

Ændringer i AKU-opregningen 2019

Baggrund Arbejdskraftundersøgelsen (AKU) er en stikprøvebaseret interviewundersøgelse af den danske befolkning i alderen 15-74 år. Resultaterne fra interviewene bliver derfor opregnet for at modsvare antallet af 15-74-årige i befolkningen. Forskellige grupper af personer udtrukket til AKU svarer ikke i lige stort omfang. Derfor opdeler vi den udtrukne stikprøve i grupper vha. forskellige registeroplysninger, og disse grupper tildeles forskellig opregningsvægt. Herved sikres, at grupperne får den størrelse i AKU, som de udgør i virkeligheden. Grupper, der har lavere svarprocenter, vægtes derfor højere end grupper med højere svarprocenter, og på denne måde søges det skæve bortfald korrigeret.

Vore analyser har vist, at personer med registreret lønmodtagerjob hidtil har været systematisk overvurderet i AKU, mens grupper uden for arbejdsstyrken har været undervurderet. Det skyldes, at personer med lønmodtagerjob har højere svarprocenter end eksempelvis mange grupper, som ikke er på arbejdsmarkedet.

En del af arbejdsmarkedsoplysningerne i den hidtidige opregning er registeroplysninger, der er cirka to år gamle på referencetidspunktet. Imidlertid er der nu mere aktuelle oplysninger til rådighed via eksempelvis Beskæftigelse for lønmodtagere (BfL). Dette og andre registre inddrages derfor i den nye opregning, således at det aktuelle befolkningsantal søges ramt for så mange relevante grupper som muligt.

Den hidtidige opregning Opregningen i AKU er senest blevet ændret i 2011 og justeret i 2015, hvor den blev gjort panelvis, således at hvert enkelt panel regnes op til befolkningen. Arbejdskraftundersøgelsen har en panelstruktur, som betyder, at hver respondent udtrækkes til at deltage fire gange. Panelerne udskiftes løbende, hvilket vil sige, at i et givent kvartal vil en fjerdedel deltage første gang, en fjerdedel anden gang osv. Ved panelvis opregning opregnes hvert af disse paneler selvstændigt.

Fra 2011 har følgende fem hovedkilder været anvendt i opregningen: Befolkningsregisteret, Personindkomstregisteret, Registerbaseret Arbejdsstyrkestatistik (RAS), Uddannelsesregisteret og den registrerede ledighed. Respondenterne er blevet grupperet i henhold til udvalgte kombinationer af køn, alder, socioøkonomisk status, antal børn, indkomst, region, statsborgerskab og nettoledighed, og tildelt en vægt, så antallet i gruppen så vidt muligt har ramt antallet i befolkningen mht. disse parametre.

Både hvad angår RAS, Uddannelsesregistret og Personindkomstregistret er oplysningerne imidlertid mellem 1 ¾ og 2 ½ år gamle på opregningstidspunktet, da de ikke findes nyere, og oplysningerne kan dermed være forældede i forhold til respondenternes aktuelle situation. Dette er særligt problematisk hvad angår socioøkonomisk status fra RAS, fordi respondenterne bl.a. kan have skiftet arbejdsmarkedsstatus i løbet af denne periode. Særligt kan dette gøre sig gældende for ledige og unge, som er grupper, hvis situation ofte ændrer sig.

I den hidtidige opregning har der således ikke været anvendt aktuell arbejdsmarkedsstatus fra registrene (udover registerledighed). Dette har medført en bias. Eksempelvis er BfL-populationen, som - vore analyser har vist - er stærkt korreleret med at være lønmodtager i AKU, blevet overvurderet. Dette hænger sammen med, at BfL-populationen har en højere svarprocent end gennemsnittet. SU-modtagere og folkepensionister har også været overvurderet, mens ikke-ledige kontanthjælpsmodtagere og flere andre grupper uden for arbejdsstyrken som nævnt er blevet undervurderet. Bias mht. disse grupper i befolkningen har varieret fra kvartal til kvartal og i forskellige aldersgrupper.

Ny hjælpeinformation i opregningen

Siden den hidtidige opregning blev udviklet i 2011, er BfL blevet tilstrækkelig aktuelt til at kunne anvendes i AKU's opregning. Dette register indeholder 60 pct. af den population, der indgår i AKU (15-74-årige) og to tredjedele af befolkningen i den erhvervsaktive alder (15-64-årige). En stor del af AKU's respondenter i denne aldersgruppe er således registreret med et lønmodtagerjob i BfL. Dermed kan den aktuelle arbejdsmarkedsstatus i referenceugen for denne store gruppe anvendes i opregningen.

Endvidere inddrages oplysninger om overførselsindkomster fra eIndkomstregistret og oplysninger om ejerskab af enkeltmandsvirksomheder og interessentskaber fra Danmarks Statistiks erhvervsregister i den nye opregning. Fra ledighedsstatistikken fås oplysninger vedrørende dagpenge- og kontanthjælpsmodtagere i referenceugen, hvor der hidtil har været anvendt antal ledighedsuger i kvartalet. I alt er over 90 pct. af respondenterne i AKU registreret i et eller flere af de nævnte aktuelle registre.

En række af registeroplysningerne har en stærk korrelation med arbejdsmarkedsstatus i AKU. Således viser vore analyser, at ca. 95 pct. af personerne, som er registreret med enten et lønmodtagerjob eller en enkeltmandsvirksomhed eller interessentskab, er beskæftiget i AKU. Tilsvarende gælder den anden vej, idet ca. 95 pct. af de beskæftigede i AKU er registreret med enten lønmodtagerjob eller en af de nævnte virksomhedstyper. Af nettoledige dagpengemodtagere er cirka 75 pct. også AKU-ledige. Også blandt nogle grupper uden for arbejdsstyrken er der stor sammenhæng. Registrerede efterlønsmodtagere og førtidspensionister uden registreret lønmodtagerjob (dvs. ikke i BfL) vil i over 90 pct. af tilfældene være uden for arbejdsstyrken i AKU. Den stærke sammenhæng betyder, at bias reduceres betydeligt, og at kvaliteten af især tallene for beskæftigelse og uden for arbejdsstyrken forbedres, når de aktuelle registeroplysninger bruges som hjælpeinformation ved opregningen. Dette skyldes, at vi ved at sikre, at antallet af beskæftigede – samt de andre nævnte arbejdsmarkedsstatusser i AKU – bedre rammer de antal, som de udgør i befolkningen, korrigerer for den forskellige svartilbøjelighed i disse forskellige grupper. Når vi fx sørger for at ramme antallet af registrerede lønmodtagere, fjerner vi den overvurdering af lønmodtagergruppen, som deres højere svartilbøjelighed hidtil har medført.

Registerbaseret arbejdsmarkedsstatus i den nye opregning

De nye kilder anvendes til at konstruere en variabel, der måler registerbaseret arbejdsmarkedsstatus i referenceugen. Den nye variabel indeholder 10 forskellige grupper og er dannet, så der i de enkelte grupper er en nogenlunde ensartet fordeling på beskæftigede, ledige og personer uden for arbejdsstyrken i AKU. For eksempel har grupperne 1 og 2 mange AKU-beskæftigede, mens fx 7 og 8

har mange uden for arbejdsstyrken. Andre grupper, fx 4, 5, 9 og 10 er mere blandede.

1. Fuldt ansvarlige deltagere i enkeltmandsvirksomheder og interessentskaber
2. Lønmodtagere
3. Modtagere af arbejdsløshedsdagpenge og andre a-kasse-ydelser
4. Andre bruttoledige samt ledighedsydelse
5. Ikke-ledige modtagere af kontanthjælpsydelse
6. Modtagere af syge- eller barseldagpenge
7. Modtagere af førtidspension, efterløn, fleksydelse, ressourceforløbsydelse eller revalideringsydelse
8. Modtagere af folkepension
9. Modtagere af Statens Uddannelsesstøtte
10. Personer uden registeroplysning

Oplysningerne hentes fra Beskæftigelse for lønmodtagere (2), ledighedsstatistikken (3-5), eIndkomstregisteret (4-9) og Erhvervsregisteret (1).

Den nye opregningsmodel Flere opregningsmodeller har været afprøvet, og variable er blevet forbedret i processen. Styrende for valget af netop denne model har været ønsker om:

1. En simpel model
2. Tilstrækkeligt med observationer i hver celle
3. Reduktion af bias og variationen i bias i forhold til arbejdsmarkedsstatus, hvilket gøres ved at genskabe registrenes arbejdsmarkedsstatus i populationen
4. Genskabe fordelingen på køn kombineret med alder i populationen
5. Genskabe region-fordeling i populationen
6. Særlig hensyn til de unge, hvilket bl.a. sker ved at genskabe fordelingen på alder kombineret med uddannelsesniveau i populationen
7. Panelvis opregning

Det er ikke fundet hensigtsmæssigt at tage flere variable med end de i pkt. 1-7 nævnte, og modellen indeholder således følgende variabler:

1. Køn
2. Alder
3. Bopælsregion
4. Uddannelse
5. Arbejdsmarkedsstatus i registrene

Inddelingerne efter køn, alder, uddannelse og bopælsregion er uændret i forhold til den hidtidige model. Dvs. at alder indgår to gange i modellen: i 11 kategorier kombineret med køn og i 6 kategorier kombineret med uddannelse (se modelspecifikation bilag 1). Uddannelsesgrupperingen er dog gjort ensartet igennem hele perioden fra 1. kvartal 2008.

Den registerbaserede arbejdsmarkedsstatus i referenceugen, som er beskrevet ovenfor, er tilføjet opregningen. Den erstatter den socioøkonomiske status fra RAS samt indkomstoplysningerne. Det hidtidige anvendte antal ledighedsuger i kvartalet erstattes også af variabelen registerstatus i referenceugen. Endvidere udgår statsborgerskab og antal børn af opregningen, da vore analyser har vist, at disse oplysninger ikke længere bidrager kvalitativt til opregningen.

Bemærkninger til historiske tal Opregningsmodellen er gennemført tilbage til 1. kvartal 2008. Der er følgende mindre variationer:

- Fra 1. kvartal 2008 til 4. kvartal 2015 indgår også, om personen er flyttet, efter forskerbeskyttelse blev indført på flytteblanketten i kommunen (dette skete typisk i 2004). De personer, der ved flytning udfyldte Forskerbeskyttelsen, måtte ikke udtrækkes til AKU. Flytteoplysningen er taget ind for, at personer, der var flyttet, mens forskerbeskyttelsen var på flytteblanketten (og ikke havde ønsket forskerbeskyttelse), kan adskilles fra dem, som ikke var flyttet, og derfor ikke havde haft konkret anledning til at træffe valg om forskerbeskyttelse.
- Fra 1. kvartal 2008 til 2. kvartal 2010 opregnes ikke panelvis, da panelerne i denne periode har meget forskellig størrelse og sammensætning på grund af indfasning af ny panelstruktur og større stikprøve.

Bilag 1 Oversigt over modellerne

2019- opregningsmodel

1. Køn × Alder (11 grp.) × panel¹
2. Bopælsregion × panel
3. Alder (6 grp.) × Uddannelse × panel
4. Arbejdsmarkedsstatus i registrene × panel

2015- opregningsmodel

1. Køn × alder (11 grp.) × panel
2. Alder (6 grp.) × uddannelse × panel
3. Socioøkonomisk status × panel
4. Antal børn i husstand × panel
5. Statsborgerskab × panel
6. Nettoledighed × panel
7. Bruttoindkomst × panel
8. Flyttet × panel
9. Bopælsregion × panel

Bilag 2 Oversigt over grupperingen af hjælpeinformation

Variable i ny og hidtidig opregning

Alder (11 grp.)

15-19 år

20-24 år

25-29 år

¹ × Betyder at oplysningerne krydses, så fx hvert køn opregnes i de 11 aldersgrupper i hvert panel. Dvs. at antallet af mænd og kvinder får det antal de har i befolkningen i de 11 aldersgrupper. Dette gøres, så hvert af de fire paneler, som respondenterne tilhører, kommer til at udgøre en fjerdedel af befolkningen.

30-34 år
35-39 år
40-44 år
45-49 år
50-54 år
55-59 år
60-64 år
65-74 år

Køn

Mænd
Kvinder

Region

Region Hovedstaden
Region Sjælland
Region Syddanmark
Region Midtjylland
Region Nordjylland

Alder (6 grp.)

15-19 år
20-24 år
25-29 år
30-34 år
35-39 år
40-44 år
45-49 år
50-74 år

Uddannelse²

Uoplyst uddannelse eller Grundskole
Gymnasiale og Erhvervsfaglige uddannelser
Videregående uddannelser samt ph.d. og forskeruddannelser

Arbejdsmarkedsstatus i
registrene i ny opregning

Arbejdsmarkedsstatus

Fuldt ansvarlige deltagere i enkeltmandsvirksomheder og interessentska-
ber
Lønmodtagere
Modtagere af arbejdsløshedsdagpenge og andre a-kasse-ydelser
Andre bruttoledige samt ledighedsydelse
Ikke-ledige modtagere af kontanthjælpsydelse
Modtagere af syge- eller barseldagpenge
Modtagere af førtidspension, efterløn, fleksydelse, ressourceforløbsydelse
eller revalideringsydelse
Modtagere af folkepension
Modtagere af Statens Uddannelsesstøtte
Personer uden registeroplysning

² Uddannelsesniveaet er en kombination af højest fuldførte niveau og igangværende uddannelse, hvor det højeste niveau af de to bliver brugt.

Øvrige variable i 2015-
opregning

Socioøkonomisk status

Selvstændig og medhjælpende ægtefæller
Arbejdsløs og uoplyst
Lønmodtagere og Beskæftigede uden løn
Aktivering
Under uddannelse
Pensionister
Børn
Øvr. overførselsindkomst

Antal børn i husstanden

0 børn
1 børn
2 børn
Flere end 2 børn

Statsborgerskab

Dansk
Nordisk
Vestlig
Ikke-vestlig

Nettoledig (antal uger i kvartalet)

0 uger
1 uge
2 uger
3 uger
4 uger
5 uger
6 uger
7 uger
8 uger
9 uger
10 uger
11-13 uger

Skattepligtig bruttoindkomst

Under 100.000 kr.
100.000-200.000 kr.
200.000-300.000 kr.
Over 300.000 kr.

Flyttet

Flyttet efter forskerbeskyttelse blev indført på flytteblanketten³
Ikke flyttet efter forskerbeskyttelse blev indført på flytteblanketten

³ I de fleste kommuner blev forskningsbeskyttelse indført på flytteblanketten den 17. december 2004