt syn på hva man kan forvente å kunne få av økt kunnskap fra dem. Det bør også være tid til å kunne foreta justeringer av opplegget for best mulig å lykkes i å etablere et bedre kunnskapsgrunnlag.

Det er også grunn til å minne om at det finnes alternative kilder til kunnskap om arbeidstid og utfall på arbeidstakernivå. Forsøksbasert kunnskap bør kunne kompletteres med studier av atferd under eksisterende regime ("kvasi-eksperimenter"). Eksempelvis ligger forholdene godt til rette for registerbaserte empiriske studier av sammenhenger mellom arbeidstid og sykefravær.

2. Eksperimenter og kvasieksperimenter

Svært mye empirisk samfunnsforskning er basert på kvasieksperimenter. Man forsøker da å utnytte eksperimentlignende situasjoner, framkalt av at ulike observasjonsenheter stilles overfor ulike rammebetingelser (f. eks. ulik arbeidstid). Studiene er typisk basert på forutsetninger om at atferden kan analyseres "som om" de ulike betingelsene som "forsøksenhetene" utsettes for hadde vært diktet og ikke selvvalgt, i hvert fall når man har kontrollert for observerte forskjeller mellom dem. En ønsker gjerne å trekke mest mulig almenngyldige slutninger om hva resultatet vil bli av å gjennomføre endringer av rammebetingelsene for aktørene på et visst nivå i samfunnet. Det store spørsmålet er om kvasiforsøket tillater gyldige slutninger, og litteraturen dreier seg om hvilke farer som truer den interne og eksterne gyldigheten. At slutningene fra kvasiforsøket er internt gyldige betyr at en rutinemessig implementering av en gitt regulering for den populasjonen som er observert fører til at populasjonen fungerer som påstått fra studien. Om slutningene også er gyldige for andre populasjoner, snakker en om ekstern validitet.

Data fra kvasieksperimenter analyseres gjerne med en eller annen form for regresjonsanalyse for å estimere effektene av mulige reguleringer. En må da bestemme hva som er den statistiske enheten, hva som er responsvariabelen(e), hvilke forklaringsvariabler som skal bringes inn osv. For å diskutere gyldighet må populasjonen identifiseres. Bare i liten grad gir kvasieksperimentene anledning til empirisk å bestemme grad av gyldighet. En må i stor grad ty til teori og skjønn.

Til forskjell fra kvasiforsøk baserer eksperimenter seg på kontrollert tilordning av rammebetingelser og randomisering. Som i kvasiforsøk må hovedingrediensene (populasjon, enhet, responsvariable, forklaringsvariabler, kontrollvariabler) i den statistiske analysen bestemmes. Forsøksenhetene identifiseres, enten som alle enhetene i populasjonen eller et tilfeldig utvalg. Etter en plan dikteres de ulike betingelsene (verdien på kontrollvariablene) de utvalgte enhetene. Forsøksplanen bestemmes delvis ved randomisering. Hvilke forsøksenheter som blir underlagt hvilken kontroll blir da bestemt gjennom en loddtrekning, hvor sannsynlighetene kan avhenge av enhetenes egenskaper (verdien på deres kovariater). Som for kvasiforsøk blir de eksperimentelle data analysert ved en form for regresjonsanalyser. Muligheten for å lære av denne typen eksperimenter er således avhengig av et relativt stort antall forsøksenheter.

Som for eksperimenter på andre områder (som i klinisk utprøving av nye behandlingsmåter) kan det oppstå problemer med sosiale forsøk. Enheter kan falle ut og bevissthet om at enheten deltar i et forsøk kan prege responsen. Når forsøket er avsluttet kan det derfor oppstå spørsmål om gyldighet på samme måte som for kvasieksperimenter. Det er i praksis en glidende overgang mellom de to formene for studier. Men generelt kan en si at jo mindre kvasi et kvasieksperiment er desto sterkere gyldighet oppnåes. Det er derfor avgjørende å planlegge forsøket så godt som mulig, og å utnytte de mulighetene som byr seg for kontroll og randomisering.

3. Prinsipielle problemstillinger

3.1 *Motivasjon og politikk-alternativer*

Kunnskapsbehovet knyttet til virkninger av arbeidstidsreform med en kortere arbeidsdag eller reformer med mer fleksibelt uttak av kortere arbeidstid, er nært knyttet til hva hovedmotivasjonen for reformen er. Her finnes ulike betraktningsmåter:

a. Preferanser for kortere arbeidstid

Et perspektiv er at kortere normalarbeidstid gir et produksjonstap som vi er villige til å bære (i form av lavere samlet inntekt), fordi vi tror det vil gi en bedre balanse mellom konsum og fritid. Med denne tolkningen er det naturlig å tenke på arbeidstidsreformer som alternativer til fortsatt lønnsvekst. Spørsmålet om kortere arbeidstid er i hovedsak normativt. Implisitt ligger en formening om at mange personer i dag har lengre arbeidstid enn det som er optimalt (enten for individet selv og eller samfunnet). Men vår kunnskap om fritidspreferanser og villigheten til å oppgi lønn for kortere arbeidstid er begrenset, bl.a. fordi reguleringer av arbeidstid hindrer arbeidstakere fritt å realisere sitt primære arbeidstidsønske. Gjennom forsøk der arbeidstakere blir stilt overfor en større meny av alternative arbeidstid/lønn-kombinasjoner kan vi få økt kunnskap om i hvilken utstrekning arbeidstidsreformer uten full lønnskompensasjon er ønskelig sett fra arbeidstakernes side. Den empiriske testing vil her kunne avdekke preferanser for, snarere enn effekter av, arbeidstidsreformer.

b. Økt tilstedeværelse og bedre helse

Kortere normalarbeidsdag vil kunne bedre arbeidstakernes helse og redusere sykefraværet. Produksjons- og innteksttapet for samfunnet blir i så fall ikke proporsjonalt med reduksjon i arbeidstid, og det kan argumenteres for at lavere arbeidstid bør være forbundet med høyere timelønn. Tolket på denne måten er spørsmålet om kortere arbeidstid rent empirisk og omhandler effekter; I hvilken grad påvirkes arbeidstageres helse og tilstedeværelse på jobben av arbeidstidens lengde, og hvordan varierer eventuelle effekter mellom typer virksomheter og grupper arbeidstakere? For tilstedeværelse er det naturlig å måle effektene på individnivå.

c. Økt produktivitet

Det kan hevdes at arbeidstakere jobber mer effektivt den tiden man er på jobb, dersom den daglige (eller ukentlige) arbeidstiden reduseres. Også samspillet mellom arbeidstakere og utnyttelse av eventuelt kapitalutstyr (gjennom skiftordninger) kan tenkes påvirket gjennom endret normalarbeidstid. I likhet med sykefraværeffekter er virkninger for produktivitet empirisk og knyttet til effekter av arbeidstidsreguleringer på produksjonsprosessen. Men så lenge produktivitet ikke kan sees isolert og måles for den enkelte arbeidstaker, må fokus settes på utfall på virksomhetsnivå. Det er ikke tilstrekkelig å se på utfall for enkelte arbeidstakere. Empiri knyttet til sammenhengen mellom arbeidstid og produktivitet må under alle omstendigheter forventes å være spesifikk for den enkelte type arbeidsoperasjon/teknologi. Resultater på dette feltet vil derfor ikke uten videre kunne generaliseres til arbeidslivet som helhet.

d. Økt arbeidstilbud for enkelte grupper arbeidstakere

Redusert normalarbeidsdag kan gjøre det mer attraktivt for personer som i dag arbeider deltid (f.eks. halv stilling) å gå over til fulltidsstilling. I tillegg har det vært anført at redusert normalarbeidsdag kan gjøre det mer attraktivt for eldre arbeidstakere å utsette pensjonsalderen. Dette er eksempler på at visse grupper kan tenkes øke sitt arbeidstilbud som følge av redusert normalarbeidsdag. I så fall vil dette bidra til å motvirke nedgangen i arbeidstilbudet fra dem som reduserer sin arbeidstid. Igjen er dette et empirisk spørsmål.

Dagens fokus på redusert arbeidstid inneholder en blanding av de ulike perspektivene omtalt over. De varmeste tilhengerne av redusert normalarbeidsdag synes å hevde at produksjons- og innteksttapet (antakelig) er lite og neglisjerbart fordi kortere arbeidstid øker tilstedeværelsen og effektiviteten når man er på jobb. La oss anta at individuell verdiskapning pr måned (eller år) kan måles. For en enkelt arbeidstaker vil verdiskapningen dekomponeres i bidrag fra "antall dager til stede på jobben" (som avhenger av sykefravær), og produksjon pr. dag på jobb (som avhenger av effektivitet). Produksjon pr dag vil avhenge av arbeidstid pr. dag, fratrullet faste tidskostnader pr. dag på jobb i form av lunsjpause, oppstart- og avviklingstid. Færre timer på jobb for heltidsarbeidende, og en reduksjon i den andel av arbeidstiden som brukes til aktivt arbeid, vil føre til lavere verdiskapning. Den siste effekten skyldes at det oftest vil være en gitt mengde tid som benyttes til oppstart/avslutning, lunsjpause etc. De negative effektene kan motvirkes av økt tilstedeværelse (mindre sykefravær, utsatt pensjonering) og endring i produktivitet per time i aktivt arbeid. Men et mulig siktemål med redusert normalarbeidsdag kan også være å motivere personer som i dag jobber f.eks. halv tid til å justere sin arbeidstid opp. For denne gruppen vil det da kunne oppnås en økning i andelen effektiv arbeidstid.

For enkelte har ønsket om redusert arbeidstid et mer fordelingspolitisk perspektiv, ved at kortere arbeidstid kan gis som et velferdsgode til utvalgte grupper av befolkningen, f.eks.

småbarnsforeldre. Igjen er dette først og fremst et normativt spørsmål, men empiri kan være relevant for å finne ut hvordan en slik velferdsordning vil bli verdsatt relativt til alternative anvendelser av de samme ressursene.

Forsøk kan gi oss kunnskap som er nyttig ved en eventuell implementering av arbeidstidsreformer, men utformingen av forsøkene avhenger av hvilke av de ovennevnte effekter man ønsker økt kunnskap om. Vårt system for lønnsdannelse, arbeidskontrakter og regulering av arbeidslivet er omfattende. Erfaringer med arbeidstakeres, fagforeningers og virksomheters holdninger til og reaksjoner på variasjonen i rammebetingelser vil være viktig når det skal utarbeides en strategi for hvordan endringer i arbeidstidsreguleringer skal iverksettes.

3.2 Hva er "effekter"?

Enten vi drøfter et forsøk eller en implementert "regulering" (kalt A) er effekt alltid definert *relativt til et alternativ* som her naturlig er status quo (for arbeidstidsregimet, kalt B).¹ Teoretisk er "effekter" av A versus B definert ved forskjeller i utfall mellom de to. I sosiale relasjoner er det *kun mulig å observere en aktørs atferd på et gitt tidspunkt under ett av reguleringsalternativene*. Enhver evaluering må derfor etablere mål på kontrafaktiske utfall for aktørene; "Hva ville aktør nr. j oppnådd på tidspkt t dersom B hadde vært gyldig (istedenfor A)?"

Her finnes to hovedtyper av empiriske strategier (som ideelt sett kan kombineres). Ved "før-etter"-strategien antas aktørene å kunne sammenliknes med seg selv, slik at utfall realisert tidligere under status quo representerer de kontrafaktiske utfallene. Dette fordrer en viss stabilitet når det gjelder status quo, og ikke minst at vi har kontroll på alle tidsvarierende forhold som kan endre utfall over tid. I mange sammenhenger er ikke dette oppfylt, og en velger å benytte "kontrollgrupper" (eller ulike verdier på kontrollvariablene omtalt i kapittel 2 over). Gjennom ideelt sett "tilfeldig" tilordning av ulike rammebetingelser for forskjellige aktører, kan vi tolke forskjell i utfall som effekter av A vs B. Det er viktig å minne om at effektene ofte vil innebære endret tilpasning for individer og virksomheter langs en rekke dimensjoner.

Dersom en lykkes i å etablere eksperimentelle data i større skala etter mønsteret beskrevet i kapittel 2, er man utstyrt med utfall under flere alternative rammebetingelser (og ikke kun A og B). Mulige effekter kunne dermed spores ved å kombinere utfallsinformasjon fra berørte virksomheter og arbeidstakere over tid med verdien på kontrollvariablene (som er fastsatt gjennom forsøket).

3.3 Forsøksenhetene; virksomheter eller enkeltarbeidstakere?

Ethvert forsøk vil innebære endringer for enkelte arbeidstakere, og i den forstand er individer uansett involvert. Spørsmålet her er om forsøksenheten(e) som får tilordnet ulike rammebetingelser er (i) virksomhet(er) der alle enkeltpersoner (med objektive kjennetegn som alder, omsorgsansvar for små barn el. lign.) innen en virksomhet blir likebehandlet, eller om (ii) individer innen samme virksomhet blir tilordnet ulike menyer av arbeidstid og total lønnsinntekt.

Valg av forsøksenhetsnivå bestemmer blant annet typen effekter man kan forvente å få kunnskap om. Kjennskap til produktivitetseffekter kan vanskelig etableres uten at virksomheter benyttes som forsøksenhet.

Det kan være vanskelig å variere rettigheter for like arbeidstakere, spesielt innen samme virksomhet. Fravær av erfaringer med sosiale eksperimenter i Norge skyldes nok i stor grad motstand mot (loddtrekningsbestemt) forskjellsbehandling når det gjelder arbeidstakeres

¹ Så langt tenker vi oss at det er mulig å måle relevante utfall for alle som potensielt blir påvirket.

rettigheter. Variasjon i rettigheter kan i tillegg skape splid innad i arbeidskollegier, påvirke arbeidsdelingen/belastningen innad i virksomheten og derigjennom utfall for arbeidstakere som ikke får sin arbeidstid endret. Dersom kolleger er færre timer på jobb kan arbeidspresset for andre øke.

3.4 Effekter av endrede rammebetingelser eller endret atferd?

En sentralt spørsmål i forsøksdesign er innslag av frivillighet og dermed deltakelse ("compliance"). Sålenge deltakelse ("faktisk ta pillen") ikke er obligatorisk, vil det kunne være en selektert gruppe forsøksenheter som faktisk gjennomfører endringen vi skal måle effektene av. Perfekt randomisering kan således være vanskelig å gjennomføre. Dette reiser, i teorien, to typer problemer. For det første vil ikke de deltagende forsøksenheterne (i gjennomsnittsforsand) nødvendigvis ha like egenskaper, uavhengig av kontrollvariablene. For det andre vil vi ikke identifisere (gjennomsnitt)effekten av en reguleringsendring som gjelder alle (i en gitt målgruppe). Dersom nedsatt arbeidstid har ulike effekter på tvers av forsøksenheter (virksomheter eller arbeidstakere) av en gitt type, blir det potensielt viktig å skille mellom "gjennomsnittseffekter for et obligatorisk tiltak" og "gjennomsnittseffekter for de som ved frivillighet faktisk velger å delta".

Eksempel på virksomhetsnivå: Tenk at avdelinger i en større offentlig virksomhet får anledning til å innføre 10% reduksjon i ukentlig arbeidstid, kombinert med en økning i lønnsbudsjettet på 2%. Anta videre at arbeidstakerne i virksomheten er sikret en økning i timelønnen på 7%, slik at virksomheten må oppnå en "produktivitetsgevinst" på omkring 5% for å gå i balanse. Det er god grunn til å tro at det er virksomheter der konsekvenser for tilstedeværelse og produktivitet på jobben av endret arbeidstid potensielt er store, som vil vise størst interesse og melde seg til å delta i et slikt forsøk. Fordelen med å begrense forsøkene til statlig virksomhet er at de (antakelig) kan dikteres.

Eksempel på individnivå: Anta at arbeidstakere enkeltvis blir tilbudt ulike menyer av arbeidstid-lønn kombinasjoner. Det vil kunne være arbeidstakere med sterke preferanser for fritid, og med størst helseforbedringspotensiale ved redusert arbeidstid, som velger alternativer med kortere arbeidstid. Effektene vi for eksempel observerer på sykefravær og tilstedeværelse kan i så fall ikke enkelt generaliseres. På den annen side vil individnivå være godt egnet om vi vil kartlegge preferanser.

Når forsøk innebærer prosesser der virksomheter, avdelinger og enkeltarbeidstakere selv kan velge mellom ulike kombinasjoner av arbeidstid og lønn per arbeidstidsenhet, blir det viktig å presisere hvilke effekter forsøket faktisk identifiserer, og ikke minst hvorvidt dette samsvarer med den endringen i arbeidstidsregime som en tenker seg gjennomført. Generelt vil forsøk med endringer i rammebetingelser, kombinert med frivillighet, ikke gi oss kunnskap om effekter av en obligatorisk endring gjennomført for alle. Likevel kan vi komme nær i tilfeller der effektene ikke varierer for mye mellom forsøksenheterne (av en gitt type).

3.5 Verdien av kunnskap om effekter via forsøk

Samfunnsforskere er uenige om hva en kan lære av "sosiale eksperimenter" eller forsøk knyttet til arbeidstidsreformer.² De fleste kan nok enes om at forsøk designet og implementert "etter læreboka", jfr. kapittel 2 over, vil gi et godt kunnskapsgrunnlag. Mulige innvendinger mot – eller skepsis til – forsøksdesign ved en studie av arbeidstidsreformer går langs flere dimensjoner og bør vurderes nøye ved utformingen av alternative forsøk. Som tidligere nevnt er utfordringene i dette konkrete tilfellet større enn ved eksperimentell evaluering av offentlig

² Dette gjelder også innad i gruppen hvis diskusjoner er innbakt i dette notatet.

tiltak, ensidige reguleringer eller ressursbruk i tjenesteproduksjon. Aktuelle spørsmål er mange;

- Er det mulig å finjustere valgbetingelsene/reguleringene i samsvar med hva som senere kan/skal implementeres?
- Finnes det økonomiske ressurser og praktiske muligheter til å variere valgbetingelse mellom tilstrekkelig mange forsøksenheter?
- Er atferden hos aktørene den samme som den vil bli under reelle endringer?
- Kan utfallet studeres lenge nok til å kunne forutse de langsiktige effektene?
- Er det i praksis mulig å sikre tilfeldig utvalg av et tilstrekkelig antall enheter (fravær av "randomiseringsskjevhet")?
- Kan forsøkene gi svar på hvilke kausale sammenhenger som ligger bak eventuelle endringer i utfallsmål?
- Er det mulig å generalisere resultatene (for den aktuelle yrkesgruppen eller for arbeidslivet generelt)?

En viktig innvending mot verdien av forsøksbaserte data er fravær av "likevektseffekter". Ved en eventuell innføring av redusert normalarbeidstid i full skala vil endringer i arbeidstid ha indirekte konsekvenser som ikke dukker opp i et avgrenset eksperiment, f.eks. ved at det finner sted endringer i lønnsdannelsen, forholdet mellom tilbud og etterspørsel i arbeidsmarkedet, arbeidsledigheten, balansen i offentlige budsjetter osv.

3.6 Lønnskompensasjonsspørsmålet

Lønn er kompensasjon for redusert fritid og grunnlag for konsum. I de fleste arbeidskontrakter i Norge er det, innenfor normalarbeidsdagen i en gitt jobb, proporsjonalitet mellom arbeidstimer og total lønn. Med hel eller delvis lønnskompensasjon i forhold til dagens lønnsnivå, vil dette prinsippet brytes. Videre kan neppe lønnskompensasjon avtales fritt, på lang sikt. For arbeidsgiverne er lønnen en kostnad og på lang sikt vil det (alle fall i privat sektor) etableres et samsvar mellom verdiskapning og kompensasjon pr tidsenhet (time). Lønnskompensasjon reiser således en rekke spørsmål:

a. Hvis ikke forsøket er ment å medføre en forskyvning av den funksjonelle inntektsfordelingen i samfunnet, må omfanget av lønnskompensasjon være tilpasset de ambisjoner man har knyttet til hva som kan oppnås av eventuelle gevinster i form av økt produktivitet og redusert sykefravær. Forsøket med lønnskompensasjon vil inneholde menyer av arbeidstid og lønn som ikke er lineære i avtalt arbeidstid, men samtidig skal være gjennomførbare i praksis - og i stor skala. Med delvis lønnskompensasjon (dvs. at årsinntekten reduseres) er det grunn til å tro at ikke alle arbeidstakere vil ønske å ta imot tilbudet om redusert arbeidstid.

b. Det er ikke opplagt hvilket alternativ eventuelle effekter skal vurderes i forhold til. Når endringer (forsøk) innbefatter lønnskompensasjon for arbeidstakere og virksomheter er det ikke lenger "status quo" alternativet vi vil sammenlikne med for å identifisere effekter. I eksempelet omtalt i avsnitt 3.4 innebærer forsøket både endret arbeidstid og større lønnsbudsjett. Arbeidstakerne får altså både høyere timelønn og kortere arbeidstid. Hva er det kontrafaktiske vi ønsker å sammenlikne med? Ideelt sett burde forsøk også omfatte virksomheter og arbeidstakere som fikk høyere lønn men uendret arbeidstid. Timelønnen kan ha en selvstendig effekt på f.eks. produktivitet og sykefravær.

c. Hvis forsøket er innrettet på å ta i bruk redusert arbeidstid som et velferdsgode for visse grupper arbeidstakere, selv om dette skulle medføre økte offentlige utgifter, vil det relevante sammenligningsgrunnlaget være den beste alternative anvendelsen av de samme ressursene.

d. I det lange løp bestemmes timelønnen i et arbeidsmarked. Dette skjer gjennom forhandlinger mellom partene i arbeidslivet, under påvirkning av etterspørsel og tilbud i arbeidsmarkedet.

4. Implementering av forsøk

4.1 Effekt av hva?

Det er behov for å presisere hva man skal måle effekter *av* vel så mye som hva man skal måle effekter *på*. Et viktig spørsmål er om man ønsker å identifisere effekten av at *det gis et tilbud om redusert arbeidstid*, eller effekten av at *arbeidstiden faktisk blir redusert*. I praksis er det naturligvis langt mer uproblematisk å randomisere et tilbud om redusert arbeidstid enn det er å randomisere den faktiske arbeidstiden (ettersom det kan være vanskelig å tvinge personer til å redusere arbeidstiden, i hvert fall med mindre de gis *full lønnskompensasjon*). Det er også svært viktig å være oppmerksom på at redusert arbeidstid *med lønnskompensasjon* involverer to tiltak samtidig; a) redusert arbeidstid og b) høyere timelønn. Hvis reformen skulle vise seg å ha effekt på f.eks. produktivitet eller sykefravær, vet vi ikke om effekten kommer av a) eller b) eller av en kombinasjon av a) og b). Ønsker vi å måle den isolerte effekten av redusert arbeidstid, burde dette strengt tatt vært gjort uten en samtidig endring i lønn (men dette ville åpenbart forsterket problemene knyttet til selektiv deltagelse).

4.2 Effekter for hvem?

En arbeidstidsreform vil potensielt ha effekter på mange aktører; de direkte berørte arbeidstakerne, eventuelle kollegaer som ikke omfattes av reformen, virksomhetene, virksomhetenes kunder/brukere, og de som finansierer virksomhetene.

Dersom forsøksenhetene er personer er det viktig å analysere mulige effekter (eksempelvis sykefravær) for de som ikke trekkes ut til å delta, men som (i fall det er et produksjonstap) kan berøres indirekte ved at de får økt arbeidsbelastning (når kollegaene skal jobbe kortere).

4.3 Effekter på hva?

For arbeidstakerne er det naturlig å fokusere på tilstedeværelse på jobb, langvarig sykdom/rehabilitering/attføring/uførhet, tilfredshet, innsats/produktivitet (svært vanskelig å måle på individnivå), turnover, fortsatt yrkesdeltagelse (spesielt for småbarnsforeldre og potensielle pensjonister), omfang av eventuelle bi-jobber, etc. I den grad forsøkene gir mulighet til å bringe fram kunnskap om arbeidstakernes egne preferanser knyttet til valg av arbeidstid, vil dette også være verdifullt.

For virksomhetene vil det være naturlig å fokusere på produktivitet, kostnader, tjenestekvalitet etc. Måling av produktivitet i offentlig virksomhet byr på særlige utfordringer ettersom det i liten grad finnes noe enkelt kvantifiserbart ”produkt”. Forsøk bør her kunne bygge på tidligere studier av verdiskapning i statlig produksjon. Kvalitetsaspektet vil stå sentralt og det er viktig å avdekke eventuelle effekter på tjenestekvalitet av endret arbeidstid (feil i saksbehandling, klager, åpningstider/tilgjengelighet for brukere etc.).

4.4 Måling av utfall

Så langt som mulig bør man utnytte eksisterende administrative rutiner (registre). Dette sikrer effektiv datainnsamling, uten fare for variasjon i tolkning av spørsmål, og med minimal risiko for frafall. Det sikrer også at forsøkspopulasjonen kan sammenlignes med andre. På noen områder vil det imidlertid være behov for å komplettere registerdata med surveydata. Dette gjelder særlig for vurderinger av brukertilfredshet og kvalitet.

Det er svært viktig å etablere datarutiner *i god tid før forsøket gjennomføres*, slik at atferden til forsøkspersoner/forsøksvirksomheter (og eventuelle kontrollgrupper) kan følges både før, under, og etter forsøket.

Rutiner som sikrer mot frafall under datainnsamlingen bør etableres. Forsøk innebærer en særlig bekymring for systematisk frafall.

Det er også grunn til å kvalitetssikre informasjon fra enheter med særlig interesse for utfallet av forsøkene. Det er et velkjent problem at aktører som har egeninteresse av vurderinger basert på informasjon de selv leverer, lett kan fristes til å vri på rapporteringen. Siden målefeil av denne typen neppe kan antas av være tilfeldig bør man være spesielt på vakt.

4.5 Utvelgelsen av forsøksenheter

I den grad man ønsker å iverksette forsøk med høy grad av generaliserbarhet, må valg av forsøksenheter randomiseres, jfr. kapittel 2 over. Grunnlaget for loddtrekningen må være den populasjonen man ønsker å trekke generelle slutninger om. Ønsker man å trekke generelle slutninger knyttet til norske arbeidstakere i sin alminnelighet, burde man med andre ord ha trukket fra hele populasjonen av norske arbeidstakere. Dette er ikke aktuelt for forsøkene med arbeidstidsreformer, ettersom de skal avgrenses til statlig sektor. Forsøkene vil derfor under alle omstendigheter bare i begrenset grad være generaliserbare. Fravær av generaliserbarhet er et særlig problem knyttet til spørsmål som har å gjøre med produktivitetseffekter, ettersom slike effekter må antas å være sterkt teknologiavhengig. Dette kan tale for at man enten tillegger produktivitetsaspektet begrenset vekt i forsøkene, eller klart avgrenser målsettingen med forsøket til å finne ut om et bestemt yrke lar seg utføre mer effektivt med kortere arbeidstid. I så fall bør man trekke fra populasjonen av utøvere innen dette yrket.

Det er viktig å sikre rutiner for korrekt implementering av selve loddtrekningen. Særlig når forsøksenheterne er personer, vil det lett oppstå situasjoner der virksomheter og personer ønsker å manipulere hvem som inngår i "risikogruppen". Det kan både skyldes ønsker om å få spesielle personer inn i forsøket, eller å unngå det. Randomisering kan "saboteres" ved at enheter/enkelt personer "inaktiveres" og fjernes fra gruppen som det skal trekkes fra, jfr. erfaringer med loddtrekning til AMO-kurs på begynnelsen av 1990-tallet.

4.6 Forsøkseffekt på atferd

Arbeidstidsforsøk inneholder en betydelig fare for såkalte Hawthorne-effekter. Forsøkspersoner vil vite at de inngår i et forsøk tilknyttet betydelig politisk oppmerksomhet, og dette kan påvirke deres atferd. Vi kan ikke forvente at virksomheter fullt ut vil opptre slik de ville ha gjort i en normalsituasjon. Tidsperspektivet er antakelig viktig her, siden Hawthorne-effekter vil kunne avta ettersom nye rammebetingelser og arbeidstid oppfattes som en normal arbeidssituasjon. Variasjon i utforming av forsøkene (slik at det ikke er så lett å identifisere hva slags atferd som "forventes" av den enkelte deltager) vil også kunne redusere dette problemet.

4.7 Forhold til arbeidstakerorganisasjoner, rettigheter i tariffavtaler og lovverk

Forsøket bør ligne mest mulig på en "normalsituasjon". Det er således viktig å unngå turbulens, negative indirekte effekter på motivasjon og samarbeidsproblemer under forsøket.

Endringene som forsøket representerer bør derfor ha aksept hos de ansatte og deres representanter. En slik klarering representerer samtidig en begrensning, idet kunnskapen fra forsøket blir begrenset til endringer som er godtatt av fagforeningene. Avveining mellom gjennomførbarhet og relevans kan innebære at forsøkene i praksis avviker fra den ideelle design.

4.8 Antallet virksomheter og arbeidstakere omfattet av forsøket

Enten kontrafaktiske utfall representeres ved egen historie ("før-etter") eller ved andre aktører utsatt for andre kontrollvariabler ("kontrollgruppe"), er det nødvendig med et betydelig antall enheter for å kunne avvise at eventuelle forskjeller skyldes tilfeldigheter. Det skal mange enheter til for å statistisk kunne identifisere små effekter. Det konkrete antallet er situasjonsavhengig.

4.9 Tidsperspektivet

Forsøket bør gå over flere år, slik at forsøksenhetene får tilstrekkelig tid til å tilpasse seg permanente endringer i rammebetingelser.

5. Mulige forsøk

For å eksemplifisere en del av de tingene som er tatt opp i dette notatet, antyder vi to alternative (komplementære) eksperimenter, innrettet mot å skaffe økt kunnskap om henholdsvis arbeidstidspreferanser og tilstedeværelse/produktivitet.

5.1 Mulig forsøk for å avdekke arbeidstakernes egne preferanser

Et stort problem innenfor arbeidsmarkedsforskningen er at det er vanskelig å finne ut hvordan individer ønsker å tilpasse sin arbeidstid. Årsaken til dette er at arbeidstidsmønsteret er sterkt institusjonelt bestemt, og mange står overfor restriksjoner i valg av arbeidstid. Individers preferanser kommer derfor ikke direkte til uttrykk i observerte data. Et eventuelt forsøk med arbeidstidsreformer vil på dette feltet kunne bidra til å bringe arbeidsmarkedsforskningen et substansielt skritt videre. Det vil være rimelig å anta at den vekt man ønsker å legge på arbeidstidsrestriksjoner i framtida vil avhenge av befolkningens egne arbeidstidspreferanser.

Et forsøk siktet mot å avdekke personers reelle arbeidstidspreferanser vil måtte ha som en grunnleggende premiss at personer som blir trukket ut til å delta *fritt får velge sin egen arbeidstid* (eventuelt innen visse grenser), gitt et sett lønnsbetingelser (og for en nærmere avgrenset tidsperiode). Man kunne da tenke seg at for en del av forsøkspersonene er fritt valg av arbeidstid den eneste endringen de blir stillet overfor, dvs. at arbeidsstiden velges på basis av den timelønnen de allerede har. Tilpasningen til denne gruppen vil ha stor interesse i seg selv. Ikke minst vil det være av interesse å kartlegge hvilke faktorer som er viktige for individers valg av arbeidstid, slik som alder, utdanning, familiesituasjon etc. I tillegg kunne man tenke seg at det ble trukket grupper av forsøkspersoner som ble stillet overfor andre typer lønnskontrakter, der timelønnen avhenger av arbeidstiden. Et nærliggende eksempel er at en gruppe arbeidstakere tilbys en kontrakt som gir noe høyere timelønn ved valg av seks timers arbeidsdag. Denne type eksperimenter vil kunne gi kunnskap om hvor stor vekt ulike grupper av personer legger på redusert arbeidstid, og mulighet for å besvare spørsmål av typen "hvor stor kompensasjon vil man kreve for å gå over fra 7,5 til 6 timers arbeidsdag".

I prinsippet kan et forsøk av denne typen gjennomføres ved å foreta en tilfeldig trekning av individer (fra den aktuelle populasjonen av statsansatte). Men fare for lokale reaksjoner på "forskjellsbehandling" og eventuelle administrative komplikasjoner kan muligens tale for at forsøket avgrenset til visse typer virksomheter (og eventuelt omfatter alle de ansatte på de virksomhetene som trekkes ut).

5.2 Mulig forsøk for å avdekke effekter på sykefravær, yrkesdeltagelse og produktivitet

For å kunne kartlegge sammenhengen mellom arbeidstid og sykefravær/produktivitet er det naturligvis essensielt at det faktisk foretas en justering av arbeidstiden, og at denne justeringen i minst mulig grad styres av individenes egne valg. Siktemålet bør være å finne ut om (og eventuelt i hvilken grad) redusert normalarbeidstid vil medføre økt tilstedeværelse og høyere produktivitet, slik at arbeidstidsreduksjonen i noen grad blir selvfinansierende. I tillegg vil det være av interesse å finne ut om redusert normalarbeidstid vil føre til endringer i valg mellom deltid og heltid, og i potensielle pensjonisters vilje til fortsatt å arbeide.

Siden det ikke er mulig å frambringe almenyldig kunnskap om dette temaet, antar vi at forsøket bør avgrenses til en veldefinert "produksjonsprosess" som har et såpass stor omfang at kunnskap om sammenhengen mellom arbeidstid og produktivitet vil være av samfunnsmessig interesse. Og ettersom produktivitet bare kan måles på virksomhetsnivå, bør det være virksomheter (og ikke individer) som trekkes ut. Som eksempel kan vi tenke oss at et forsøk avgrenses til populasjonen av nye arbeids- og velferdsforvaltningskontorer (NAV-kontorer).

Et mulig forsøk kan bestå i at et antall NAV-kontorer, som trekkes tilfeldig, gis pålegg om å innføre redusert normalarbeidstid, samtidig som timelønnen økes noe for dem som velger den nye normalarbeidsdagen (vi antar at det bare ved full lønnskompensasjon kan være snakk om å tvinge noen til å redusere sin arbeidstid). Tilsvarende økes budsjettrammen for det aktuelle NAV-kontor noe, men mindre enn det reduksjonen i arbeidstid isolert sett skulle tilsi. Differansen mellom den reduserte arbeidstiden og budsjettjusteringen definerer en implisitt målsetting om produktivitetsgevinst. Et viktig siktemål med evalueringen blir da å kartlegge om denne målsettingen nås eller ikke, i tillegg til mer direkte analyser av sykefravær, tjenesteproduksjon, brukertilfredshet mv. Forsøket kan eventuelt designes slik at ulike kontorer stilles overfor ulike målsettinger.

Et annet mulig forsøk kan bestå i at kontorer som trekkes ut gir tilbud om redusert arbeidstid (med hel eller delvis lønnskompensasjon) til visse grupper arbeidstakere, f.eks. småbarnsforeldre og eldre arbeidstakere. Siktemålet her vil sannsynligvis være å oppnå økt stabilitet og deltagelse blant denne type arbeidstakere (hindre uttrekning fra arbeidsstyrken), og ikke primært økt produktivitet. Det er derfor rimelig at kontorenes budsjetter justeres fullt ut for de økte timelønnskostnadene i dette tilfellet. Evalueringen vil da sikte mot å finne ut om tiltaket faktisk bidro til økt yrkesdeltagelse for målgruppen.

Det er viktig å være oppmerksom på at utsagnskraftige evalueringer fordrer at denne type eksperimenter har et visst omfang for at man skal kunne skille mellom reelle effekter og tilfeldige variasjoner. Det vil neppe være hensiktsmessig å gjennomføre forsøk med mindre enn 100 forsøksenheter.

6. Effektstudier basert på atferd under eksisterende regime kan supplere kunnskap fra forsøk

Eksperimenter har naturligvis mest å tilføre på områder der det ikke er mulig å skaffe til veie den aktuelle kunnskap uten eksperimenter. En del av de effektene som er drøftet i dette notatet kan alternativt identifiseres og estimeres med basis i eksisterende datakilder. Dette betyr ikke at eksperimenter er overflødige, ettersom det på de områdene vi har diskutert her neppe vil kunne framskaffes empiri som ikke er beheftet med en eller annen form for usikkerhet. Men det betyr at noen av de temaene som ønskes belyst gjennom arbeidstidseksperimenter også kan belyses med basis i eksisterende data. Dette gjelder for eksempel sammenhengen mellom arbeidstid og sykefravær. Ettersom det som kjent allerede er en betydelig variasjon i befolkningens arbeidstid, finnes det et godt grunnlag for å

analysere slike sammenhenger med mer kvasieksperimentelle metoder (jfr. avsnitt 2 foran). Informasjon om individers arbeidstid, lønn, familiesituasjon, kvalifikasjoner og virksomhetstilknytning kan kombineres med data om sykefravær. Det eksisterer allerede et register for alt legemeldt sykefravær på individnivå som kan kobles mot en rekke andre administrative registre.

På den annen side finnes det kunnskap som bare kan skaffes til veie gjennom eksperimenter. Dette gjelder f.eks. kunnskap om individers arbeidstidspreferanser (jfr. avsnitt 5.2), ettersom det ikke finnes noe data som beskriver hva folk ville ha gjort dersom de kunne valgt arbeidstid fritt. Det gjelder også eventuelle responser på ikke-lineære lønnskontrakter (dvs. at timelønnen avhenger av arbeidstiden), herunder mulige effekter på yrkesdeltagelsen blant småbarnsforeldre og eldre.