

Ledighedsindikatoren - metodenotat

Indledning

Formålet med den nye hurtige ledighedsindikator er at kunne indikere den månedlige bruttoledighed, før datagrundlaget er komplet. Ledighedsindikatoren skal således ses som foreløbige tal, på det helt overordnede detaljeringsniveau, for den månedlige bruttoledighed. Den nye indikator offentliggøres månedligt og første gang 12. oktober med ledigheden for september.

I snart et år har DST fortløbende testet og optimeret metoden til beregning af ledighedsindikatoren, og resultaterne herfra tyder på, at vi som udgangspunkt kan ramme den endelige faktiske bruttoledighed (efter 60 dage), indenfor ± 1.000 fuldtidspersoner. Dette skal ses i lyset af, at den ordinære (officielle) månedlige bruttoledighedsstatistik til offentliggørelse ca. 30 dage efter reference-månedens afslutning typisk bliver revideret med op til ± 500 fuldtids-personer ved anden offentliggørelse/revision omkring 60 dage efter reference-månedens afslutning.

Hertil kommer dog, at der for enkelte af testmånederne er observeret større afvigelser mellem indikatoren sæsonkorrigeret, som er det tal der offentliggøres, og den efterfølgende bruttoledighed sæsonkorrigeret, hvor afvigelsen alene skyldes det faktum at førstnævnte dannes som en sum af to sæsonkorrigerede delserier, kontanthjælpsdelen og dagpengedelen, mens sidstnævnte dannes som en sum af 48 sæsonkorrigerede delserier for hhv. to køn, seks aldre og fire ledighedstyper.

Selvom estimationsmetoderne løbende er blevet optimeret over den seneste tid, vil DST fortsætte med at følge metodernes prædiktionssevner nøje for om muligt at forbedre disse yderligere. Senest til marts 2023 vil det blive vurderet, om den bagvedliggende metode kan optimeres yderligere.

Den samlede bruttoledighed består af henholdsvis dagpenge- og kontanthjælpsledige. Derfor kan den samlede ledighedsindikator opbygges på følgende vis:

Bruttoledighed = Dagpengemodtagere i alt
- Syge på dagpenge
- EØS-dagpengemodtagere
+ Kontanthjælpsledige

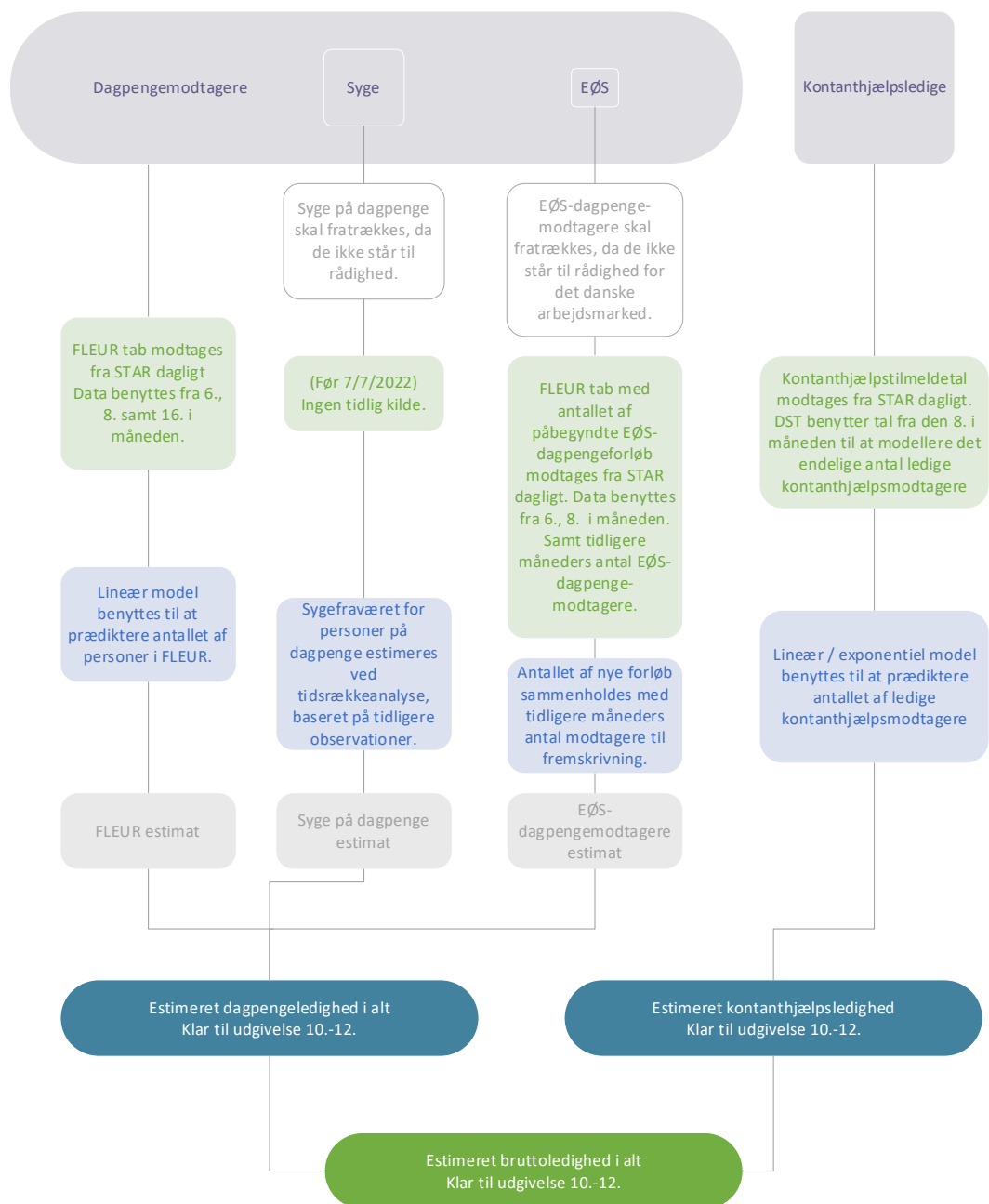
Her modtages FLEUR-optællinger på sumniveau for samtlige dagpengemodtagere dagligt fra STAR. Ud fra disse tidlige optællinger estimeres det endelige antal dagpengemodtagere for måneden. Optællingerne indeholder alle dagpengemodtagere, og dermed skal der fratrækkes de syge på dagpenge samt EØS-dagpengemodtagere, da disse ikke står til rådighed for det danske arbejdsmarked og dermed ikke er bruttoledige. Til indikatoren benyttes FLEUR-optællingerne fra den 6. i måneden efter referencemåneden.

For at estimere antallet af syge på dagpenge fremskrives tidsserien af de foregående måneders antal af syge på dagpenge. Fremover tænkes denne metode forbedret ved at benytte tidlige/foreløbige sygdomsmarkeringer fra STAR modtaget 7 dage efter reference-måneden.

Antallet af EØS-dagpengemodtagere for den aktuelle måned estimeres ved at fremskrive det faktiske antal fuldtidsmodtagere af EØS-dagpenge for måneden før med antallet af nystartede EØS-dagpengemodtagere for den aktuelle måned, i forhold til antallet af nystartede EØS-dagpengemodtagere i de foregående måneder. Her kendes det eksakte antal af fuldtidsmodtagere af EØS-dagpenge for måneden før fra seneste offentliggørelse af bruttoledigheden, mens de nystartede EØS-dagpengeforløb pr. måned fremgår af de daglige FLEUR-tab-indberetninger fra STAR.

Den sidste del af bruttoledigheden er de ledige kontanthjælpsmodtagere, som estimeres ud fra kontanthjælps-tilmeldetallene, der ligeledes modtages dagligt fra STAR. Til indikatoren benyttes optællingerne fra den 8. i måneden efter reference-måneden.

Notatet her indeholder et afsnit for hver af de ovennævnte fire dele.



Estimering af dagpengemodtagere i alt

Testresultaterne fra de seneste 11 måneder anvendes til at fastlægge den mest hensigtsmæssige metode til at få estimeret det samlede antal dagpengemodtagere. Datakilden er de daglige FLEUR-optællinger fra STAR. Til testen er benyttet optællingerne fra henholdsvis den 6. og 8. i måneden efter referencemånedens samt optællinger fra den 16. i måneden to måneder efter referencemånedens. Erfaringsmæssigt kan optællingerne for sidstnævnte målepunkt betragtes som endelige tal, da ændringer herefter vil være marginale. Herudover er ugedagen for den 6. i måneden medtaget, da dette kan have en effekt, fordi opdateringer af FLEUR ikke sker jævnt over ugens syv dage. Optællingerne er på A-kasse niveau, hvor A-kasser, der er blevet lagt sammen i løbet af referenceperioden januar 2019 og frem til nu, er sammenlagt for hele tidsserien. For nogle af referencemånederne findes der ingen FLEUR-tab og dermed intet data, og nogle steder er der fejl i optællingerne på A-kasseniveau. For hver mangel er det metodemæssigt besluttet, hvordan denne håndteres.

Ovenstående giver otte potentielle regressionsmodeller for estimering af det samlede antal dagpengemodtagere. Her fremkommer de otte modeller ved at kombinere 6./8. med a-kassefordelte-/totaloptællinger og med hensynstagen til hvilken ugedag der er tale om/eller ej. Til offentliggørelsen af ledighedsindikatoren for september den 12/10 anvendes modellen der benytter totaloptællingerne fra den 6. uden ugedag, da testen viste, at denne model i gennemsnit over de 11 måneder var mindst lige så god som de syv andre modeller. Det kontrolleres løbende, at den valgte model fortsat er den mest retvisende model. Dette tjekkes både med forudbestemte tidsintervaller samt i tilfælde af, at estimationen for en måned rammer meget skævt.

Estimering af syge på dagpenge

Dagpengemodtagere står ikke til rådighed for arbejdsmarkedet i den periode, de er syge. Derfor skal de i disse perioder ikke tælles med i bruttoledigheden.

Der er indtil nu ikke en tidlig kilde for sygdomsmarkeringer med stort nok sammenligningsgrundlag til at lave en regressionsanalyse. Derfor bruges tidligere måneder til at fremskrive det forventede antal syge dagpengemodtagere ved tidsrækkeanalyse. Til denne analyse anvendes data fra januar 2016 frem. Data fittes til en sæsonbetinget ARIMA model, hvis form og parametre fastsættes måned til måned. For manglende datapunkter og outliers imputeres værdier ved interpolation med hensynstagen til sæsonmønsteret. Fra tidsrækkemodellen forudsiges værdien af den næste måned i rækken. Den største udfordring ved denne metode er, at det ikke er muligt for modellen at opfange pludselige skift i data, eksempelvis da corona-smitten var høj i starten af 2022.

Modellen har haft store problemer i testperioden i starten af 2022, da det netop var i disse måneder Covid-19 blev "sluppet fri", og der var mange korttidssyge. Derfor forventes det snarest muligt at kunne forbedre metoden til estimation af de syge på dagpenge ved inddragelse af tidlige/foreløbige sygdomsmarkeringer fra STAR modtaget syv dage efter reference-månedens.

Estimering af EØS-dagpengemodtagere

Personer, der modtager EØS-dagpenge, indgår som en delmængde af samtlige dagpengemodtagere i FLEUR-registret. Da de personer, som modtager EØS-dagpenge, ikke står til rådighed for det danske arbejdsmarked, skal de ikke tælles med i den danske ledighed. Derfor er det vigtigt at estimere antallet af fuldtidsmodtagere af EØS-dagpenge for den aktuelle måned. Dette antal har svinget mellem 300 og 500 fuldtidsmodtagere pr. måned.

Da FLEUR-tab'en ikke indeholder tidlige løbende optællinger af bestanden af EØS-dagpengemodtagere, må denne bestand for den aktuelle måned estimeres ud fra 1) Den seneste produktion af det officielle månedlige bruttoledighedstal for måneden før, der indeholder det eksakte antal af fuldtidsmodtagere af EØS-dagpenge for måneden før, samt 2) De daglige FLEUR-tab-indberetninger fra STAR, der indeholder de nystartede EØS-dagpengeforløb pr. måned.

Selve estimeringen foretages ved at fremskrive det faktiske antal fuldtidsmodtagere af EØS-dagpenge for måneden før med antallet af nystartede EØS-dagpengemodtagere for den aktuelle måned *i forhold til* antallet af nystartede EØS-dagpengemodtagere i de foregående måneder. Herunder anvendes den tillægsinformation, at man højst kan modtage EØS-dagpenge i 3 måneder, samt at den gennemsnitlige varighed af EØS-dagpengeforløbene er 2,5 måneder.

Estimering af ledige kontanthjælpsmodtagere

De tidligste markeringer for antallet af ledige kontanthjælpsmodtagere modtages dagligt fra STAR i form af de såkaldte tilmeldetal, som beskriver antallet af personer, der har "ansøgt om kontanthjælp, som ledige". Det endelige antal af kontanthjælpsledige ligger altid lavere end disse tilmeldetal. Disse tal holdes op mod den efterfølgende registrerede bruttoledighed for kontanthjælpsmodtagerne ved regression.

Til test er der fittet to regressionsmodeller. En Gaussisk model samt en Exponentiel model. Når DST modtager tilmeldetallet fra STAR, indsættes værdien i de nævnte modeller, og det forventede antal bruttoledige kontanthjælpsmodtagere kan dermed beregnes. De to modeller er blevet testet sideløbende, og på baggrund af disse test er det besluttet at anvende den exponentielle model til offentliggørelsen af ledighedsindikatoren for september den 12. oktober.

Samlet bruttoledighed

Den samlede estimerede bruttoledighed er summen af den estimerede dagpengeledighed samt den estimerede kontanthjælpsledighed. Den estimerede dagpengeledighed består, som nævnt, af det estimerede antal personer på dagpenge, estimeret ved FLEUR-tab'en, fratrukket de syge på dagpenge samt EØS-dagpengemodtagerne.

Den faktiske bruttoledighed estimeret efter ca. 10 dage har ramt "korrekt" indenfor ± 1.000 fuldtidspersoner for samtlige ti måneder på nær januar-2022 og juli-2022. Førstnævnte større afvigelse skyldtes de mange små Corona-udbrud på et tidspunkt, hvor DST ikke havde de foreløbige sygdomsmarkeringer efter syv dage, og den sidstnævnte større afvigelse skyldtes en ændret sammenhæng mellem tilmeldetallene og den resulterende bruttoledighed for kontanthjælpsmodtagerne, som formodentlig hang sammen med meget sene "afmeldinger" af tilmeldte, nytilkomne ukrainere.

Sæsonkorrektur

Når dagpenge- og kontanthjælpsledigheden for den aktuelle måned er estimeret, sæsonkorrigeres serierne. Dette gøres ud fra to faktiske tidsserier bestående af de dagpenge- og kontanthjælpstal, der tidligere er offentliggjort i form af den officielle bruttoledighed, samt estimererne for de to dele for den aktuelle måned. Efterfølgende er det disse to sæsonkorrigerede ledighedstal samt deres sum, der udgives under navnet ledighedsindikator. Sæsonkorrektionen foretages indirekte på de to underserier via softwaren JDemetra+.