



A-ordningen i Landbrukstellingen 2020

TALL

SOM FORTELLER

NOTATER / DOCUMENTS

2020 / 30

Tatsiana Pekarskaya og Leiv Solheim

I serien Notater publiseres dokumentasjon, metodebeskrivelser, modellbeskrivelser og standarder.

© Statistisk sentralbyrå
Ved bruk av materiale fra denne publikasjonen
skal Statistisk sentralbyrå oppgis som kilde.

Publisert 28. september 2020

ISBN 978-82-587-1183-1 (elektronisk)
ISSN 2535-7271 (elektronisk)

Standardtegn i tabeller	Symbol
Tall kan ikke forekomme	.
Oppgave mangler	..
Oppgave mangler foreløpig	...
Tall kan ikke offentliggjøres	:
Null	-
Mindre enn 0,5 av den brukte enheten	0
Mindre enn 0,05 av den brukte enheten	0,0
Foreløpig tall	*
Brudd i den loddrette serien	—
Brudd i den vannrette serien	
Desimaltegn	,

Forord

Seksjon for eiendom-, areal- og primærnæringsstatistikk fikk sommeren 2018 et Grantsprosjekt (Eurostat, avtalenummer 2018.0217) for «Modernisation, improvement and integration of statistics related to Integrated Farm Statistics and nutrient budgets». Et av målene i prosjektet er å undersøke om spørsmål om arbeidstid som stilles til oppgavegiverne i de fullstendige Landbrukstellingene hvert tiende år (hovedgrunn) og de årlige utvalgsbaserte Landbruksundersøkelsene ca. hvert tredje år, kan erstattes helt eller delvis med opplysninger fra A-ordningen. Dette arbeidet ble utført av Seksjon for metoder med støtte fra Seksjon for eiendom-, areal- og primærnæringsstatistikk.

Statistisk sentralbyrå, 17. september 2020

Arvid Olav Lysø

Sammendrag

Årsaken til å undersøke potensiell bruk av A-ordningen i Landbrukstellingene og Landbruksundersøkelsene, forklares med at bruk av opplysninger fra A-ordningen kan redusere både produksjonskostnader og oppgavebyrde i tellinger og utvalgsundersøkelser.

Slik det er kjent, blir ikke den faktiske arbeidstiden meldt inn i A-ordningen, men vi kan finne et tilnærmet riktig tall for faktisk arbeidstid ved hjelp av utbetalt lønn, avtalte timer og andre opplysninger tilgjengelige i A-ordningen. Likevel er det ikke nok nødvendige opplysninger tilgjengelig, og resultater viser at:

- A-ordningen omfatter ikke alle enheter (jordbruksbedriftene) i populasjonene til landbruks-tellingene og landbruksundersøkelsene.
- A-ordningen omfatter ikke alle grupper som skal oppgi timer arbeidet på jordbruksbedriftene. Innehaver av ENK, deltakere i ANS/DA, ektefelle/samboer og delvis familiemedlemmers arbeidstid mangler.
- Avtalt arbeidstid fra A-ordningen kan ikke brukes som faktisk arbeidstid. Avtalt arbeidstid fra A-ordningen er tilnærmet lik faktisk arbeidstid, men må justeres ved bruk av utbetalt lønn og andre relevante opplysninger. For det formålet kan brukes vanlige logiske regler, nærmeste nabo imputering (NNI) og Ratemodell (RM).
- Det anbefales ikke å bruke A-ordningen til å erstatte spørsmål i Landbruks-tellingen eller Landbruksundersøkelsene. Men en del av opplysninger kan brukes fra A-ordningen med editeringsjustering, og som et trinn videre til reduksjon av oppgavebyrden.

Innhold

Forord.....	3
Sammendrag.....	4
1. Bakgrunn.....	6
2. Data om timeverk i Landbrukstellingen 2020 og Landbruksundersøkelser.....	6
3. Data om timeverk i A-ordningen.....	7
4. Spørsmål som potensielt kan erstattes i landbrukstellingen og landbruksundersøkelser fra A-ordningen	9
5. Analyser av sammenhengen mellom timer fra A-ordningene og timer fra spørreskjema	9
5.1. Data brukt	9
5.2. Antall jordbruksbedrifter i Landbrukspopulasjonen som er tilgjengelig i A- ordningen.....	9
5.3. Total årslønn og årarbeidstid.....	10
5.4. Total årarbeidstid for populasjonen U.....	16
5.5. Total årarbeidstid utenom gård for bruker/innehaver og samboer/ektefelle	17
6. Konklusjoner og anbefalinger	18
Vedlegg A: Ratemodellen er en beskrivelse av data i Landbruksstatistikken.....	20
Vedlegg B: Forskjellige begreper i Landbruksstatistikken	25
Figurregister	26
Tabellregister.....	27

1. Bakgrunn

Seksjon for eiendom-, areal- og primærnæringsstatistikk (S426) fikk sommeren 2018 et Grantsprosjekt (Eurostat, avtalenummer 2018.0217) for «Modernisation, improvement and integration of statistics related to Integrated Farm Statistics and nutrient budgets». Et av målene i prosjektet er å undersøke om spørsmål om arbeidstid som stilles til oppgavegiverne i de fullstendige Landbrukstellingene hvert tiende år (hovedgrunn) og de årlige utvalgsbaserte Landbruksundersøkelsene ca. hvert 3. år, kan erstattes helt eller delvis med opplysninger fra A-ordningen. Det vil redusere både produksjonskostnader og oppgavebyrden i tellinger og utvalgsundersøkelser.

I kapittel 2 gis det en kort gjennomgang av spørsmålene om arbeidstid i jordbruksbedrift som er planlagt hentet inn i Landbrukstellingen 2020. I kapittel 3 beskrives både arbeidstid og utbetaling av lønn hentet fra A-ordningen for å beregne timeverkene til enhetene (jordbruksbedriftene) for andre enn den næringsdrivende selv, ektefellen, familiemedlemmene og innleide næringsdrivende, og utfordringene ved dette. I kapittel 4 gis en vurdering om hvilke spørsmål fra LU-ene og LT-ene som kan erstattes med A-ordningen. Analyser av sammenhengen mellom timer fra A-ordningen og timer fra spørreskjema er beskrevet i kapittel 5. Til slutt gis en konklusjon og anbefaling i kapittel 6 etterfulgt av noen vedlegg.

2. Data om timeverk i Landbrukstellingen 2020 og Landbruksundersøkelser

Statistikk om arbeidsinnsatsen i landbruket er svært viktig for gjennomføringen av de årlige forhandlingene mellom staten og organisasjonene i landbruket (Norges bondelag og Norsk Bonde- og småbrukarlag). Derfor er innhenting av data om arbeidsinnsatsen på de enkelte jordbruksbedrifter i Landbruksundersøkelsene og de tiårige Landbrukstellingene helt avgjørende for å produsere statistikk om timeinnsatsen for en jordbruksenhet. (En jordbruksbedrift er den enheten som består av de landbrukseiendommer som drives under felles ledelse, normalt av en selvstendig næringsdrivende.)

I landbruksstatistikken er det to viktige enheter, nemlig jordbruksbedrift og landbrukseiendom. Dersom en sammenlikner disse med enheter i andre næringer svarer landbrukseiendommen til den delen av en juridisk enhet innenfor en kommune, mens jordbruksbedrift likner mest på begrepet statistisk foretak ellers i næringsstatistikken. Slik sett ville det egentlig vært mest riktig å bruke betegnelsen landbruksforetak om jordbruksbedrift. Når bonden søker om produksjonstilskudd, skal han/hun søke om produksjonstilskudd til foretaket fordi den som søker må ha et organisasjonsnummer som myndighetene skal betale ut tilskuddet til.

I Landbrukstellingen 2010 var det to skjema, det ene for organisasjonsform ENK og det andre for ANS, DA, AS, osv. Dersom en ser nærmere på spørreskjemaet for organisasjonsform ENK i Landbrukstellingen 2010, er det hele 18 spørsmål som stilles for å få en nøyaktig kartlegging av den totale arbeidsinnsatsen på en jordbruksbedrift i en tolv månedersperiode (ett år). Spørsmålene kan deles i fire grupper:

1. Timeinnsatsen til driftsleder for jordbruksbedriften. Driftsleder kan være innehaver, samboer/ektefelle, et familiemedlem eller en annen, og dennes arbeidstid på og utenfor landbruksbedriften (spørsmål 8)
2. Kartlegge andre som har utført lønnet eller ulønnet arbeid på jordbruksbedriften siste tolv måneder etter type ansatte, type jobb og kjønn (spørsmål 9-21). Her er tre grupper (typer): familiemedlemmer,

fastarbeidende, ikke-fastansatte/sporadiskhjelp. Av interesse er alle som er 16 år eller eldre. Typer av jobb i landbruksbedriften kan være én av tre: 1. husdyrhold, jord- og hagebruk og veksthus, 2. skogbruk, 3. andre tilleggsnæringer som utnytter jordbruksbedriftens arealer, bygninger, maskiner og/eller råvarer. For personlig bruker og ektefelle/samboer ble det i tillegg spurt om arbeidstid utenom bedriften. I år med rapportering til Eurostat rapporteres det om arbeid utenom jordbruksbedriften er hoved- eller bisysselsetting. I undersøkelser uten rapportering til Eurostat er arbeidstid utenom bruket med for å kontrollere sannferdighet i utfylling av opplysninger.

3. Kartlegge arbeidsinnsats siste tolv måneder til regnskapsfører og veterinær for jordbruksbedriften (spørsmål 18-19)
4. Arbeidsinnsatsen til andre gårdbrukere, entreprenører, håndverkere eller servicefolk (spørsmål 20-21)

I spørreskjemaet for andre organisasjonsformer (ANS, DA, AS m.fl.) var det 14 spørsmål som ble stilt for å kartlegge den totale arbeidsinnsatsen på en jordbruksbedrift:

1. Timeinnsatsen til driftsleder og deltakere i ANS eller DA på landbruksbedriften. I tillegg skulle driftsleder oppgi timer utenom jordbruksbedriften.
2. Familiemedlem er ikke relevant. Kartleggingen gjelder to typer ansatte: fastarbeidende og ikke-fastansatte/sporadiskhjelp.

Gruppe 3 og 4 er identiske med organisasjonsform ENK.

Vi ser at personene som utfører arbeidsinnsatsen kan både være lønnet og ulønnet, altså er det ikke sikkert at all arbeidstiden som rapporteres vil bli registrert i A-ordningen. Kun dersom det er en lønnet driftsleder utenom bruker selv eller ektefelle vil denne arbeidsinnsatsen registreres i en A-melding som skal leveres dersom:

- Har du ansatte, eller utbetaler lønn, pensjoner eller andre ytelser, har du opplysningsplikt i A-ordningen og skal levere A-melding.

I Landbrukstelling 2020 blir gruppe 3 slått sammen med gruppe 4. For resten av spørsmålene blir strukturen bevart. Derfor kan spørsmål i Landbruksundersøkelsene og Landbrukstelling 2020 deles i de samme gruppene med noen små modifikasjoner. Undersøkelsene om arbeidskraft har alltid omfattet en periode på tolv måneder, men start- og sluttmåned har variert.

Tilgjengelige resultater fra Landbruksundersøkelsen 2018 og data fra A-ordningen i relevante perioder, har vært benyttet i arbeidet som presenteres i dette notatet.

3. Data om timeverk i A-ordningen

«A-ordningen er en samordnet måte for arbeidsgivere å rapportere opplysninger om inntekt og ansatte til NAV, SSB og Skatteetaten»

[<https://www.skatteetaten.no/bedrift-og-organisasjon/arbeidsgiver/a-meldingen/om-A-ordningen/om-A-ordningen/>].

Ut fra denne opplysningen på Skatteetatens hjemmeside virker det som at dersom en jordbruksbedrift ansetter en person og/eller utbetaler lønn for denne personens arbeidsinnsats, må det sendes en A-melding til Skatteetaten som samarbeider med NAV og SSB om disse data. Men her kan noen av de ansatte være omfattet av unntak. Jordbruksbedriften trenger ikke rapportere opplysninger om:

1. Ansatte som fikk mindre enn 1000 nkr i året
2. Utbetalinger til selvstendig næringsdrivende

3. For enkeltpersonforetak (ENK) skal innehaver og ektefelle ikke behandles som ansatte.
[\[https://www.skatteetaten.no/bedrift-og-organisasjon/arbeidsgiver/a-meldingen/hvem-ma-levere-a-melding/\]](https://www.skatteetaten.no/bedrift-og-organisasjon/arbeidsgiver/a-meldingen/hvem-ma-levere-a-melding/) Det samme gjelder deltakere i ANS/DA [\[https://www.altinn.no/starte-og-drive/starte/valg-av-organisasjonsform/ansvarlig-selskap/\]](https://www.altinn.no/starte-og-drive/starte/valg-av-organisasjonsform/ansvarlig-selskap/).

De tre viktigste opplysningene som skal sendes inn hver måned for hver av de personene som tilfredsstiller kravene for å være ansatt er:

- Identifikasjon (F-nummer eller D-nummer), lønn, avtalt arbeidstid.

Avtalt arbeidstid kan gis som:

- Avtalt stillingsprosent og arbeidstimer per uke relevant for 100 prosent stilling
- Arbeidstimer jobbet i måneden med relevante utbetalinger hvis man får betalt timelønn.

Det betyr at arbeidsgiver ikke rapporterer arbeidstiden direkte, men indirekte gjennom hvor mange timer en 100 prosent stilling svarer til og hvilken stillingsprosent arbeidstakeren har arbeidet. Disse to opplysningene blir så brukt til å beregne antall timer arbeidstakeren har arbeidet i perioden. Lønna som rapporteres er utbetalt lønn til arbeidstakeren siste måned, mens arbeidstiden er basert på arbeidstiden til 100 prosent stilling i en uke i den siste måneden. Det siste rapporteres sammen med stillingsprosent til arbeidstakeren i samme uke.

I situasjonen med timelønn bør det være lettere å få registrert faktisk arbeidet tid. Basert på data som sendes inn i A-meldingene er det derfor mulig å beregne data for både antall timer utført arbeid for siste måned rapportert via A-ordningen og utbetalt lønn for samme måned. Utfordringen er ofte at mens timetallet gjelder rapporteringsmåneden, vil lønna delvis gjelde arbeidstiden i forrige måned dersom det for eksempel skal utbetales tillegg på grunn av overtid eller vaktordninger. En liknende utfordring er at årlige lønnsutbetalinger blir registrert i A-ordningen bare én gang i slutten av året eller høstsesongen. Men i utgangspunktet kan vi finne den samlede arbeidstiden for hver måned for denne arbeidsgiveren for alle ansatte der lønn ble utbetalt og dermed den samlede utbetalte lønna fra arbeidsgiveren til alle arbeidstakerne hos arbeidsgiveren.

I utgangspunktet burde disse tallene svare til alle arbeidsforhold som oppgis av jordbruksbedriftene for de tre siste gruppene (spørsmålene 9-21), også inkludert en ansatt driftsleder utenom bruker selv eller ektefelle. Arbeidstiden til de to siste gruppene som er selvstendig næringsdrivende vil ikke finnes i data for jordbruksbedriften i A-ordningen. Det vil også være mangler i data hentet fra A-ordningen dersom timer som rapporteres i landbruksundersøkelsene eller Landbrukstelingen hvert tiende år inneholder timer som utføres for innleid arbeidskraft fra andre jordbruksbedrifter. Samtidig gir utfordringer leieforhold mellom jordbruksbedriftene. Når en jordbruksbedrift (A) leier arbeidskraft fra en annen jordbruksbedrift (B), må jordbruksbedrift A som leier arbeidskraften, betale for tjenesten. I dette tilfellet ofte betaler jordbruksbedrift A for tjenester til jordbruksbedriften B, og etterpå betaler jordbruksbedrift B til ansatte for utført arbeid.

Konklusjonen er at det på forhånd er grunn til å tenke seg at timetall for arbeidsinnsatsen registrert i A-meldingene avviker fra den arbeidsinnsatsen som blir registrert i en årlig Landbruksundersøkelse eller en tiårig Landbrukstelling selv om vi sørger for at registreringene gjelder samme periode. Analysen vil vise i hvilken grad registrerte i A-meldingene arbeidstid kan erstatte spørsmålene i Landbrukstelling eller Landbruksundersøkelsene.

4. Spørsmål som potensielt kan erstattes i landbrukstellingen og landbruksundersøkelser fra A-ordningen

Hvis en tar i betraktning kapitlene 2 og 3, kan en se hvilke spørsmål fra undersøkelsen som kan erstattes med data fra A-ordningen.

- Spørsmål om arbeidstid til samboer som er ansatt kan finnes i A-ordningen. Samtidig finnes arbeidstid til samboer/ektefelle og bruker/innehaver som ikke er ansatt i jordbruksbedriften (jobber utenom jordbruksbedriften). Men når samboer/ektefelle er den selvstendige brukeren selv, kan ikke arbeidstiden for han eller hun hentes fra A-ordningen.
- Spørsmål om arbeidstid til driftsleder som er ansatt i jordbruksbedriften.
- Spørsmål om familiemedlemmer, fast ansatte og ikke fast ansatte kan også finnes i A-ordningen. Utfordringen kan være å klassifisere ansatte til en av de gruppene.

For å dele arbeidstid i grupper etter arbeidstype (1. husdyrhold, jord- og hagebruk og veksthus; 2. skogbruk; 3. tilleggsnæringer) kan yrkeskode for arbeidstakere kanskje gi et svar langt på vei, men antakelig ikke fullt ut fra A-ordningen.

5. Analyser av sammenhengen mellom timer fra A-ordningene og timer fra spørreskjema

5.1. Data brukt

Data fra Landbruksundersøkelsen 2018 om arbeidstid i perioden 1. januar 2017 – 31. desember 2017 ble sammenholdt med opplysninger fra A-ordningen for den relevante perioden.

5.2. Antall jordbruksbedrifter i Landbrukspopulasjonen som er tilgjengelig i A-ordningen

Kobling av LU18 mot A-ordningen 2017 (AO17) viser at omtrent 50 prosent av jordbruksbedriftene ikke melder inn noen informasjon om ansatte til AO17 (Tabell 5.1). Antall enheter som har tilgjengelige utbetalinger relevante for arbeidstid i A-ordningen er enda mindre: datasettet i A-ordningen som er relevant for Landbruksundersøkelsen 2018 er på kun 3 047 enheter.

Tabell 5.1 Antall jordbruksbedrifter i LU18 (timer i 2017) og AO17 (frtk_id_ssb = orgnr_delreg)

LU18	AO17
7941	3978 (3047 med tilgjengelig arbeidstid)
100%	50,1% (38,4% med tilgjengelig arbeidstid)

Kilde: Statistisk sentralbyrå.

For hele perioden 2015–18 gjelder at i gjennomsnitt 41,27 prosent av jordbruksbedriftene i populasjonen fra 2018 har ingen informasjon i A-ordningen. Dessuten finnes det ikke noen tendens til at dette antallet øker mye over tid (Tabell 5.1 og Tabell 5.2). Grunnen til dette er at arbeidshjelpa for mange av enhetene i Landbruksundersøkelsen er innleie fra andre selvstendig næringsdrivende, eller at de kjøper tjenesten (arbeidshjelp) fra andre jordbruksbedrifter.

Tabell 5.2 Antall jordbruksbedrifter i 2018 og kobling til A-ordningen i forskjellig år

Jordbruks- bedrifter i Landbruks- registren 2017	AO15	AO16	AO17	AO18 ²	AO19(Siste EDAG_PERIOD E = '01-07- 2019')
39621	16122	16425	16412	16453	8316 ¹

¹ Mange jordbruksbedrifter melder informasjonen i desember for hele året, derfor er ikke tilgjengelig så tidlig

² I perioden 2015- 2018 har antall jordbruksbedrifter gått ned, mens antall koblinger til AO har økt svakt. Prosent jordbruksbedrifter med kobling til AO har derfor økt noe.

Kilde: Statistisk sentralbyrå.

For å finne ut om det er noen forskjell mellom de som rapporterer gjennom A-ordningen og ikke, ble antall enheter i gruppene beregnet med hensyn til forskjellige variabler: driftsform, organisasjonsform, type ansatte (fast, familie, sporadisk hjelp), omsetning, totalt antall timer svart i Landbruksundersøkelsen. I alle gruppene var det mulig å finne eksempler på både de som rapporterte og de som ikke rapporterte, derfor viste analysen ikke noe forskjell. Den eneste tendensen var at jo flere totalt antall timer jordbruksbedriften hadde svart i Landbruksundersøkelsen, jo større sjanse var det for å finne enheten i A-ordningen (Tabell 5.3).

Tabell 5.3 Jordbruksbedrifter i A-ordningen og jordbruksbedrifter ikke i A-ordningen, fordelt etter totalt antall timer i Landbruksundersøkelsen 2018

Grupper etter totalt antall timer svart i Landbruks- undersøkelsen	Antall enheter som finnes i A-ordningen		Antall enheter som ikke finnes i A-ordningen	
	Frekvens	Prosent	Frekvens	Prosent
group = 1: [; 1000]	2111	53,07	3014	76,17
group = 2:]1000; 2500]	1086	27,30	687	17,36
group = 3:]2500; 5000]	509	12,80	222	5,61
group = 4:]5000;10000]	171	4,30	29	0,73
group = 5:]10000; ∞[	101	2,54	5	0,13

Kilde: Statistisk sentralbyrå.

For videre analyse ble det brukt enheter som finnes i både Landbruksundersøkelsen og i A-ordningen med tilgjengelige relevante lønnsutbetalinger.

5.3. Total årslønn og årsarbeidstid

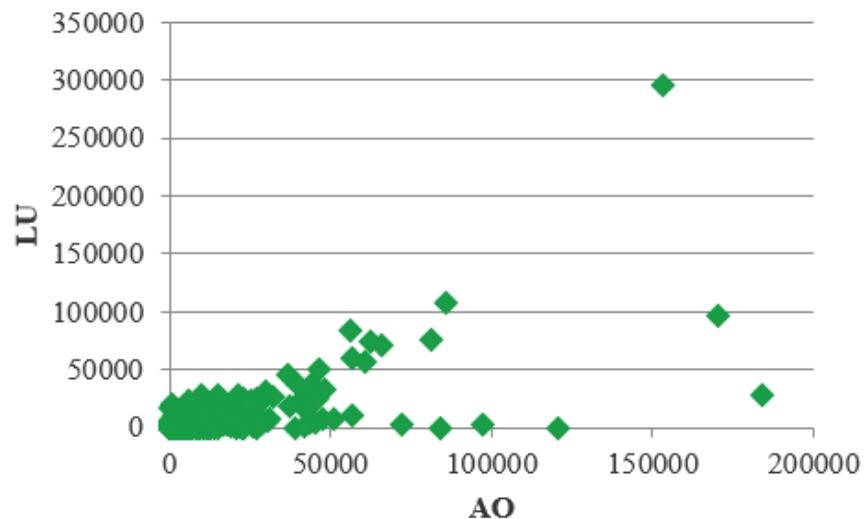
Slik det allerede er nevnt tidligere i notatet, blir ikke den faktiske arbeidstiden meldt inn i A-ordningen, men vi kan finne et tilnærmet riktig tall for faktisk arbeidstid ved hjelp av utbetalt lønn, avtalte timer og andre opplysninger tilgjengelige i A-ordningen. Fra landbruksundersøkelsen får vi ikke navn på ansatte. Derfor brukes aggregert arbeidstid per jordbruksbedrift i året for sammenligning av kilder. For å oppnå tallene i A-ordningen ble følgende algoritme benyttet:

1. Finne enheter fra A-ordningen som er i relevant landbruks-populasjon.
2. Ekskludere enhetene som ikke har noen utbetalt lønn relevant til arbeidstid, de som har "tom" arbeidstype, og de som ikke har relevante yrkeskoder.
3. Ekskludere negative utbetalinger i januar
4. Hvis lønn per time var for høy eller for lav for spesifisert arbeidstid, kjøre NNI basert på lønn og yrkeskode på nivåer 5,4,3.
5. Når fast lønn eksisterer og arbeidsstillingsprosent er 0 i måned så settes den til 100. Videre tid skal editeres.
6. Beregne arbeidsdager per måned per ansatt: ekskludere nasjonal ferie, sykedager, permisjoner, permitteringer, lønnsendringer og ferie for de som jobber lengre enn 1 år. Tar i betraktning arbeidsstart og slutt datoer.
7. Beregne månedlig arbeidstid per ansatt og editere med NNI basert på lønn og yrkeskode hvis lønn var for høy eller for lav.

Mens for å få tall i Landbruksundersøkelsen, ble det arbeidstiden til samboer, familiemedlemmer, fastansatte og sporadisk hjelp summert. Deltagere og selvstendig næringsdrivende blir ikke med.

Videre kan man finne spredningsdiagram som viser total årsarbeidstid i følge i Landbruksundersøkelsen (LU) mot total årsarbeidstid i A-ordningen (AO) (Figur 5.1 og Tabell 5.4). Figuren viser at det finnes tall i undersøkelsen som er mye mindre enn de som ble rapportert i A-ordningen.

Figur 5.1 Spredningsdiagram total årsarbeidstid for A-ordningen 2017 (AO) vs. Landbruksundersøkelsen 2018 (LU)



Kilde: Statistisk sentralbyrå.

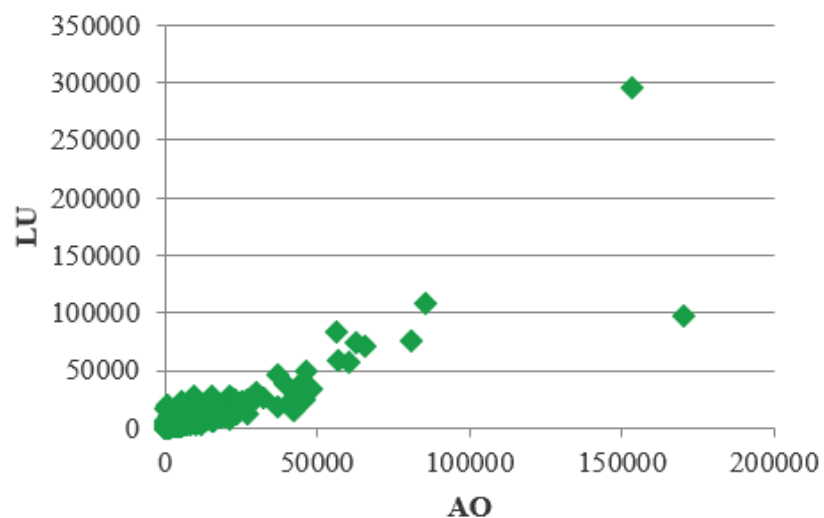
Tabell 5.4 Deskriptiv statistikk og Pearson korrelasjon for total årsarbeidstid for A-ordningen 2017 (AO) vs. Landbruksundersøkelsen 2018 (LU)

	N	Sum	Mean	Median	1st Qu.	3rd Qu.	StdDev	Min	Max
AO	2821	7388810	2619,22	478	170	1887	8878,72	0	183894
LU	2821	5848045	2073,04	650	200	1820	7793,18	0	295290
Person Correlation Coefficients			0,69844						

Kilde: Statistisk sentralbyrå.

Videre ble de enhetene som hadde for lav arbeidstid i Landbruksundersøkelsen sammenliknet med lønn tilgjengelig fra A-ordningen, ekskludert. Grunnen til dette kan være at det for enheter som ikke har svart noe, ble gjort feil editeringer eller at administrative yrker ble inkludert feilaktig.

Figur 5.2 Spredningsdiagram total årsarbeidstid fra A-ordningen 2017(AO) vs årsarbeidstid Landbruksundersøkelsen 2018(LU) uten enheter som har lav arbeidstid i Landbruksundersøkelsen sammenliknet med lønn tilgjengelig fra A-ordningen



Kilde: Statistisk sentralbyrå.

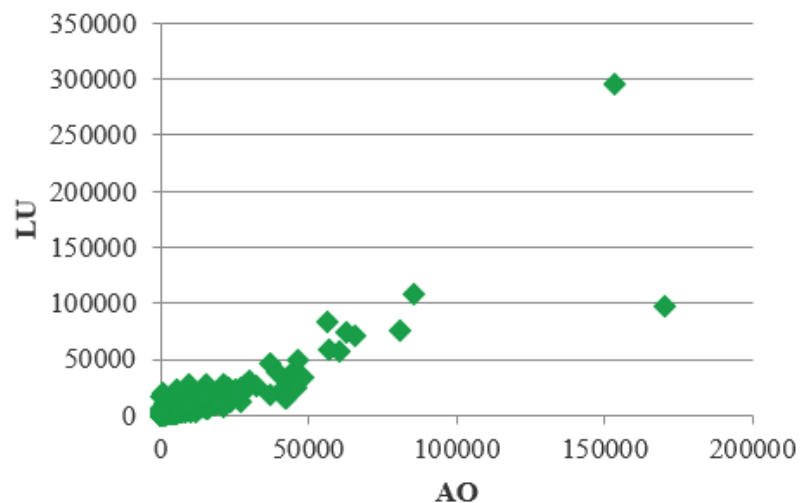
Tabell 5.5 Deskriptiv statistikk og Pearson korrelasjon for total årsarbeidstid fra A-ordningen 2017(AO) vs årsarbeidstid Landbruksundersøkelsen 2018(LU) uten enheter som har lav arbeidstid i Landbruksundersøkelsen sammenliknet med lønn tilgjengelig fra A-ordningen

	N	Sum	Mean	Median	1st Qu.	3rd Qu.	StdDev	Min	Max
AO	2363	5140500	2175,41	432	165	1699	7379,92	0	170548
LU	2363	5636512	2385,32	875	306,5	2050	8448,23	0	295290

Person Correlation Coefficients 0,88627

Kilde: Statistisk sentralbyrå.

Figur 5.3 Spredningsdiagram årslønn fra A-ordningen 2017(AO) vs årsarbeidstid Landbruksundersøkelsen 2018(LU) uten enheter som har lav arbeidstid i Landbruksundersøkelsen sammenliknet med lønn tilgjengelig fra A-ordningen



Kilde: Statistisk sentralbyrå.

Tabell 5.6 Deskriptiv statistikk og Pearson korrelasjon for årslønn fra A-ordningen 2017(AO) vs årsarbeidstid Landbruksundersøkelsen 2018(LU) uten enheter som har lav arbeidstid i Landbruksundersøkelsen sammenliknet med lønn tilgjengelig fra A-ordningen

	N	Sum	Mean	Median	1st Qu.	3rd Qu.	StdDev	Min	Max
AO	2363	5140500	2175,41	432	165	1699	7379,92	0	170548
LU(SSA)	2363	5636512	2385,32	875	306,5	2050	8448,23	0	295290

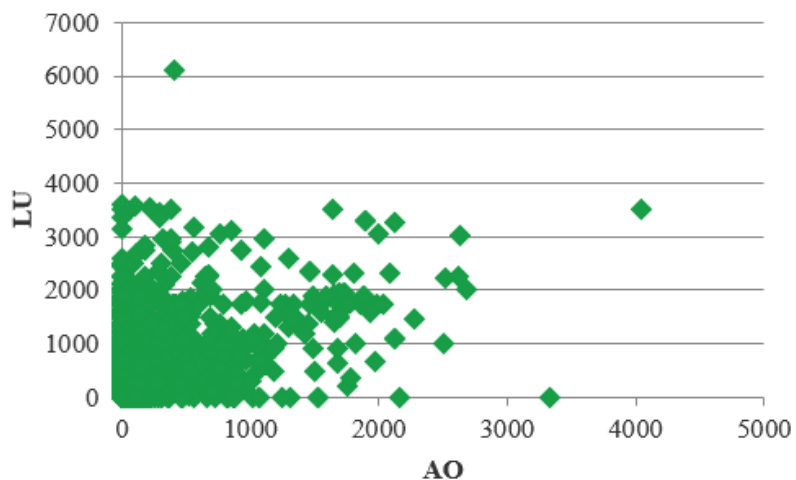
Person Correlation Coefficients 0,88627

Kilde: Statistisk sentralbyrå.

Fra Figur 5.2 og 5.3 ser man at både arbeidstid og lønn fra A-ordningen har god sammenheng med arbeidstid fra undersøkelsen. En uteligger i Figur 5.2 (AO = 170548) skyldes editeringer i Landbruksundersøkelsen (i rådata hadde denne enheten 185 800 timer).

I Figur 5.4 og Figur 5.5 ser man at arbeidstid for familiemedlemmer og for samboer i Landbruksundersøkelsen viser liten sammenheng med tallene hentet fra A-ordningen. Grunnen til forskjellene kan være at begrepet familiemedlemmer var ikke godt nok spesifisert i spørreskjema, eller at samboer og ektefeller har utført ulønnet arbeid som ikke skal registreres i A-ordningen. Det resulterer i at en fortsatt må spørre om arbeidstiden for disse gruppene (familiemedlemmer og samboere/ektefeller). For resten av ansatte, se Figur 5.6 og Tabell 5.9.

Figur 5.4 Spredningsdiagram total årsarbeidstid for familiemedlemmer fra A-ordningen 2017 (AO) vs årsarbeidstid Landbruksundersøkelsen 2018 i relevant gruppe (LU)



Kilde: Statistisk sentralbyrå.

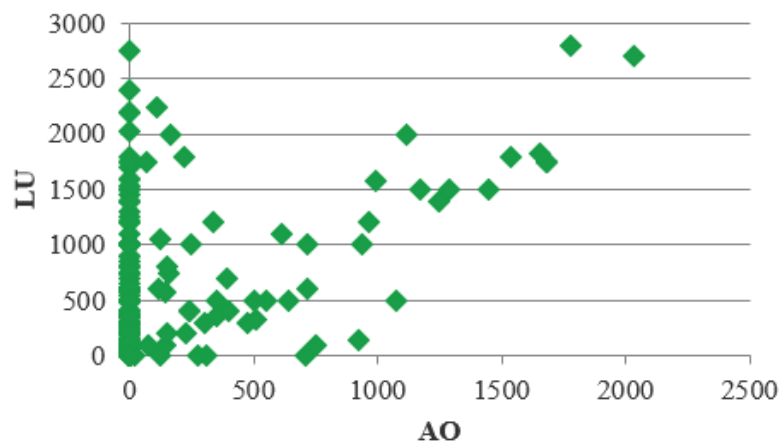
Tabell 5.7 Deskriptiv statistikk og Pearson korrelasjon for total årsarbeidstid for familiemedlemmer fra A-ordningen 2017 (AO) vs årsarbeidstid Landbruksundersøkelsen 2018 i relevant gruppe (LU)

	N	Sum	Mean	Median	1st Qu.	3rd Qu.	StdDev	Min	Max
AO	2363	431941	182,79	33	0	200,5	357,44	0	4049
LU	2363	984994	416,84	182	0	500	620,01	0	6100

Person Correlation Coefficients 0,44131

Kilde: Statistisk sentralbyrå.

Figur 5.5 Spredningsdiagram total årsarbeidstid for samboer fra A-ordningen 2017 (AO) vs årsarbeidstid Landbruksundersøkelsen 2018 i relevant gruppe (LU)



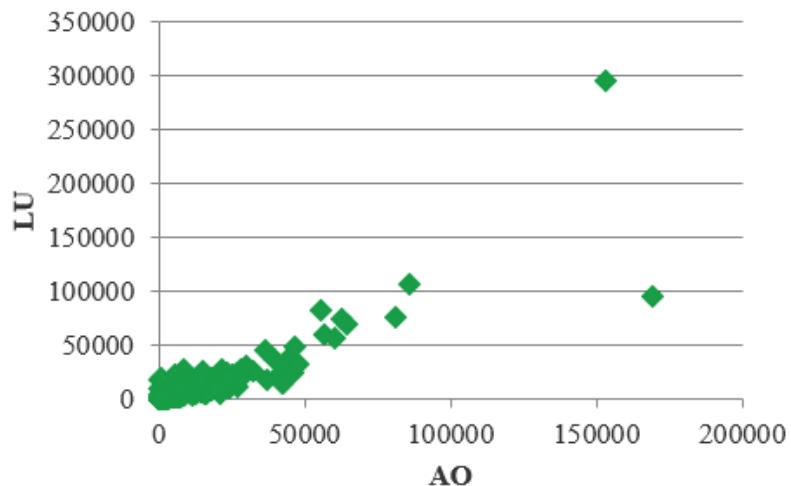
Kilde: Statistisk sentralbyrå.

Tabell 5.8 Deskriptiv statistikk og Pearson korrelasjon for total årsarbeidstid for familiemedlemmer fra A-ordningen 2017 (AO) vs årsarbeidstid Landbruksundersøkelsen 2018 i relevant gruppe (LU)

	N	Sum	Mean	Median	1st Qu.	3rd Qu.	StdDev	Min	Max
AO	371	33197	89,48	0	0	0	299,63	0	2032
LU	371	148877	401,29	190	0	500	559,94	0	2800

Person Correlation Coefficients 0,51510

Kilde: Statistisk sentralbyrå.

Figur 5.6 Spredningsdiagram total årsarbeidstid for ansatte som ikke er familiemedlemmer eller samboer/ektefelle, fra A-ordningen 2017 (AO) vs årsarbeidstid Landbruksundersøkelsen 2018 i relevant gruppe (LU)

Kilde: Statistisk sentralbyrå.

Tabell 5.9 Deskriptiv statistikk og Pearson korrelasjon for total årsarbeidstid for ansatte som ikke er familiemedlemmer eller samboer/ektefelle, fra A-ordningen 2017 (AO) vs årsarbeidstid Landbruksundersøkelsen 2018 i relevant gruppe (LU)

	N	Sum	Mean	Median	1st Qu.	3rd Qu.	StdDev	Min	Max
AO	2237	4569535	2042,71	206	0	1484	7558,29	0	169236
LU	2237	4488617	2006,53	300	0	1700	8671,1	0	295290

Person Correlation Coefficients 0,88896

Kilde: Statistisk sentralbyrå.

En videre bør skille arbeidstiden for ansatte som ikke er familiemedlemmer eller samboer/ektefelle mellom fast ansatte og tilfeldig hjelp. Dessverre er det ikke mulig å gjøre dette basert på data fra A-ordningen alene. Det ble brukt data fra både A-ordningen og Landbruksundersøkelsen, og Random Forest, SVM, Naive Bayes modeller for klassifisering fast/ikke-fast ansatte. De beste modellene klarte bare rundt 60 prosent nøyaktighet.

I tillegg ser man at total arbeidstid i A-ordningen kan avvike mye fra arbeidstid i undersøkelsen: for noen enheter med høyest arbeidstid i Landbruksundersøkelsen klarte vi bare finne halvparten av arbeidstiden i A-ordningen (se eksempel i Figur 5.6, enhet med den høyeste tallen på LU-aksen).

Vider ble det vurdert enheter med arbeidstid i A-ordningen $< 100\,000$. For å undersøke om $T_{andre,i}^{LU}$ (total årsarbeidstid T relatert for andre ansatte enn familiemedlemmer eller samboer/ektefelle beregnet i Landbruksundersøkelsen LU for jordbruksenhet i) og $T_{andre,i}^{A-ordn}$ (total årsarbeidstid T relatert for andre ansatte enn familiemedlemmer eller samboer/ektefelle beregnet i A-ordningen AO for jordbruksenhet i) er de samme timetallene som en får fra de to kildene, må regresjon koeffisienten $\beta_{andre} = 1$ i følgende formel:

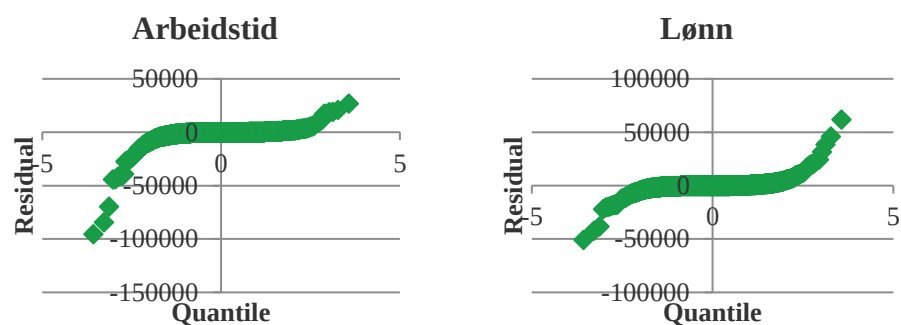
$$T_{andre,i}^{LU} = \beta_{andre} \cdot T_{andre,i}^{A-ordn} + \varepsilon_{andre,i} \quad \text{der } \varepsilon_{andre,i} \sim N(0, T_{andre,i}^{A-ordn} \cdot \tau_{andre}^2)$$

For å undersøke dette tester vi først om $\varepsilon_{andre,i} \sim N(0, T_{andre,i}^{A-ordn} \cdot \tau_{andre}^2)$. Videre trinn for sammenheng av tall kan en finne i vedlegget.

Likhet av varians og gjennomsnitt (F-test, t-test med like varianser) ble tilbakevist, samtidig ble test for sammenligning av fordelinger (Kolmogorov-Smirnov test) tilbakevist. For samme enheter viste Student test og Wilcoxon test at gjennomsnitt og median av feil er forskjellig fra 0. Sammen med Q-Q plot av feil (Figur 5.7) viser feilanalyse at antakelsen om Normalfordeling er tilbakevist.

Samme situasjon med feil ble sett for arbeidstid-landbrukspredikeringer basert på lønn (Figur 5.7).

Figur 5.7 Q-Q plot for feil i predikeringer av total årsarbeidstid for ansatte som ikke er familiemedlemmer eller samboer, basert på årsarbeidstid og årslønn fra A-ordningen



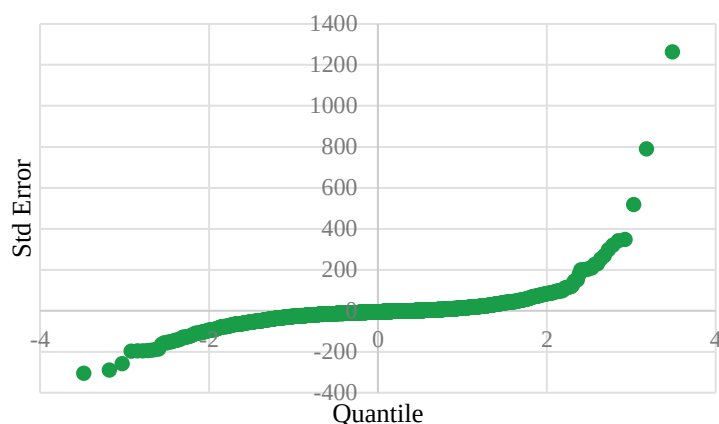
Kilde: Statistisk sentralbyrå.

Figur 5.7 viser Q-Q plot for standardisert feil $\varepsilon_{andre,i}^{std}$ for totalarbeidstid slik at alle har forventet lik varians:

$$\varepsilon_{andre,i}^{std} = \varepsilon_{andre,i} / \sqrt{T_{andre,i}^{A-ordn}}, \text{ når } \hat{\beta}_{andre} = 1.$$

Figur 5.8 viser at når undersøkelsens arbeidstid er større enn A-ordningens arbeidstid, er avvikene veldig store sammenliknet med motsatte tendenser (undersøkelsens arbeidstid er mindre enn arbeidstid i A-ordningen). Det kan igjen tyde på underdekning og at bruk av A-ordningen i stedet for å spørre om arbeidstid i skjema er utfordrende.

Figur 5.8 Q-Q plot for standardisert feil i predikeringer av total årsarbeidstid for ansatte som ikke er familiemedlemmer eller samboer/ektefelle basert på årsarbeidstid fra A-ordningen



Kilde: Statistisk sentralbyrå.

5.4. Total årsarbeidstid for populasjonen U

Før en fastslår at tallene fra A-ordningen er dårligere for vårt formål enn tallene fra Landbruksundersøkelsen, må en også tenke seg at det er feil i disse, derfor er det best å sjekke kildene på makronivå.

Estimat for total årsarbeidstid for ansatte som ikke er familiemedlemmer eller samboer/ektefelle $\hat{T}_{andre}^{LU}$ ble beregnet i Landbruksundersøkelsen:

$$\hat{T}_{andre}^{LU} = \sum_{i \in S} T_{andre,i}^{LU} w_i,$$

hvor w_i – vekt av jordbruksenhet i brukte i Landbruksundersøkelsen 2018.

Hvis vi erstatter Landbruksundersøkelsen med tilgjengelige data fra A-ordningen, beregner vi estimat for total årsarbeidstid for samme type ansatte:

$$\hat{T}_{andre}^{LU_{AO}} = \sum_{i \in S} \left(T_{andre,i}^{LU} I(i) + T_{andre,i}^{A-ordn} (1 - I(i)) \right) w_i,$$

hvor $I(i) = \begin{cases} 1, & \text{hvis enhet } i \text{ ikke var funnet i A-Ordningen} \\ 0, & \text{ellers} \end{cases}$

Tabell 5.10 Estimat for total årsarbeidstid for ansatte som ikke er familiemedlemmer, samboer/ektefeller beregnet fra Landbruksundersøkelsen og kombinasjonen av landbruksundersøkelsen med A-ordningen med bruk av vektorer fra Landbruksundersøkelsen

	Menn	Kvinner	Total
$\hat{T}_{andre}^{LU}$	9 121 208	3 316 156	12 437 364
$\hat{T}_{andre}^{LU_{AO}}$	11 540 450	4 127 362	15 670 631

Kilde: Statistisk sentralbyrå.

Beregningene av forskjell mellom gammelt og nytt tall gir

$$\frac{\hat{T}_{andre}^{LU} - \hat{T}_{andre}^{LU_{AO}}}{\hat{T}_{andre}^{LU}} * 100\% = -25,97\%$$

Hvis vi grupperer etter driftsformer (Tabell 5.11), ser vi at for noen driftsformer er forskjellen stor (for eksempel 3 Hagebruksvekster). Det betyr at Landbruksundersøkelsen ikke uten videre kan erstatte spørsmål om arbeidstid for denne driftsformen med A-ordningen.

Tabell 5.11 Estimert for total årsarbeidstid etter driftsformen for ansatte som ikke er familiemedlemmer, samboer/ektefelle beregnet fra Landbruksundersøkelsen og kombinasjonen av landbruksundersøkelsen med A-ordningen med bruk av vektorer fra Landbruksundersøkelsen

Driftsform	$\hat{T}_{andre}^{LU}$	$\hat{T}_{andre}^{LU,AO}$
1 Korn og oljevekster	449 793	613 152
2 Øvrige jordbruksvekster	420 179	500 596
3 Hagebruksvekster	4 286 091	6 435 539
4 Storfe mjølkeproduksjon	2 623 825	2 771 662
5 Storfe kjøttproduksjon	579 930	665 908
6 Storfe melk og kjøtt	551 666	612 397
7 Sau	756 747	662 473
8 Øvrige grovføretende dyr	397 862	487 053
9 Svin og fjørfe	1 050 667	1 233 293
10 Blandet planteproduksjon	388 806	565 863
11 Blandet husdyrproduksjon	323 630	379 699
12 Plante- og husdyrproduksjon i kombinasjon'	608 162	740 177

Kilde: Statistisk sentralbyrå.

Forskjellen er stor, og derfor er A-ordningens arbeidstidstall for arbeidstypegruppe ikke fornuftig.

5.5. Total årsarbeidstid utenom gård for bruker/innehaver og samboer/ektefelle

For alle enheter ble det sjekket om det finnes opplysninger om arbeidstid til samboer/ektefelle utenom gården. I Tabell 5.12 kan vi se antall enheter som hadde spesifisert arbeidstid for samboer/ektefelle i Landbruksundersøkelsen og tallene som var funnet i A-ordningen.

Tabell 5.12. Antall samboere/ektefeller som er ansatt utenfor jordbruksbedriften

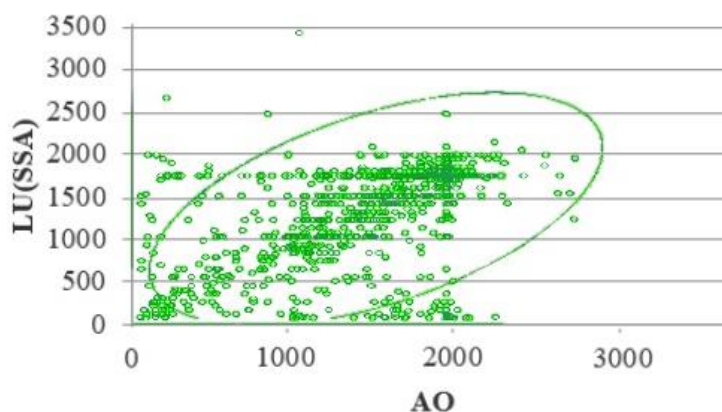
	Ja A-ordningen	Nei A-ordningen
Ja Landbruksundersøkelsen	1 156	203 (for 42 var ikke samboer definert)
Nei Landbruksundersøkelsen	1 212	1 067 (for 391 var ikke samboer definert)

Kilde: Statistisk sentralbyrå.

Fordi det ligger mange enheter i feltet 'Nei Landbruksundersøkelsen' sammenliknet med 'Ja A-ordningen', betyr det at svært mange enheter ikke klarte eller ignorerte å svare på spørsmål om arbeidstid for ektefeller/samboer i undersøkelsen, mens disse opplysningene finnes i A-ordningen. Da ser vi at det i dette tilfelle finnes potensiale for editering av arbeidstid i svart i Landbruksundersøkelsen.

Spredningsdiagram for arbeidstid av samboer spesifisert i A-ordningen mot svart i Landbruksundersøkelsen finnes i Figur 5.9. Den viser at vi for det meste greier å få et fornuftig tall (alt som ligger på diagonalen, i ovalen og under diagonalen).

Figur 5.9 Spredningsdiagram for årsarbeidstid til samboer/ektefeller utenfor jordbruksbedriften fra Landbruksundersøkelsen vs. A-ordningen



Kilde: Statistisk sentralbyrå.

Situasjonen for brukere/innehaver ser ganske lik ut (Tabell 5.13)

Tabell 5.13 Antall brukere/innehavere som er ansatt utenfor jordbruksbedriften

	Ja A-ordningen	Nei A-ordningen
Ja Landbruksundersøkelsen	1286	563
Nei Landbruksundersøkelsen	578	1211

Kilde: Statistisk sentralbyrå.

I både Tabell 5.12 og Tabell 5.13 kan ikke kryssingsverdier i feltet 'Ja Landbruksundersøkelsen vs. Nei A-ordningen' dekkes i A-ordningen på grunn av at både samboere/ektefeller og brukere/ innehavere kan ha arbeid som selvstendig næringsdrivende utenfor jordbruksbedriften. Dette skal fanges opp i LU, men ikke i A-ordningen.

6. Konklusjoner og anbefalinger

Analyse støtter konklusjonen om at det ikke anbefales å erstatte spørsmål om arbeidstid i landbruksundersøkelsen eller tellingen basert på A-ordningens data. Men data fra A-ordningen kan brukes for editering av data om arbeidstid fra Landbruksundersøkelsen/landbrukstelingen.

Samtidig er det mulig å vise beregnede relevante timer eller utbetalt lønn per ansatt til respondent (Figur 6.1). Selv om relevante utbetalinger / beregnede timer ikke er korrekte for alle som finnes i A-ordningen for en del av populasjonen $T_{andre,i}^{LU} = 1 \cdot T_{andre,i}^{A-ordn} + \varepsilon_{andre,i}$, er det en korrekt sammenheng for makrotall. Da er det viktig å inkludere mulighet for respondenten til å editere anbefalte tall. Det er på grunn av at det kan være underdekninger i A-ordningen eller beregnede tall kan ha for stort avvik. De anbefalte tallene bør gjøre utfylling av skjema lettere.

Figur 6.1 Eksempel av spørsmål med anbefalte tall (AT)

Hvor mange timer har hver av de ansatte utført i jordbruksbedriften i 2017?

NAME1	BORSDAG	KJØNN	I husdyr- hold ...	I skog- bruk ...	I tilleggs- næringer	AT*	Familie- medlem	Fast- arbeidende	Ikke-fast- arbeidende
NAME2	BORSDAG	KJØNN				AT	☐	☐	☐

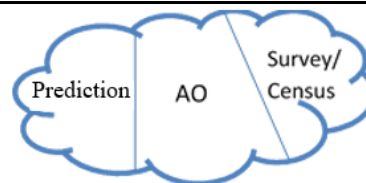
Hvis noen av ansatte ble ikke listet, gi opplysninger ned.

	KJØNN				☐	☐	☐
--	-------	--	--	--	-----------------------	-----------------------	-----------------------

Kilde: Statistisk sentralbyrå.

Før Seksjon for eiendom-, areal- og primærnæringsstatistikk bestemmer seg for å bruke anbefalingen i Figur 6.1, bør seksjonen diskutere det med skjemametodiker.

Med implementering av anbefalte tall i skjema og etter at landbrukstelling er gjennomført, får seksjonen mulighet til å teste maskinlæring for å klassifisere til gruppene etter ansatt type (fast ansatt og tilfeldig hjelp) og justere predikert arbeidstid nærmere faktisk arbeidstid. Den mulighet skal finnes fordi de beregnede og faktiske tallene er kjent per ansatt. Som utgangspunkt i en maskinlæringsmodell kan man bruke beregnede tall og andre tilgjengelige opplysninger i A-ordningen, Landbruksregistre, Landbruksundersøkelsen/Landbrukstelling. Når maskinlæringsmodellen er tatt i bruk, kan mindre enheter undersøkes. Man kan spørre noen av de som har opplysninger i A-ordningen for å oppdatere modellen, etterpå spørre noen i undersøkelsen (de som er representative for den delen som er utenfor A-ordningen, men i populasjonen) og etterpå estimere for resten av populasjon. Da skal populasjonen bestå av tre deler: A-ordningen, undersøkelse og predikerte tall (Figur 6.2).

Figur 6.2 Populasjonen

For å få mer data for enheter som er utenom A-ordningen kan man se på opplysninger om lønn fra Selvgivelsen for en del av $\bar{U}$ populasjon utenfor utvalget.

$$L_{i,\text{selv}} = \gamma T_i + \varepsilon,$$

hvor $L_{i,\text{selv}}$ lønn fra Selvgivelsen er kjent og T_i relevante arbeidstid er ukjent.

$$L_{j,AO} = \beta T_j + \epsilon,$$

hvor lønn $L_{j,AO}$ fra A-ordningen er kjent og T_j relevante arbeidstid er kjent.

Mens regner at $\gamma \approx \beta$ kan en få T_i å være kjent.

Vedlegg A: Ratemodellen er en beskrivelse av data i Landbruksstatistikken

Vi skal dele dette kapittelet i to der vi i første del viser noen egenskaper til ratemodellen dersom vi antar at feilleddene er uavhengige og normalfordelte, mens vi i del to gir en argumentasjon for at sammenhengen mellom timeverk og lønn i A-ordningen beskrives realistisk ved bruk av en ratemodell der vi med god grunn kan anta at feilleddene er uavhengige og normalfordelte slik vi har brukt i første del av kapittelet.

Vedlegg A.1. Ratemodell med uavhengige og normalfordelte feilledd

Vi skal først gi et formelt svar og deretter gi en begrunnelse for når data vil følge en ratemodell. Vi tar utgangspunkt i følgende modell:

$$(1) Y_i = \beta \cdot x_i + \varepsilon_i \quad \text{der } \varepsilon_i \stackrel{\text{u.f.}}{\sim} N(0, \tau^2 \cdot x_i), \text{ for } i = 1, 2, 3, \dots, n$$

I formel (1) antar vi at:

$$\varepsilon_i \stackrel{\text{u.f.}}{\sim} N(0, \tau^2 \cdot x_i) \text{ betyr } \varepsilon_i \text{ er uavhengige og normalfordelte} \\ \text{med forventning lik } 0 [= E(\varepsilon_i)] \text{ og var i ens lik } \tau^2 \cdot x_i [= \text{Var}(\varepsilon_i)]$$

Det neste trinnet i prosessen er å finne estimatorer for parameteren β i modellen som definert i formel (1), det skjer på følgende vis:

$$(2) \text{ Minimer kvadratuttrykket: } Q(\beta) = \sum_{i=1}^n \frac{(Y_i - \beta \cdot x_i)^2}{x_i} = \sum_{i=1}^n \frac{\varepsilon_i^2}{x_i}$$

Merk at uttrykket i telleren i den siste summen er uavhengige og at forventet verdi til leddene i summen er like (τ^2), siden vi har:

$$E\left(\frac{\varepsilon_i^2}{x_i}\right) = \frac{E(\varepsilon_i^2)}{x_i} = \frac{\text{Var}(\varepsilon_i^2)}{x_i} = \frac{\tau^2 \cdot x_i}{x_i} = \tau^2$$

Vi minimerer kvadratuttrykket i formel (2) med hensyn på β ved å derivere med hensyn på β for å finne en estimator:

$$(3) Q'(\beta) = \sum_{i=1}^n \frac{2(Y_i - \beta \cdot x_i)}{x_i} (-x_i) = -2 \cdot \sum_{i=1}^n (Y_i - \beta \cdot x_i) = -2[\sum_{i=1}^n Y_i - \beta \cdot \sum_{i=1}^n x_i]$$

Vi finner minimum av Q ved å sette den deriverte lik 0:

$$Q'(\beta) = 0 \Leftrightarrow \sum_{i=1}^n Y_i - \beta \cdot \sum_{i=1}^n x_i = 0 \Leftrightarrow \beta = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n x_i}$$

Denne verdien av parameteren β som minimerer Q kaller vi estimatoren og betegner den $\hat{\beta}$:

$$(4) \hat{\beta} = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n x_i} = \frac{\sum_{i=1}^n [\beta \cdot x_i + \varepsilon_i]}{\sum_{i=1}^n x_i} = \beta + \frac{\sum_{i=1}^n \varepsilon_i}{\sum_{i=1}^n x_i}$$

Det neste er å finne forventning og varians til estimatorene:

$$(5) \begin{cases} E(\hat{\beta}) = \beta \\ Var(\hat{\beta}) = \frac{\tau^2}{\sum_{i=1}^n x_i} \end{cases}$$

Først beregner vi forventningen til estimatoren ved å vise disse resultatene fra høyre side av formel (4):

$$E(\hat{\beta}) = E\left[\beta + \frac{\sum_{i=1}^n \varepsilon_i}{\sum_{i=1}^n x_i}\right] = E(\beta) + \frac{\sum_{i=1}^n E(\varepsilon_i)}{\sum_{i=1}^n x_i} = \beta + \frac{\sum_{i=1}^n 0}{\sum_{i=1}^n x_i} = \beta$$

Litt mer utfordrende å finne variansen, men vi tar tak i samme uttrykk som for forventningen:

$$\begin{aligned} \sigma_{\hat{\beta}}^2 = Var(\hat{\beta}) &= Var\left[\beta + \frac{\sum_{i=1}^n \varepsilon_i}{\sum_{i=1}^n x_i}\right] = Var\left[\frac{\sum_{i=1}^n \varepsilon_i}{\sum_{i=1}^n x_i}\right] = \frac{\sum_{i=1}^n Var(\varepsilon_i)}{[\sum_{i=1}^n x_i]^2} \\ &= \frac{\sum_{i=1}^n (\tau^2 \cdot x_i)}{[\sum_{i=1}^n x_i]^2} = \frac{\tau^2 \cdot [\sum_{i=1}^n x_i]}{[\sum_{i=1}^n x_i]^2} = \frac{\tau^2}{\sum_{i=1}^n x_i} \end{aligned}$$

Vi må nå finne en estimator for τ^2 fordi da har vi også en estimator for variansen til $\hat{\beta}$ og foreslår:

$$(6) \hat{\tau}^2 = \frac{1}{n-1} \sum_{i=1}^n \frac{(Y_i - \hat{\beta} \cdot x_i)^2}{x_i}$$

Før vi kan finne forventningen til denne estimatoren må vi regne litt videre på uttrykket til høyre i formel (6):

$$\begin{aligned} \left\{ \begin{aligned} \sum_{i=1}^n \frac{(Y_i - \hat{\beta} x_i)^2}{x_i} &= \sum_{i=1}^n \frac{[Y_i - \beta x_i - (\hat{\beta} - \beta) \cdot x_i]^2}{x_i} = \\ \sum_{i=1}^n \frac{(Y_i - \beta x_i)^2 - 2 \cdot (Y_i - \beta x_i) \cdot (\hat{\beta} - \beta) \cdot x_i + (\hat{\beta} - \beta)^2 \cdot x_i^2}{x_i} &= \\ \sum_{i=1}^n \frac{(Y_i - \beta x_i)^2}{x_i} - 2 \cdot \sum_{i=1}^n \frac{(Y_i - \beta x_i) \cdot (\hat{\beta} - \beta) \cdot x_i}{x_i} + \sum_{i=1}^n \frac{(\hat{\beta} - \beta)^2 \cdot x_i^2}{x_i} &= \\ \sum_{i=1}^n \frac{(Y_i - \beta x_i)^2}{x_i} - 2 \cdot (\hat{\beta} - \beta) \cdot \left[\sum_{i=1}^n Y_i - \beta \cdot \sum_{i=1}^n x_i \right] + (\hat{\beta} - \beta)^2 \cdot \sum_{i=1}^n x_i &= \\ \sum_{i=1}^n \frac{(Y_i - \beta x_i)^2}{x_i} - 2 \cdot (\hat{\beta} - \beta) \cdot \left[\hat{\beta} \cdot \sum_{i=1}^n x_i - \beta \cdot \sum_{i=1}^n x_i \right] + (\hat{\beta} - \beta)^2 \cdot \sum_{i=1}^n x_i &= \\ \sum_{i=1}^n \frac{\varepsilon_i^2}{x_i} - (\hat{\beta} - \beta)^2 \cdot \sum_{i=1}^n x_i \Leftrightarrow \hat{\tau}^2 = \frac{1}{n-1} \left[\sum_{i=1}^n \frac{\varepsilon_i^2}{x_i} - (\hat{\beta} - \beta)^2 \cdot \sum_{i=1}^n x_i \right] \end{aligned} \right. \end{aligned}$$

Etter denne mellomregningen er vi klar for å vise at estimatoren er forventningsrett:

$$(7) \begin{cases} E(\hat{\tau}^2) = E\left\{\frac{1}{n-1} \left[\sum_{i=1}^n \frac{\varepsilon_i^2}{x_i} - (\hat{\beta} - \beta)^2 \cdot \sum_{i=1}^n x_i \right]\right\} = \frac{1}{n-1} \left\{ \sum_{i=1}^n \frac{E(\varepsilon_i^2)}{x_i} - \left[\sum_{i=1}^n x_i \right] \cdot E[(\hat{\beta} - \beta)^2] \right\} \\ = \frac{1}{n-1} \left\{ \sum_{i=1}^n \frac{Var(\varepsilon_i)}{x_i} - \left[\sum_{i=1}^n x_i \right] \cdot Var(\hat{\beta}) \right\} = \frac{1}{n-1} \left\{ \sum_{i=1}^n \frac{\tau^2 \cdot x_i}{x_i} - \left[\sum_{i=1}^n x_i \right] \cdot \frac{\tau^2}{\sum_{i=1}^n x_i} \right\} = \tau^2 \end{cases}$$

Vedlegg A.2. Ratemodellen i avsnitt 3.1 gir en riktig forståelse av sammenhengen mellom lønn og timer

I kapittel 3 beskrev vi at arbeidsgivere hver måned skal melde inn timeverk og lønn for alle personer som utfører arbeid i siste måned som de mottar lønn for, dette gjelder også alt arbeid på landbruksbedriften som det utbetales lønn for og som det er brukeren, samboer/ektefelle eller driftsleder som er ansvarlig arbeidsgiver for. For hver arbeidstaker som mottar lønn for utført arbeid skal dette dokumenteres på en A-melding som sendes samlet for alle arbeidstakere for hver måned fra arbeidsgiver. For en arbeidsgiver A kan vi da tenke oss at lønn og timeverk (egentlig hvor mange timer 100-prosent stilling svarer til og hvilken prosent av full stilling svarer til innfører vi en betegnelse på disse. Vi ser imidlertid på antall dager arbeidstakeren har utført arbeid på først, derfor er lønna merket med en DS for antall dager:

$$\begin{cases} L_{A,P}^D - \text{Lønn utbetalt til person } P \text{ fra arbeidsgiver } A \\ D_{A,P} - \text{Dager arbeidet av person } P \text{ for arbeidsgiver } A \end{cases}$$

Disse to størrelsene er et resultat av at personer arbeider et visst antall dager D for arbeidsgiveren i en måned, altså kan vi tenke oss at både lønn og timer som rapporteres er en sum av lønn og dager arbeidsforholdet er aktivt i en måned:

$$\begin{cases} L_{A,P}^D = \sum_{d=1}^{D_{A,P}} L_{A,P}^d \\ D_{A,P} = \sum_{d=1}^{D_{A,P}} 1 \end{cases}$$

Dersom vi forenkler litt og antar at timetallet er det samme hver dag (omtrent det samme), men innholdet i arbeidet kan veksle og det kan medføre at lønna varierer fra dag til dag så kan vi tenke oss at følgende enkle modell gir en riktig beskrivelse av lønn person P mottar fra arbeidsgiver A for arbeidet for hver dag i måneden:

$$(8) \quad L_{A,P}^d = \beta_{A,P}^{L,D} + \eta_{A,P}^d \quad \text{der } E(\eta_{A,P}^d) = 0 \text{ og } Var(\eta_{A,P}^d) = \sigma_{A,P}^2 \quad \text{for } d = 1, 2, 3, \dots, D$$

Dette betyr at vi antar at lønna person P mottar hver dag er likt fordelt rundt forventningsverdien ($\beta_{A,P}^{L,D}$) med en ukjent varians (σ^2). Dette er en rimelig antagelse siden vi har gått ut fra at personen P arbeider omtrent like lenge hver dag. Hvis vi bruker sammenhengen mellom daglige data i formel 8 og de månedlige data i ballpunktet overfor og dersom vi antar at feilleddene i formel (8) er uavhengige:

$$\begin{cases} \text{Anta } L_{A,P}^D = \sum_{d=1}^{D_{A,P}} L_{A,P}^d, \text{ der hver } L_{A,P}^d \text{ er gitt ved formel (9).} \\ \text{Da er forventning og varians lik:} \\ E(L_{A,P}^D) = E\left[\sum_{d=1}^{D_{A,P}} L_{A,P}^d\right] = \sum_{d=1}^{D_{A,P}} E(\beta_{A,P}^{L,D} + \eta_{A,P}^d) = \\ = \sum_{d=1}^{D_{A,P}} [\beta_{A,P}^{L,D} + E(\eta_{A,P}^d)] = \sum_{d=1}^{D_{A,P}} (\beta_{A,P}^{L,D} + 0) = \beta_{A,P}^{L,D} \cdot D_{A,P} \\ Var(L_{A,P}^D) = Var\left[\sum_{d=1}^{D_{A,P}} L_{A,P}^d\right] = \sum_{d=1}^{D_{A,P}} Var(\beta_{A,P}^{L,D} + \eta_{A,P}^d) = \\ = \sum_{d=1}^{D_{A,P}} Var(\eta_{A,P}^d) = \sum_{d=1}^{D_{A,P}} \sigma_{A,P}^2 = \sigma_{A,P}^2 \cdot D_{A,P} \end{cases}$$

I stedet for dager vil vi ha inn timer og det gjør vi en enkel omregning:

$$\left\{ \begin{array}{l} \text{Sett antall timer for dag } d \text{ lik: } T_{A,P}^d \\ \text{Da er antall timer i måneden: } T_{A,P} = \sum_{d=1}^{D_{A,P}} T_{A,P}^d \\ \text{Altså: } T_{A,P} = \sum_{d=1}^{D_{A,P}} T_{A,P}^d = \sum_{d=1}^{D_{A,P}} T_{A,P}^d \cdot 1 = \left[\frac{1}{D_{A,P}} \sum_{d=1}^{D_{A,P}} T_{A,P}^d \right] \cdot \left[\sum_{d=1}^{D_{A,P}} 1 \right] = \bar{T}_{A,P} \cdot D_{A,P} \end{array} \right.$$

Det er egentlig helt opplagt at antall timer i måneden blir lik gjennomsnittlig antall timer per dag multiplisert med antall dager i måneden. Det fører til at forventning og varians kan skrives om i timer:

$$\left\{ \begin{array}{l} E(L_{A,P}^D) = \beta_{A,P}^{L,D} \cdot D_{A,P} = \frac{\beta_{A,P}^{L,D}}{\bar{T}_{A,P}} \cdot \bar{T}_{A,P} \cdot D_{A,P} = \beta_{A,P}^{L,T} \cdot T_{A,P} \quad \text{for } \beta_{A,P}^{L,T} = \frac{\beta_{A,P}^{L,D}}{\bar{T}_{A,P}} \text{ og } T_{A,P} = \bar{T}_{A,P} \cdot D_{A,P} \\ \text{Var}(L_{A,P}^D) = \sigma_{A,P}^2 \cdot D_{A,P} = \frac{\sigma_{A,P}^2}{\bar{T}_{A,P}} \cdot \bar{T}_{A,P} \cdot D_{A,P} = \tau_{A,P}^2 \cdot T_{A,P} \quad \text{for } \tau_{A,P}^2 = \frac{\sigma_{A,P}^2}{\bar{T}_{A,P}} \text{ og } T_{A,P} = \bar{T}_{A,P} \cdot D_{A,P} \end{array} \right.$$

Hvis vi kombinerer dette ballpunktet og definisjonen av modellen for $L_{A,P}^d$ i formel (8) og definisjonen av $L_{A,P}^D = \sum_{d=1}^{D_{A,P}} L_{A,P}^d$ får vi følgende ratemodell for $L_{A,P}^T = L_{A,P}^D$

$$(9) \quad L_{A,P}^T = \beta_{A,P}^{L,T} \cdot T_{A,P} + \varepsilon_{A,P}^T \quad \text{der } E(\varepsilon_{A,P}^T) = 0 \text{ og } \text{Var}(\varepsilon_{A,P}^T) = \sigma_{A,P}^2 \cdot P \\ T_{A,P} \quad \text{for personer } P$$

Med utgangspunkt i forrige ballpunkt definerer vi feilleddet:

$$\varepsilon_{A,P}^T = L_{A,P}^T - E(L_{A,P}^T)$$

Da regner vi ut forventning og varians til $\varepsilon_{A,P}^T$:

$$\left\{ \begin{array}{l} E(\varepsilon_{A,P}^T) = E[L_{A,P}^T - E(L_{A,P}^T)] = E(L_{A,P}^T) - E[E(L_{A,P}^T)] = E(L_{A,P}^T) - E(L_{A,P}^T) = 0 \\ \text{Var}(\varepsilon_{A,P}^T) = \text{Var}[L_{A,P}^T - E(L_{A,P}^T)] = \text{Var}(L_{A,P}^T) = \sigma_{A,P}^2 \cdot T_{A,P} \end{array} \right.$$

Da gjenstår det bare å argumentere for at feilleddene er uavhengige og normalfordelte. Hvert feilledd kan vi skrive som en sum av identiske fordelte størrelser definert i formel (8):

$$\begin{aligned} \varepsilon_{A,P}^T &= L_{A,P}^T - E(L_{A,P}^T) = \sum_{d=1}^{D_{A,P}} L_{A,P}^d - E\left[\sum_{d=1}^{D_{A,P}} L_{A,P}^d\right] = \sum_{d=1}^{D_{A,P}} [L_{A,P}^d - E(L_{A,P}^d)] = \\ &= \sum_{d=1}^{D_{A,P}} \eta_{A,P}^d = D_{A,P} \cdot \bar{\eta}_{A,P} \end{aligned}$$

Siden et gjennomsnitt i grensen er normalfordelt [$\bar{\eta}_{A,P} \xrightarrow{D_{A,P} \rightarrow \infty} N(0, \sigma_{A,P}^2)$], altså har vi at feilleddene er tilnærmet normalfordelte siden antall arbeidsdager i måneden er nærmere 25.

Konklusjonen er at antakelsen vi begynte med for den generelle ratemodellen i formel (1) om at feilleddene er identisk normalfordelte og uavhengige gjelder for ratemodellen i formel (9) for lønn forklart av antall timer for første del av antakelsen, nemlig normalfordeling. Uavhengighet og identisk fordeling følger av at vi antar at avvikene for hver person P er upåvirket av hverandre og at vi kan anta at hver person sin lønn og timer er rimelig like.

Ratemodellen kan på samme vis brukes for å analysere sammenhengen mellom timetallet som rapporteres fra jordbruksbedriften i Landbrukstelingen og Landbruksundersøkelsene og de alternative timetallene fra A-meldingene:

$$(10) \quad T_{jB,i}^{A\text{-ordn}} = \beta_{jB} \cdot T_{jB,i}^{LU} + \varepsilon_{jB,i} \quad \text{der } \varepsilon_{jB,i} \sim N(0, T_{jB,i}^{LU} \cdot \tau_{jB}^2)$$

Å undersøke om det er de samme timetallene som en får fra de to kildene svarer til at $\beta_{jB} = 1$. For å undersøke dette får vi bruk for variansen til estimatet til β_{jB} , se formel (5).

Vedlegg B: Forskjellige begreper i Landbruksstatistikken

Fast ansatt «menes at ansettelsen er løpende og tidsubegrenset, samtidig som lovens regler om opphør av arbeidsforhold gjelder, og at arbeidstaker sikres forutsigbarhet for arbeid i form av et reelt stillingsomfang» [fra arbeidsmiljøloven].

Deltager er relevant for jordbruksbedrifter organisert som DA eller ANS. Ofte er DA eller ANS basert på tidligere enkeltpersonforetak som har inngått et organisert samarbeid om hele eller deler av jordbruksdrifta. DA eller ANS kan også være en organisasjonsform for samboere.

Sporadiskhjelp er det samme som ikke fast ansatt.

Spørsmål om samboer/ektefelle er bare med på skjema til enkeltpersonforetak. Samboer/ektefelle er person som er definert i Husholdningsregistre i ektefelle- eller parfelt og har relevante sivilstand og parstatus.

Spørsmålene om familiemedlemmer er bare med på skjema til enkeltpersonforetak. Som familiemedlem regnes barna, barnebarna, foreldrene, søsken, besteforeldrene til bruker/innehaver samt barna, foreldrene og barnebarna til ektefellen eller samboer.

Figurregister

Figur 5.1	Spredningsdiagram total årsarbeidstid for A-ordningen 2017 (AO) vs. Landbruksundersøkelsen 2018 (LU).....	11
Figur 5.2	Spredningsdiagram total årsarbeidstid fra A-ordningen 2017(AO) vs årsarbeidstid Landbruksundersøkelsen 2018(LU) utan enheter som har lav arbeidstid i Landbruksundersøkelsen sammenliknet med lønn tilgjengelig fra A-ordningen	12
Figur 5.3	Spredningsdiagram årslønn fra A-ordningen 2017(AO) vs årsarbeidstid Landbruksundersøkelsen 2018(LU)) utan enheter som har lav arbeidstid i Landbruksundersøkelsen sammenliknet med lønn tilgjengelig fra A-ordningen	12
Figur 5.4	Spredningsdiagram total årsarbeidstid for familiemedlemmer frå A-ordningen 2017 (AO) vs årsarbeidstid Landbruksundersøkelsen 2018 i relevant gruppe (LU)	13
Figur 5.5	Spredningsdiagram total årsarbeidstid for samboer fra A-ordningen 2017 (AO) vs årsarbeidstid Landbruksundersøkelsen 2018 i relevant gruppe (LU).....	13
Figur 5.6	Spredningsdiagram total årsarbeidstid for ansatte som ikke er familiemedlemmer eller samboer/ektefelle, fra A-ordningen 2017 (AO) vs årsarbeidstid Landbruksundersøkelsen 2018 i relevant gruppe (LU).....	14
Figur 5.7	Q-Q plot for feil i predikeringer av total årsarbeidstid for ansatte som ikke er familiemedlemmer eller samboer, basert på årsarbeidstid og årslønn fra A-ordningen	15
Figur 5.8	Q-Q plot for standardisert feil i predikeringer av total årsarbeidstid for ansatte som ikke er familiemedlemmer eller samboer/ektefelle basert på årsarbeidstid fra A-ordningen.....	16
Figur 5.9	Spredningsdiagram for årsarbeidstid til samboer/ektefeller utenfor jordbruksbedriften fra Landbruksundersøkelsen vs. A-ordningen.....	18
Figur 6.1	Eksempel av spørsmål med anbefalte tall (AT)	19
Figur 6.2	Populasjonen.....	19

Tabellregister

Tabell 5.1	Antall jordbruksbedrifter i LU18 (timer i 2017) og AO17 (frtk_id_ssb = orgnr_delreg)	9
Tabell 5.2	Antall jordbruksbedrifter i 2018 og kobling til A-ordningen i forskjellig år	10
Tabell 5.3	Jordbruksbedrifter i A-ordningen og jordbruksbedrifter ikke i A-ordningen, fordelt etter totalt antall timer i Landbruksundersøkelsen 2018	10
Tabell 5.4	Deskriptiv statistikk og Pearson korrelasjon for total årsarbeidstid for A-ordningen 2017 (AO) vs. Landbruksundersøkelsen 2018 (LU)	11
Tabell 5.5	Deskriptiv statistikk og Pearson korrelasjon for total årsarbeidstid fra A-ordningen 2017(AO) vs årsarbeidstid Landbruksundersøkelsen 2018(LU) uten enheter som har lav arbeidstid i Landbruksundersøkelsen sammenliknet med lønn tilgjengelig fra A-ordningen	12
Tabell 5.6	Deskriptiv statistikk og Pearson korrelasjon for årslønn fra A-ordningen 2017(AO) vs årsarbeidstid Landbruksundersøkelsen 2018(LU)) uten enheter som har lav arbeidstid i Landbruksundersøkelsen sammenliknet med lønn tilgjengelig fra A-ordningen	12
Tabell 5.7	Deskriptiv statistikk og Pearson korrelasjon for total årsarbeidstid for familiemedlemmer fra A-ordningen 2017 (AO) vs årsarbeidstid Landbruksundersøkelsen 2018 i relevant gruppe (LU)	13
Tabell 5.8	Deskriptiv statistikk og Pearson korrelasjon for total årsarbeidstid for familiemedlemmer fra A-ordningen 2017 (AO) vs årsarbeidstid Landbruksundersøkelsen 2018 i relevant gruppe (LU)	14
Tabell 5.9	Deskriptiv statistikk og Pearson korrelasjon for total årsarbeidstid for ansatte som ikke er familiemedlemmer eller samboer/ektefelle, fra A-ordningen 2017 (AO) vs årsarbeidstid Landbruksundersøkelsen 2018 i relevant gruppe (LU)	14
Tabell 5.10	Estimat for total årsarbeidstid for ansatte som ikke er familiemedlemmer, samboer/ektefeller beregnet fra Landbruksundersøkelsen og kombinasjonen av landbruksundersøkelsen med A-ordningen med bruk av vektorer fra Landbruksundersøkelsen	16
Tabell 5.11	Estimat for total årsarbeidstid etter driftsformen for ansatte som ikke er familiemedlemmer, samboer/ektefelle beregnet fra Landbruksundersøkelsen og kombinasjonen av landbruksundersøkelsen med A-ordningen med bruk av vektorer fra Landbruksundersøkelsen	17
Tabell 5.12.	Antall samboere/ektefeller som er ansatt utenfor jordbruksbedriften	17
Tabell 5.13	Antall brukere/innehavere som er ansatt utenfor jordbruksbedriften	18