

Metod för länkning av AKU:s tidsserier för perioden 2005– 2020

20 oktober 2023

Innehåll

1	Bakgrund.....	5
1.1	Länkning med makroansats eller mikroansats	5
1.2	Tidigare länkningar av AKU	6
1.3	Nuvarande länkning av AKU	6
2	AKU:s design	8
3	Ändringar i AKU från och med januari 2021	9
3.1	Definitionsförändring	9
3.2	Populationsförändring	9
3.3	Blankettförändring	10
3.4	Förändrad hjälpinformation	10
4	Tillgänglig information	11
4.1	Dubbelmätning.....	11
4.2	Skattningar med ny och gammal hjälpinformation	11
4.3	Information om definitionsförändring	11
4.4	Information om populationsförändring	11
4.5	Flödesdata.....	12
4.6	Tidsserieanalys.....	12
5	Metod för skattning av tidsseriebrott.....	13
5.1	Ny hjälpinformation	13
5.2	Ny ramlag	14
5.3	Definitionsförändring	15
5.4	Populationsförändring	17
5.5	Kombinera dubbelmätning med skattning av definitionsförändring och populationsförändring	17
5.6	Förändring av säsongsmönster	18

5.7	Flödesanalyser och tidsserieanalys	18
6	Metod för länkning av tidsserier.....	20
6.1	Länkning med makroansats	20
6.2	Länkning med mikroansats	21
7	Skattningar av tidsseriebrott.....	23
7.1	Ny hjälpinformation	23
7.2	Ny ramlag	27
7.3	Definitionsförändring	29
7.4	Populationsförändring	29
7.5	Extra blanketteffekter	31
7.6	Förändring av säsongsmönster	35
7.7	Arbetade timmar och outnyttjat arbetskraftsutbud	35
7.8	Sammanfattning av brottskattningar	36
8	Länkning med mikroansats	38
8.1	Styrvillkor	38
8.2	Kvalitetssäkring av mikrolänkade tidsserier.....	42
8.2.1	Kvalitetssäkring av vikter.....	42
8.2.2	Kvalitetssäkring av mikrolänkade tidsserier som länkats på makronivå	43
8.2.3	Kvalitetssäkring av mikrolänkade tidsserier som inte länkats på makronivå	44
8.2.4	Kvalitetssäkring av tidsserier för arbetade timmar och outnyttjat arbetskraftsutbud.....	45
8.3	Tidsserier som ej är jämförbara över tid	46
8.3.1	Överenskommen arbetstid.....	46
8.3.2	Övertid.....	46
	Referenser	48
	Bilagor.....	49

A. Modell för skattning av definitionsförändring	49
A.1. Analys av personer som påverkas av definitionsförändringen	49
A.2. Klassificering i nya blanketten av de som påverkas av definitionsförändringen	54
A.3. Imputeringsmodell för definitionsförändringen	56
B. Modell för skattning av populationsförändring	58
B.1. Särskilda boenden.....	58
B.2. Värnpliktiga	64
B.3. Fängelseinterner	70
C. Modeller för latent arbetssökande.....	73
C.1. Brottskattningar för latent arbetssökande	73
C.2. Imputeringsmodell för latent arbetssökande	78
C.3. Imputeringsmodell för önskat antal arbetade timmar för latent arbetssökande.....	79

1 Bakgrund

AKU är sedan 1 januari 2021 anpassad till EU:s nya ramlag (förordning 2019/1700) för den sociala statistiken (EU, 2019). Den nya förordningen innebär bland annat förändringar avseende målpopulation, definitioner och blankettens utformning. Utöver de förändringar som den nya ramlagen medfört har även AKU:s hjälpinformation förändrats. Förändringen av hjälpinformationen implementerades samtidigt som införandet av den nya ramlagen i syfte att minimera antalet tidpunkter med tidsseriebrott.

För att säkerställa att AKU:s tidsserier är jämförbara över tid, så har SCB beslutat att ta fram länkade tidsserier för åren 2005–2020. Därigenom kommer AKU:s tidsserier att vara jämförbara tillbaka till januari 2005. Vid härledning av länkade tidsserier krävs dels skattning av tidsseriebrotten, dels justering av de ursprungliga tidsserierna för skattningarna av tidsseriebrotten.

1.1 Länkning med makroansats eller mikroansats

Länkning av tidsserier kan antingen genomföras med en makroansats eller med en mikroansats. Länkning med dessa båda ansatser kan genomföras på många olika sätt och vi ger här en kortfattad övergripande bild av de båda ansatserna.

Med makroansats avses att länkning görs för utvalda tidsserier och länkningen genomförs därmed på en aggregerad nivå. De utvalda tidsserierna justeras bakåt i tiden utifrån analys av tidsseriebrottens storlek. Att genomföra länkning med en makroansats är mindre resurskrävande eftersom enbart utvalda tidsserier länkas. Konsistens mellan dessa tidsserier erhålls inte per automatik utan måste säkerställas genom separata bearbetningar.

Med mikroansats avses att uppräkningsvikterna ändras på individnivå för urvalspersonerna. Utifrån dessa nya uppräkningsvikter beräknas därefter länkade tidsserier. Länkning med mikroansats är mer flexibel med avseende på vilka tidsserier som länkas, eftersom länkningen genomförs på individnivå, vilket medför att det på förhand inte är nödvändigt att besluta om vilka tidsserier som ska länkas. Att genomföra länkning med en mikroansats innebär att det i teorin är möjligt att ta fram alla tänkbara tidsserier. Därutöver erhålls konsistens mellan tidsserierna per automatik. Denna ansats är dock mer resurskrävande än vad som är fallet vid en makroansats. Det stora antalet länkade tidsserier som är möjliga att ta fram medför även att det är svårt att ha kontroll över kvaliteten för samtliga tidsserier.

1.2 Tidigare länknings av AKU

Länkning av tidsserier i AKU har tidigare genomförts vid flera olika tillfällen och både makroansatser och mikroansatser har då använts. Vi beskriver här kortfattat tre av dessa länknings av AKU:s tidsserier.

Vid omräkningen av AKU för tidsperioden 1987–1992 länkades ett hundratal viktiga tidsserier med en makroansats, se Statistics Sweden (2012). Därefter omräknades varje intervjuad persons uppräkningsvikt med de makrolänkade tidsserierna som styrvillkor vilket medförde att den slutliga länkningen baserades på en mikroansats. Metoden som användes vid denna omräkning är lämplig då generella förändringar genomförts, såsom vid ny undersökningsdesign eller nytt skattningsförfarande. I dessa fall är det inte specifika grupper som förändras på särskilda sätt.

Vid omräkningen av AKU för tidsperioden 1987–mars 2005 användes en mikroansats där de intervjuade personernas uppräkningsvikt ändrades genom att grupper som påverkades av de genomförda förändringarna identifierades, se Statistics Sweden (2013a). För att veta hur uppräkningsvikterna skulle förändras analyserades hur personerna förflyttades från en grupp till en annan till följd av genomförda förändringar av AKU. På detta sätt skapades grupper med *givers*, för vilka uppräkningsvikterna reducerades, och grupper med *receivers*, för vilka uppräkningsvikterna ökades. Länkningen genomfördes därmed med en mikroansats. Denna metod bygger på konstruktion av modeller som förklarar relationen mellan gamla och nya förutsättningar och är därför lämplig vid förändringar som påverkar specifika grupper, såsom nya definitioner, nytt formulär eller nya insamlingsmetoder.

Vid omräkningen av utvalda AKU-tidsserier för tidsperioden 1970–1986 användes en makroansats, se Statistics Sweden (2013b). För att skapa den slutliga omräkningen genomfördes korrigering i flera steg i syfte att korrigera för de förändringar som genomförts genom åren. De utvalda tidsserierna skapades utifrån åtta centrala variabler i AKU uppdelat på redovisningsgrupper efter kön och åldersgrupper (sammanlagt 144 tidsserier). Länkningen genomfördes på ett sådant sätt att summakonsistensen bevarades.

1.3 Nuvarande länkning av AKU

Arbetet med att ta fram länkade tidsserier för tidsperioden 2005–2020 har varit uppdelat i två faser. Först genomfördes en länkning med makroansats. Ett drygt tusental länkade tidsserier för de viktigaste variablerna och indelningarna i AKU publicerades i två etapper (februari 2022 respektive juni 2022). Därefter genomfördes en länkning med mikroansats, varvid de tidsserier som länkats på makronivå, tillsammans med olika modeller, användes för att justera uppräkningsvikterna för de svarande i AKU under åren 2005–2020. Med de justerade vikterna har länkade tidsserier för samtliga variabler och indelningar i

AKU kunnat tas fram. Dessa länkade tidsserier publiceras i samband med denna rapport i oktober 2023.

Denna rapport beskriver den metod som har använts för att härleda de slutliga länkade tidsserierna. Rapporten är indelad enligt följande. Kapitel 2 beskriver kortfattat AKU:s design, kapitel 3 beskriver förändringarna som genomförts i AKU i samband med införandet av den nya ramlagen och kapitel 4 beskriver den information som finns tillgänglig för att undersöka effekten av förändringarna. Kapitel 5 beskriver de metoder som har använts för att analysera tidsseriebrott och kapitel 6 de metoder som har använts för att omsätta skattningarna av tidsseriebrott till länkade tidsserier. Kapitel 7 redovisar sedan brottskattningarna för de viktigaste variablerna och Kapitel 8 härledningen av justerade uppräkningsvikter. Ett antal modeller som har använt i länkningen redovisas i tre avslutande bilagor (se bilaga A–C).

2 AKU:s design

Syftet med AKU är att beskriva aktuella arbetsmarknadsförhållanden för hela befolkningen som tillhör målpopulationen. AKU är den enda källan som löpande ger en sammanhållen bild av arbetsmarknaden (sysselsättningen, arbetslösheten, arbetade timmar m.m.).

AKU är en individbaserad urvalsundersökning som genomförs via telefonintervjuer varje månad under hela året. Månadsurvalet var 2021 ungefär 18 200 personer och är från och med januari 2022 ungefär 17 000 personer. Urvalspersonerna besvarar frågor om sin situation på arbetsmarknaden under en specifik vecka, kallad referensveckan, under referensmånaden. Upplägget är sådant att samtliga veckor under året undersöks. Resultaten av de månadsvisa undersökningarna publiceras kort efter referensmånadens slut. Dessa ligger också till grund för skattningar av kvartals- och årsgenomsnitt.

AKU är en panelundersökning med roterande urval där urvalspersonerna deltar en gång per kvartal vid totalt åtta tillfällen. Detta medför att 7/8 av urvalet återkommer med tre månaders mellanrum och att 1/8 av urvalet byts ut mot nya urvalspersoner varje kvartal.

3 Ändringar i AKU från och med januari 2021

Här beskrivs kortfattat de förändringar som införts i AKU från och med januari 2021. Mer detaljerad information finns på AKU:s produktsida¹.

3.1 Definitionsförändring

Den främsta definitionsförändringen avser sysselsatta som är frånvarande från arbetet. I synnerhet påverkas personer som av vissa frånvaroskäl är helt frånvarande från arbetet och där frånvaron förväntas vara i mer än tre månader. Detta gäller bland annat helt tjänstlediga utan lön, permitterade, frånvarande pga. arbetsbrist, uppsagda utan arbetsplikt, föräldralediga som inte är berättigade till föräldrapenning samt säsongsarbetande som inte utfört något arbete under innevarande säsong. Dessa personer exkluderas ur gruppen sysselsatta, varför definitionsförändringen leder till färre sysselsatta.

Förändringar har även genomförts för de frågor som används för att klassificera arbetslösa. Till skillnad från definitionsförändringen för sysselsatta är det svårt att på förhand avgöra hur dessa förändringar kommer att påverka skattningarna av antalet arbetslösa. Frågan om en respondent har sökt arbete har stärkts upp i syfte att tydliggöra vad som avses med frågan. Dessutom har frågan om hur man sökt arbete ändrats. Tidigare var detta en öppen fråga, men i den nya blanketten finns en egen fråga för varje möjligt söksätt. Om respondenten svarar ja på någon av söksättsfrågorna så ställs inga ytterligare frågor om söksätt och personen klassificeras som arbetssökande.

3.2 Populationsförändring

Enligt den nya ramlagen består målpopulationen endast av boende i privata hushåll. Det som exkluderas ur målpopulationen är personer som bor på så kallade institutionsboenden, såsom gruppboenden för funktionsnedsatta, fängelser och hem för sluten ungdomsvård. Även värnpliktiga exkluderas från målpopulationen. Tidigare bestod målpopulationen av samtliga folkbokförda personer och målpopulationen blir med den nya ramlagen därför mindre än tidigare.

SCB har inte någon förhandsinformation om vilka individer som bor i privata hushåll och denna grupp måste därför identifieras i samband med intervjun. Ett arbete har dock utförts under 2023 med syfte att skatta antalet personer som inte bor i privata hushåll från befintliga

¹ Övergripande om ändringarna i AKU i samband med ramlagen [ändringar-i-aku-i-samband-med-ramlagen 202102.pdf \(scb.se\)](#) (hämtad 2023-10-18).

register så att dessa skattningar sedan kan användas i hjälpinformationen i AKU, se SCB (2023).

Ytterligare en ändring som införts avseende populationen är att åldersgruppen utökats från 15–74 år till 15–89 år.

3.3 Blankettförändring

Den nya ramlagen har medfört att frågeordningen i blanketten har förändrats så att blanketten som används i AKU ska följa ett visst flöde. Denna förändring har bidragit till att möjligheten till så kallade beroende intervjuer nu är mer begränsade. Tidigare förlitade sig AKU:s blankett på beroende intervjuer med kontrollfrågor där respondenten ombads bekräfta tidigare uppgifter. I den nya blanketten är detta ej längre möjligt i samma utsträckning.

3.4 Förändrad hjälpinformation

Skattningarna i AKU baseras på en regressionsestimator som använder hjälpinformation från olika register. Hjälpinformationen består av variabler som identifierar viktiga redovisningsgrupper och/eller samvarierar med viktiga undersökningsvariabler och/eller svarssannolikheten.

Information om kön i kombination med ålder samt information om boenderegion respektive födelseland hämtas från registret över totalbefolkningen (RTB) och används som hjälpinformation i AKU.

Information från Arbetsförmedlingens sökanderegister (Af) används också i hjälpinformationen och är föremål för en av ändringarna som implementerats. Från och med januari 2021 inkluderas fler sökandekategorier i klassificeringen av arbetslösa i hjälpinformationen. Syftet med denna förändring är att förbättra samvariationen mellan arbetslösa i hjälpinformationen och arbetslösa i AKU.

Tidigare har information från registerbaserad arbetsmarknadsstatistik, RAMS, använts som hjälpinformation. Uppgifter från detta register har indelats i åtta grupper, sju grupper efter näringsgren och en grupp för de som inte finns med i RAMS. RAMS framställs på årsbasis och vid användning i AKU:s hjälpinformation har informationen i bästa fall varit 13 månader gammal och i sämsta fall 24 månader gammal. Från och med januari 2021 ersätts RAMS i hjälpinformationen av månatliga arbetsgivardeklarationer på individnivå, AGI, och dessa är alltid en månad gamla. Hjälpinformationen blir därmed mycket mer aktuell och samvariationen mellan sysselsatta i hjälpinformationen och sysselsatta i AKU förbättras. Vid införandet av AGI som hjälpinformation i AKU fanns inte tillgång till information om näringsgren kopplad till uppgifterna och därför är informationen från AGI inte uppdelad efter näringsgren. I stället indelas informationen från AGI efter åldersgrupp, eftersom olika svarsmönster har observerats för olika åldersgrupper.

4 Tillgänglig information

Analys av tidsseriebrott i AKU kan genomföras på många olika sätt. I de genomförda analyserna har sex huvudsakliga källor använts och de beskrivs kortfattat nedan.

4.1 Dubbelmätning

För samtliga månader under 2021 genomfördes en dubbelmätning där det månatliga urvalet på ungefär 18 200 urvalspersoner delades upp i två delar. Ena delen av dubbelmätningen utgjordes av 80 procent av urvalet och genomfördes med den nya blanketten medan den andra delen utgjordes av 20 procent av urvalet och genomfördes med den gamla blanketten. Detta innebär att insamling har skett med både den nya och den gamla blanketten under 2021 och motsvarande skattningar finns tillgängliga för årets alla månader.

4.2 Skattningar med ny och gammal hjälpinformation

Uppgifterna från AGI finns tillgängliga från och med januari 2019, men de har inte använts vid framställningen av AKU förrän från och med januari 2021. Detta medför att skattningar med ny hjälpinformation kan tas fram från och med februari 2019. Att det inte går att gå tillbaka till januari 2019 beror på att AGI används med en månads eftersläpning i AKU. För perioden februari 2019–december 2021 finns således skattningar med ny och gammal hjälpinformation med den gamla blanketten. Dessutom finns, från och med januari 2021, skattningar med ny och gammal hjälpinformation för den nya blanketten.

4.3 Information om definitionsförändring

För att kvantifiera effekten av definitionsändringen för sysselsatta som är frånvarande från arbetet genomfördes ett test mellan februari 2020 och december 2020. Inom ramen för detta test ställdes ytterligare frågor till de som klassificerades som sysselsatta men som uppfyllde villkoren för att klassificeras som arbetslösa eller ej i arbetskraften enligt den nya ramlagen. Dessa ytterligare frågor ställdes i slutet av intervjun och liknar frågorna i nya blanketten.

Dessutom går det att i den nya blanketten identifiera de som skulle ha klassificerats som sysselsatta enligt den gamla blanketten men som ej gör det enligt den nya blanketten.

4.4 Information om populationsförändring

Personer som inte bor i privata hushåll tillhör i allmänhet en av följande tre huvudgrupper:

- a. Personer som bor i särskilda boenden, såsom äldreboenden, servicehus eller gruppboendestäder
- b. Värnpliktiga eller kontraktsanställda soldater
- c. Fängelseinterner

Det finns även andra grupper som inte bor i privata hushåll, exempelvis personer som bor i kloster, men dessa grupper är avsevärt mindre till antalet än de tre huvudgrupperna och kan därför försummas här.

Antalet personer i de tre huvudgrupperna kan till viss del skattas med hjälp av befintliga register. Fastighetsregistret innehåller sedan 2012 adressuppgifter för särskilda boenden. Genom att samköra dessa adressuppgifter med folkbokföringsadressen i RTB, så kan vi sortera ut alla personer som var folkbokförda på ett särskilt boende i slutet av ett visst år. Inkomst- och taxeringsregistret (IoT) innehåller uppgifter om vilka personer som erhållit dagersättning för värnplikt eller grundläggande militär utbildning och antalet värnpliktiga kan skattas från denna information. För att skatta antalet fängelseinterner har vi erhållit årlig data från Kriminalvården.

4.5 Flödesdata

Eftersom AKU är en panelundersökning där urvalspersonerna deltar under åtta efterföljande kvartal går det att följa urvalspersonerna över tid. Denna design möjliggör analyser av flödesdata för att beskriva hur arbetskraftsstatus ändras från ett kvartal till nästa.

I syfte att analysera effekten av den nya ramlagen är det av särskilt intresse att jämföra flöden mellan kvartal 4 2020 och kvartal 1 2021 med de flöden som vanligtvis observeras. På så sätt är det möjligt att analysera skillnader mellan det gamla och det nya förfarandet.

4.6 Tidsserieanalys

Vid säsongrensningen av AKU:s tidsserier används SAS-makrot Proc X13 för att ta fram säsongrensade tidsserier och trendserier. Med hjälp av den automatiska outliersökningen i X13, så går det att undersöka om det finns belägg för att ett nivåskifte har inträffat i en viss tidsserie från och med en viss månad. Programmet skattar även storleken på ett eventuellt nivåskifte. Genom att applicera X13 på publicerade AKU-tidsserier går det alltså att skatta tidsseriebrottet i januari 2021. Vidare kan vi applicera X13 på de länkade tidsserier för att säkerställa att inga tidsseriebrott kvarstår i januari 2021 i de länkade tidsserierna.

5 Metod för skattning av tidsseriebrott

För att avgöra om den nya ramlagen och bytet av hjälpinformation har givit upphov till brott i AKU:s tidsserier har en så kallad differensmetod använts. För att beskriva differensmetoden i AKU:s kontext så inför vi först viss notation. Låt $X_t^{r,h}$ beteckna värdet på tidsserien X vid tiden t . Indexen r och h betecknar här ramlag respektive hjälpinformation och båda dessa index antar värden i mängden $\{g, n\}$, där g står för gammal och n står för ny. Detta ger totalt fyra varianter av varje tidsserie. Det totala tidsseriebrottet ges av väntevärdet av

$$X_t^{n,n} - X_t^{g,g},$$

där t är godtycklig månad under 2021. Eftersom tidsserier med ny respektive gammal ramlag och hjälpinformation kan ha olika säsongsmönster, så kan väntevärdet, och därmed storleken på tidsseriebrottet, vara olika för olika månader under året. På grund av svårigheten att skatta förändringar i säsongsmönstret (se även avsnitt 5.6) så betraktar vi här, något förenklat, väntevärdet som konstant under 2021.

Vi har tillgång till skattningar av $X_t^{g,g}$ och $X_t^{n,n}$ för alla 12 månader under 2021, men eftersom vi har data för gammal och ny hjälpinformation under en längre tidsperiod, så väljer vi att göra uppdelningen

$$X_t^{n,n} - X_t^{g,g} = \underbrace{(X_t^{n,n} - X_t^{g,n})}_{\text{ny ramlag}} + \underbrace{(X_t^{g,n} - X_t^{g,g})}_{\text{ny hjälpinfo}}.$$

Den första termen i högerledet motsvarar tidsseriebrottet till följd av nya ramlagen och den kan vi skatta med hjälp av dubbelmätningen under 2021. Den andra termen i högerledet motsvarar tidsseriebrottet till följd av bytet av hjälpinformation och den kan vi skatta med hjälp av skattningar med gammal och ny hjälpinformation.

5.1 Ny hjälpinformation

Vi skattar först storleken för det tidsseriebrott som uppkommit vid byte av hjälpinformation. Skattningen av detta brott betecknas $\hat{\Delta}^h$ och ges av

$$\hat{\Delta}^h = \sum_{t \in [T_0^h, T_1]} \alpha_t (X_t^{g,n} - X_t^{g,g}),$$

där T_0^h betecknar februari 2019 och T_1 december 2021. Koefficienterna α_t är icke-negativa tal som uppfyller villkoret $\sum_{t \in [T_0^h, T_1]} \alpha_t = 1$. Vid

beräkning av kvartalsmedelvärden och årsmedelvärden för AKU:s tidsserier viktas de två första månaderna i varje kvartal med 4 och den sista månaden i varje kvartal med 5². Vi använder vikter med dessa proportioner även i skattningen $\hat{\Delta}^h$. Med hjälp av grundläggande sannolikhetssteori går det att visa att variansen för brottskattningen minimeras om vikten för en given månad är omvänt proportionell mot variansen för den månads skattning. Eftersom urvalsstorleken för den gamla blanketten under 2021 är 1/5 av urvalsstorleken under 2019 och 2020 och variansen är proportionell mot urvalsstorleken, så multiplicerar vi vikterna för alla 2021 års månader med en faktor 1/5.

Eftersom $X_t^{g,n}$ och $X_t^{g,g}$ skattas baserat på samma stickprov, så kan vi använda det gängse skattningsprogrammet för att skatta medelfelet σ_t av $X_t^{g,n} - X_t^{g,g}$. Korrelationen mellan $X_t^{g,n} - X_t^{g,g}$ och $X_s^{g,n} - X_s^{g,g}$ för olika tidpunkter s och t är mycket svag för AKU:s tidsserier och detta gäller även då $s - t$ är en multipel av tre månader, dvs då urvalen vid tidpunkterna s och t delvis överlappar varandra. Följaktligen kan vi skatta variansen av brottskattningen $\hat{\Delta}^h$ som

$$\hat{V}^h \approx \sum_{t \in [T_0^h, T_1]} \alpha_t^2 \sigma_t^2 \approx 0.0362 \bar{V}_1 + 0.00073 \bar{V}_2 \approx 0.0399 \bar{V}_1,$$

där $\bar{V}_1$ är stickprovsmedelvärde av varianserna för $X_t^{g,n} - X_t^{g,g}$ under 2019–2020 och $\bar{V}_2$ är stickprovsmedelvärde av varianserna för $X_t^{g,n} - X_t^{g,g}$ under 2021. Sambandet längst till höger i ovanstående ekvation gäller eftersom $\bar{V}_2 \approx 5\bar{V}_1$ till följd av dubbelmätningens design.

5.2 Ny ramlag

Vi skattar härnäst storleken för det tidsseriebrott som uppkommit vid införande av den nya ramlagen. Skattningen av detta brott betecknas $\hat{\Delta}^r$ och ges av

$$\hat{\Delta}^r = \sum_{t \in [T_0^r, T_1]} \beta_t (X_t^{n,n} - X_t^{g,n}),$$

där T_0^r betecknar januari 2021 och T_1 december 2021. Koefficienterna β_t är icke-negativa tal som uppfyller villkoret $\sum_{t \in [T_0^r, T_1]} \beta_t = 1$. Vi låter vikterna β_t ha samma inbördes proportioner som vikterna α_t . Variansen $\hat{V}^r$ av brottskattningen $\hat{\Delta}^r$ ges då av

$$\hat{V}^r = \sum_{t \in [T_0^r, T_1]} \beta_t^2 \sigma_t^2 + \sum_{s, t \in [T_0^r, T_1]} \beta_s \beta_t \lambda_{s,t} \rho_{s,t} \sigma_s \sigma_t \approx \gamma^2 \bar{V},$$

där γ är en parameter som varierar mellan AKU:s olika tidsserier och $\bar{V}$ är stickprovsmedelvärde av varianserna av $\{X_t^{n,n} - X_t^{g,n}\}_{t \in [T_0^r, T_1]}$, dvs

² För år med 53 referensveckor viktas ytterligare en av årets månader med 5.

stickprovsmedelvärde av $\{(\sigma_t^{n,n})^2 + (\sigma_t^{g,n})^2\}_{t \in [T_0^r, T_1]}$. Vidare gäller det, i ekvationen ovan, att $\lambda_{s,t}$ är andelen av urvalet som överlappar mellan tiderna s och t samt att $\rho_{s,t}$ är korrelationen mellan $X_s^{n,n} - X_s^{g,n}$ och $X_t^{n,n} - X_t^{g,n}$. För att bestämma parametern γ har vi använt historiska värden som tidigare har beräknats för AKU:s tidsserier³.

Syftet med dubbelmätningen är att skatta den totala effekten av alla förändringar som gjorts vid införandet av den nya ramlagen. Denna totala effekt består dels av effekter som är direkt kvantifierbara, såsom definitionsförändringen för antalet sysselsatta, dels av effekter som inte är kvantifierbara, såsom förändrad frågeordning. Eftersom dubbelmätningen bara består av tolv månaders data med relativt små urvalsstorlekar, särskilt för den gamla blanketten, så är osäkerheten i brott-skattningarna $\hat{\Delta}^r$ mycket stor. Följaktligen är det mycket få tids-serier som uppvisar ett signifikant tidsseriebrott i dubbelmätningen.

5.3 Definitionsförändring

För att ändå kunna skatta ramlageffekten så har vi separat undersökt definitionsförändringen, vilken är den enda direkt kvantifierbara ramlageffekten. Enligt den nya definitionen av sysselsatta så kommer vissa personer som har varit frånvarande från sitt arbete i minst tre månader ej längre att klassas som sysselsatta. För att skatta antalet personer som påverkas av denna definitionsförändring, så kan vi använda månatliga skattningar av denna grupp baserat på de svar som inhämtats i den nya blanketten under 2021. Dessa uppgifter kan även användas för att skatta vilken arbetsmarknadsstatus de som påverkas av definitionsförändringen får i den nya blanketten.

Skattningar baserade på data från 2021 har dock följande brist. Eftersom de personer som påverkas av definitionsförändringen ej längre klassas som sysselsatta i den nya blanketten, så har dessa personer inte fått frågor som kan ge svar på vilken sysselsättningsstatus de hade haft i den gamla blanketten. Med andra ord kan vi inte avgöra vilken anknytningsgrad till arbetsmarknaden dessa personer hade haft med den gamla definitionen eller vilken sektor de verkade inom. För att kunna skatta definitionsförändringens påverkan på samtliga variabler behöver vi ha tillgång till sådan information och för att lösa detta problem finns två möjliga datakällor. Dels finns flödesdata som beskriver vilken sysselsättningsstatus de personer som under kvartal 1 2021 påverkades av

³ Följande värden på γ har använts. För antalet sysselsatta och antalet undersysselsatta är $\gamma = 0.451$, för antalet arbetslösa och antalet arbetslösa heltidsstuderande är $\gamma = 0.367$, för antalet i arbetskraften och antalet ej i arbetskraften samt för antalet latent arbetssökande och antalet i andra undergrupper till gruppen ej i arbetskraften är $\gamma = 0.452$, för antalet i befolkningen är $\gamma = 0.488$, för antalet sysselsatta i arbete är $\gamma = 0.377$, för antalet sysselsatta frånvarande är $\gamma = 0.342$, för antalet fast anställda är $\gamma = 0.473$, för antalet tidsbegränsat anställda är $\gamma = 0.423$, för antalet företagare och medhjälpande hushållsmedlemmar är $\gamma = 0.481$ och för antalet anställda, både totalt och uppdelat på sektorer, så är $\gamma = 0.459$.

definitionsförändringen hade under kvartal 4 2020. Dels finns svar från den extrafråga som ställdes i AKU under perioden februari–december 2020 för att uppskatta antalet personer som påverkades av definitionsförändringen. Båda dessa datakällor ger relativt likstämmiga resultat, men vi har valt att basera skattningarna helt och hållet på den sistnämnda källan eftersom den baseras på fler urvalspersoner och är mindre säsongsberoende. Dessutom finns det en risk att resultaten från flödesanalyserna förvrängs av personer som har bytt sysselsättningsstatus mellan kvartal 4 2020 och kvartal 1 2021.

Baserat på data från 2020–2021 kan vi således ta fram månatliga skattningar av definitionsförändringens påverkan på samtliga tidsserier i AKU. Viktade medelvärden av dessa månatliga skattningar (med vikterna β_t definierade i avsnitt 5.2) kan sedan användas som brott-skattningar $\hat{\Delta}^d$ för definitionsförändringen. Denna metodik användes för de länkade tidsserier som publicerades 2022. Vi har emellertid nu reviderat denna metodik eftersom skattningar av definitionsförändringen som baserades enbart på data från 2020–2021 gav alltför stora skattningarna av antalet personer som påverkas av definitionsförändringen (se nedan). Denna överskattning är förmodligen ett resultat av de speciella förhållanden som rådde under coronapandemin.

Den nya metodiken för att skatta definitionsförändringen baseras på det faktum att respondenternas svar på andra frågor i AKU, såsom frånvaroskäl, frånvarotidens längd och anknytningsgrad till arbetsmarknaden samt respondenternas ålder kraftigt påverkar sannolikheten att omfattas av definitionsförändringen. Baserat på svarsdata har vi därför konstruerat en modell (se bilaga A för detaljer) som skattar antalet personer som påverkas av definitionsförändringen.

Precis som med den tidigare metodiken används svarsdata från 2020 för att avgöra vilken arbetskraftsstatus personer som påverkas av definitionsförändringen hade i den gamla blanketten och svarsdata från 2021 för att avgöra vilken arbetskraftsstatus de får i den nya blanketten. Det är värt att notera att sannolikheten att påverkas av definitionsförändringen givet de svarsdata som används i modellen inte har något tydligt säsongsmönster. Skattningarna av antalet personer som påverkas av definitionsförändringen har likväl ett säsongsberoende eftersom de ärver ett säsongsmönster från vissa av variablerna som används i modellen, exempelvis frånvaroskäl. Detta är ytterligare en fördel jämfört med den gamla metodiken som krävde separata skattningar för sommar- och vintermånader.

För varje månad t av intresse och varje relevant tidsserie i AKU kan vi med hjälp av modellen och svarsdata från månad t ta fram en skattning $\hat{\Delta}_t^d$ för definitionsförändringens påverkan på den tidsserien under månad t . Under 2020 ger modellen och den tidigare metodiken liknande skattningar, men för perioden 2005–2015 skattar den nya modellen det

totala antalet personer som påverkas av definitionsförändringen ungefär 10 tusen personer lägre än den gamla metodiken.

5.4 Populationsförändring

Förändringen i AKU:s målpopulation innebär att enbart personer boende i privata hushåll numera räknas in i AKU-befolkningen. Personer som inte bor i privata hushåll tas dock inte bort ur urvalet eftersom SCB inte har registerinformation på individnivå om vilka som bor i privata hushåll.

Baserat på registeruppgifter från Fastighetsregistret, RTB, Inkomst- och Taxeringsregistret och Kriminalvården skattar vi från och med januari 2021 antalet personer som inte bor i privata hushåll. Vi skattar även hur många av dessa personer som tillhör de olika kategorierna i hjälpinformationen. Med andra ord skattar vi fördelningen för de som inte bor i privata hushåll efter kön, ålder, region, födelseland, sysselsättningsstatus enligt AGI samt arbetslöshetsstatus enligt Af. Med denna information kan vi justera totalerna i hjälpinformationen så att totalerna bara motsvarar de som bor i privata hushåll och sedan kalibrera AKU-skattningarna efter dessa justerade totaler.

För tidsperioden januari 2005–december 2020 finns bara delar av den registerinformation som beskrevs i föregående stycke tillgänglig. Utgående från den registerinformation som vi har tillgång till från denna tidsperiod har vi skapat en modell (se bilaga B för detaljer) för att kunna skatta antalet personer som inte bor i privata hushåll. För varje månad t av intresse och varje relevant variabel och indelning i AKU kan vi med hjälp av modellen ta fram en skattning $\hat{\Delta}_t^p$ av antalet personer under månad t som uppfyller villkoren för variabeln och indelningen och som inte bor i ett privat hushåll (med negativt tecken eftersom det motsvarar en minskning av populationen).

5.5 Kombinera dubbelmätning med skattning av definitionsförändring och populationsförändring

Skattningen $\hat{\Delta}^r$ som vi har tagit fram med hjälp av dubbelmätningen innefattar utöver rena blanketteffekter såväl definitionsförändringen som populationsförändringen. Eftersom de direkta skattningarna av definitionsförändringen och populationsförändringen är mycket mer tillförlitliga än skattningen $\hat{\Delta}^r$ (även om vi inte exakt kan kvantifiera osäkerheten i de förra skattningarna), så använder vi endast resultat från dubbelmätningen när det signifikant avviker från de direkta skattningarna av definitionsförändringen och populationsförändringen.

I praktiken använder vi följande metodik för att avgöra när skattningar från dubbelmätningen ska användas. Vi låter först $\hat{\Delta}^d$ och $\hat{\Delta}^p$ beteckna

medelvärde över 2020 av $\hat{\Delta}_t^d$ respektive $\hat{\Delta}_t^p$. Dubbelmätningen används inte om villkoret

$$\hat{\Delta}^r - 1.96\sqrt{\hat{V}^r} \leq \hat{\Delta}^d + \hat{\Delta}^p \leq \hat{\Delta}^r + 1.96\sqrt{\hat{V}^r},$$

är uppfyllt, dvs om den totala skattningen av definitionsförändringen och populationsförändringen inte avviker signifikant från resultatet av dubbelmätningen. Om ovanstående villkor inte är uppfyllt, så har vi lagt till en extra blanketteffekt $\hat{\Delta}^e$, så att villkoret

$$\hat{\Delta}^r - 1.96\sqrt{\hat{V}^r} \leq \hat{\Delta}^d + \hat{\Delta}^p + \hat{\Delta}^e \leq \hat{\Delta}^r + 1.96\sqrt{\hat{V}^r},$$

uppfylls. Det exakta värdet på $\hat{\Delta}^e$ har valts med hjälp av bland annat flöden och egenskaper hos de länkade tidsserierna. Metodiken för att välja $\hat{\Delta}^e$ beskrivs i detalj i avsnitt 7.5. Notera slutligen att för tidsserier som är små i absoluta tal eller som är av marginellt intresse, så har vi i vissa fall inte infört någon extra blanketteffekt även om det första villkoret ovan inte är uppfyllt.

5.6 Förändring av säsongsmönster

Såväl bytet av hjälpinformation som införandet av den nya ramlagen kan ge upphov till ett förändrat säsongsmönster. Eftersom dubbelmätningen och analysen av definitionsförändringen bara ger oss en mätning per månad, så kan vi inte särskilja en eventuell förändring i säsongsmönstret från slumpmässiga avvikelser. Vi är därför tvungna att försumma eventuella förändringar i säsongsmönstret till följd av blankettbytet.

För bytet av hjälpinformation är situationen något mer gynnsam eftersom vi där har tillgång till totalt 35 månaders data. Detta är marginellt mindre än de tre års data som normalt krävs för att skatta en förändring i säsongsmönstret. För att ändå försöka avgöra om säsongsmönstret påverkats av bytet av hjälpinformation, så har vi, för var och en av årets tolv månader, beräknat viktade medelvärden av de differenser $X_t^{g,n} - (X_t^{g,g} + \hat{\Delta}^h)$ som finns tillgängliga för den månaden. För ett antal av AKU:s variabler, såsom antalet sysselsatta, antalet i arbetskraften, antalet sysselsatta i arbete och antalet fast anställda ser det ut att finnas ett linjärt samband över året för dessa viktade medelvärden. Detta linjära samband kan dock inte observeras i data från den nya blanketten och vi har därför utfört ytterligare analyser för att avgöra om eventuella skillnader i säsongsmönster till följd av bytet av hjälpinformation ska tas med i de länkade tidsserierna. Dessa analyser finns beskrivna i detalj i avsnitt 7.6.

5.7 Flödesanalyser och tidsserieanalys

Som nämnts ovan har flödesdata analyserats med flera olika syften i arbetet med att skatta eventuella tidsseriebrott. I några fall har dessa analyser påvisat avvikelser som varken har kunnat förklaras av bytet av

hjälpinformation eller av förändringarna av sysselsättningsdefinitionen eller populationen. I dessa fall har vi gjort mindre justeringar av brott-skattningarna i enlighet med flödesanalysen. Även tidsserieanalys av publicerade AKU-tidsserier har använts i samma syfte. Justeringar baserade på flödesdata och/eller tidsserieanalys finns beskrivna i detalj i avsnitt 7.

6 Metod för länkning av tidsserier

Med hjälp av analyserna i föregående avsnitt kan vi skatta storleken för de tidsseriebrott som förändringarna i AKU har givit upphov till. För att använda dessa brottskattningar till att konstruera länkade tidsserier använder vi den metodik som beskrivs i kommande avsnitt.

6.1 Länkning med makroansats

Låt X_t vara en av AKU:s tidsserier. Vi omvandlar först summan av brottskattningen för bytet av hjälpinformation $\hat{\Delta}^h$ och den eventuella extra blanketteffekten $\hat{\Delta}^e$ för denna tidsserie (som uttrycks i antalet tusental personer) till en faktor av befolkningen. Mer specifikt, så beräknar vi först medelvärdet $\bar{Y}$ under 2021 av RTB-befolkningen Y_t^g för samma åldersgrupp som X_t , dvs

$$\bar{Y} := \frac{1}{T_1 - T_0^r} \sum_{t \in [T_0^r, T_1]} Y_t^g.$$

Brottskattningen uttryckt som en faktor av $\bar{Y}$ blir då

$$\hat{\Delta} := \frac{\hat{\Delta}^h + \hat{\Delta}^e}{\bar{Y}}.$$

Därefter härleder vi länkade tidsserier för perioden januari 2005–december 2020 genom att justera tidsserien $X_t^{g,g}$ enligt

$$X_t^{link} := X_t^{g,g} + \hat{\Delta}_t^d + \hat{\Delta}_t^p + Y_t^g \hat{\Delta}.$$

Eftersom alla brottskattningar är summakonsistenta och eftersom de, för varje åldersgrupp, viktas med samma faktor Y_t^g , så kommer de resulterande länkade tidsserierna att vara summakonsistenta. Länkade tidsserier för åldersgruppen 15–74 år erhålls genom att summera de länkade tidsserierna för alla undergrupper i ålder. De korrekationer som införts i de länkade tidsserierna på grund av bytet av hjälpinformation eller extra blanketteffekter är proportionella mot storleken på befolkningen i motsvarande åldersgrupp. Eftersom befolkningen har ökat över tid, så innebär detta att korrektionerna generellt är mindre i början av länkingsperioden 2005–2020 än i slutet av densamma.

Värdena på de länkade tidsserierna i december 2020 är framtagna på ett något annorlunda sätt än enligt formlerna ovan. Praxis i AKU har varit att uppdatera RAMS-versionen i hjälpinformationen i december månad och säsongsmönstret i AKU är därför anpassat till detta förfarande. I december 2020 uppdaterades dock inte hjälpinformationen från RAMS

2018 till RAMS 2019 i de publicerade tidsserierna⁴. För att undvika att orsaka en avvikelse i säsongsmönstret, så utgår värdena på de länkade tidsserierna i december 2020 från skattningar med RAMS 2019 som hjälpinformation i stället för de publicerade skattningarna som har RAMS 2018 som hjälpinformation.

6.2 Länkning med mikroansats

Som nämnades i avsnitt 3.4 så används en regressionsestimator för att ta fram skattningar i AKU. För att beskriva metodiken lite mer utförligt, så låter vi $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^T$ beteckna en hjälpvektor som antas vara känd från register för hela populationen U . Även de så kallade hjälptotalerna $\mathbf{X} = \sum_{i \in U} \mathbf{x}_i$ är därmed kända. Vi tilldelar varje svarande urvalsperson i AKU en uppräkningsvikt w_i som uppfyller kalibreringsekvationen

$$\sum_{i \in r} w_i \mathbf{x}_i = \mathbf{X},$$

där r betecknar mängden av svarande urvalspersoner. De kalibrerade uppräkningsvikterna kan sedan användas för att skatta totalen av varje given undersökningsvariabel y_i i AKU med hjälp av $\hat{y} = \sum_{i \in r} w_i y_i$.

Tanken med länkning på mikronivå är att använda de länkade tidsserier som tagits fram på makronivå enligt beskrivningen i avsnitt 6.1 som hjälptotaler i kalibreringsekvationen. Vi får då, för varje månad under länkingsperioden, en uppsättning kalibrerade uppräkningsvikter som är konsistenta med de länkade tidsserierna på makronivå. Med denna nya uppsättning kalibrerade uppräkningsvikter, som vi hädanefter refererar till som mikrolänkade vikter, kan vi dessutom ta fram länkade skattningar för samtliga undersökningsvariabler i AKU.

Det finns ett par viktiga metodologiska frågor att ta ställning till vid härledning av mikrolänkade vikter. För det första är det en avvägning att bestämma den optimala mängden av länkade tidsserier att använda som hjälptotaler, eller styrvillkor, vid justeringen av de kalibrerade uppräkningsvikterna. Om hela mängden länkade tidsserier används som styrvillkor finns det en risk att de mikrolänkade vikterna överanpassas i kalibreringen med alltför stor variation i vikterna, och kanske till och med negativa vikter, som följd. Å andra sidan, om alltför få av de länkade tidsserierna används som styrvillkor finns det en risk att de länkade tidsserier som inte används som styrvillkor inte kan återskapas med tillräckligt god precision av de mikrolänkade vikterna. I kapitel 8 nedan redovisar vi det vägval som vi har gjort i denna fråga i den nuvarande länkningen av AKU.

⁴ Att byte av RAMS-version inte genomfördes för AKU i december 2020 såsom brukligt berodde på att framställningen av RAMS ändrats mellan 2018 och 2019.

Ett annat problem som kan uppkomma vid länkning med mikroansats är att det i svarsdata en given månad inte finns någon urvalsperson som uppfyller villkoren för en tidsserie som i länkningen har fått ett värde som är strikt större än noll. I sådana fall kan kalibreringsekvationen omöjligen vara uppfylld och det går inte att härleda mikrolänkade vikter. För den nuvarande AKU-länkningen är det framför allt för vissa indelningar av latent arbetssökande som detta problem uppkommer (eftersom denna grupp har blivit avsevärt mycket större i den nya blanketten).

Vi har löst det senare problemet genom att ta fram en imputeringsmodell för latent arbetssökande (se bilaga C för detaljer). Med hjälp av modellen justeras svarsdata så att personer som med hög sannolikhet hade klassificerats som latent arbetssökande i den nya blanketten men som inte klassificerades som detta i den gamla blanketten får en ny klassificering i svarsdata.

På liknande sätt har vi skapat en imputeringsmodell för definitionsförändringen. Med hjälp av modellen ändrar vi svarsdata för vissa urvalspersoner från att vara frånvarande från arbetet till att vara utanför arbetskraften (se bilaga A för detaljer). En fördel med att använda imputeringsmodeller för förändringar som inträffat vid bytet av förfarande i AKU är att det minskar skillnaden mellan de mikrolänkade vikterna och de ursprungliga kalibrerade uppräkningsvikterna.

7 Skattningar av tidsseriebrott

I detta avsnitt redovisar vi skattningar av tidsseriebrott i AKU, men inleder med några generella ställningstaganden som har gjorts i analysen. I AKU:s säsongrensningssystem delas det totala urvalet bestående av personer i åldrarna 15–74 år (åldersgruppen 75–89 år ingår inte i säsongrensningssystemet) upp i följande åldersgrupper: 15 år, 16–17 år, 18–19 år, 20–24 år, 25–34 år, 35–44 år, 45–54 år, 55–59 år, 60–64 år samt 65–74 år. Vi har tagit fram separata brottskattningar för alla dessa åldersgrupper och sedan låtit brottskattningen för 15–74 år vara summan av brottskattningarna för alla ingående åldersgrupper. På motsvarande sätt används könsuppdelningen män respektive kvinnor i AKU. Vi har tagit fram separata brottskattningar för respektive kön och får den totala brottskattningen som summan av brottskattningarna för båda könen. I AKU används även uppdelningen inrikes född/utrikes född. För en delmängd av variablerna har vi tagit fram separata brottskattningar för respektive födelselandskategori och i dessa fall fås den totala brottskattningen som summan av brottskattningarna för de båda födelselandskategorierna.

Eftersom urvalet för åldersgruppen 15 år är mycket begränsat, särskilt i dubbelmätningen, så har alla brottskattningar för denna åldersgrupp satts till noll förutom för latent arbetssökande (se avsnitt 7.5). För de båda åldersgrupperna 16–17 år respektive 18–19 år, så har vi inte kunnat göra separata brottskattningar eftersom vi inte har en finare åldersindelning i hjälpinformationerna än 16–19 år. Detta implicerar att direkta brottskattningar för åldersgrupperna 16–17 år respektive 18–19 år är mycket volatila. Vi har därför valt att göra brottskattningar för åldersgruppen 16–19 år och att sedan fördela dessa brottskattningar lika över åldersgrupperna 16–17 år respektive 18–19 år. För ytterligare några tidsserier har brottskattningarna satts till noll även för några andra åldersgrupper, exempelvis arbetslösa i åldersgruppen 65–74 år samt fast anställda och företagare i åldersgruppen 16–19 år.

7.1 Ny hjälpinformation

Vi redovisar först skattningar av tidsseriebrott till följd av byte av hjälpinformation. Tabell 7.1 visar brottskattningarna för sysselsatta och tabell 7.2 brottskattningarna för arbetslösa uppdelat på ålder och kön samt på födelseland och kön.

Tabell 7.1. Skattningar av tidsseriebrott till följd av byte av hjälpinformation (ny minus gammal hjälpinformation) för antalet sysselsatta, uppdelat på ålder och kön samt på födelseland och kön, mätt i tusentals personer. Signifikanta brott på 5 %-nivån indikeras med en asterisk.

Sysselsatta	Män	Kvinnor	Totalt
16–19 år	-0.3 (±0.8)	0.0 (±0.9)	-0.3 (±1.7)
20–24 år	5.1 (±1.3)*	4.0 (±1.1)*	9.1 (±2.3)*
25–34 år	-4.9 (±0.9)*	-4.6 (±1.0)*	-9.6 (±1.7)*
35–44 år	-1.3 (±0.6)*	-1.7 (±0.9)*	-3.0 (±1.3)*
45–54 år	-1.9 (±0.7)*	-1.8 (±0.9)*	-3.7 (±1.4)*
55–59 år	-1.7 (±0.5)*	-0.9 (±0.5)*	-2.5 (±0.9)*
60–64 år	-5.9 (±0.8)*	-6.3 (±0.9)*	-12.1 (±1.5)*
65–74 år	-4.7 (±0.8)*	-3.8 (±0.7)*	-8.5 (±1.4)*
15–74 år, inrikes födda	-13.5 (±2.8)*	-11.8 (±3.3)*	-25.3 (±5.6)*
15–74 år, inrikes födda	-2.0 (±1.5)*	-3.4 (±1.4)*	-5.4 (±2.6)*
15–74 år, totalt	-15.5 (±3.8)*	-15.2 (±4.3)*	-30.6 (±7.8)*

Tabell 7.2. Skattningar av tidsseriebrott till följd av byte av hjälpinformation (ny minus gammal hjälpinformation) för antalet arbetslösa, uppdelat på ålder och kön samt på födelseland och kön, mätt i tusentals personer. Signifikanta brott på 5 %-nivån indikeras med en asterisk.

Arbetslösa	Män	Kvinnor	Totalt
16–19 år	-0.4 (±0.2)*	-0.1 (±0.2)	-0.5 (±0.3)*
20–24 år	-2.4 (±0.6)*	-1.3 (±0.3)*	-3.6 (±0.8)*
25–34 år	0.4 (±0.5)	0.1 (±0.5)	0.5 (±0.9)
35–44 år	0.1 (±0.5)	-0.8 (±0.5)*	-0.7 (±0.8)
45–54 år	-0.1 (±0.5)	-0.3 (±0.5)	-0.4 (±0.8)
55–59 år	0.2 (±0.3)	-0.3 (±0.3)*	-0.1 (±0.5)
60–64 år	-0.3 (±0.3)*	-0.2 (±0.3)	-0.6 (±0.4)*
65–74 år	0.0 (±0.1)	0.0 (±0.1)	0.1 (±0.1)
15–74 år, inrikes födda	-1.5 (±1.1)*	-1.2 (±0.8)*	-2.7 (±1.7)*
15–74 år, inrikes födda	-1.0 (±1.0)*	-1.7 (±0.9)*	-2.6 (±1.7)*
15–74 år, totalt	-2.5 (±1.8)*	-2.8 (±1.5)*	-5.3 (±3.1)*

För sysselsatta är alla brottskattningar utom för åldersgruppen 16–19 år signifikanta. Vi sätter därför brottskattningen för åldersgruppen 16–19 år till noll i länken. För arbetslösa är endast brottskattningarna för 16–19 år, 20–24 år och 60–64 år signifikanta. Eftersom den signifikanta skattningen för åldersgruppen 60–64 år är mycket liten i absoluta tal, så

sätter vi både den och samtliga icke-signifikanta brottskattningar för arbetslösa till noll i länken.

Eftersom brottskattningarna för sysselsatta och arbetslösa har satts till noll i vissa åldersgrupper och den totala brottskattningen ges av summan av brottskattningarna för de ingående åldersgrupperna, så finns det mindre avvikelser mellan de brottskattningar för 15–74 år som redovisas i tabell 7.1–7.2 och de slutliga brottskattningarna för 15–74 år som redovisas i tabell 7.13.

Tabell 7.3 visar, uppdelat på kön, brottskattningarna för åldersgruppen 15–74 år för samtliga tidsserier för antal personer utom för sysselsatta, arbetslösa och befolkningen. Notera här att befolkningen inte påverkas av bytet av hjälpinformation. Tabell 7.4 visar, uppdelat på födelseland, brottskattningarna för åldersgruppen 15–74 år för den delmängd av tidsserierna i tabell 7.3 som har studerats efter födelseland.

På totalnivå är samtliga brottskattningar för 15–74 år signifikanta bortsett från brottskattningen för fast anställda. Latent arbetssökande ligger precis på gränsen till signifikans. Brottskattningarna är relativt lika för båda könen bortsett från för fast anställda samt för företagare och medhjälpande hushållsmedlemmar där brottskattningarna har olika tecken för män och kvinnor. För fast anställda har brottskattningarna även olika tecken för inrikes födda och utrikes födda.

Huvudregeln är att brottskattningar för bytet av hjälpinformation används i länken även om de inte är signifikanta. Undantag har bland annat gjorts för samtliga undergrupper till sysselsatta för åldersgruppen 16–19 år. Dessa brottskattningar har satts till noll i länken av konsistensskäl eftersom brottskattningen för sysselsatta i samma åldersgrupp har satts till noll. Samma sak gäller för arbetslösa, heltidsstuderande för de åldersgrupper där brottskattningen för arbetslösa har satts till noll.

Samtliga brottskattningar som understiger 300 personer har dock satts till noll i länken. Detta val har gjorts för att undvika att justera tidsserier som är mycket små i absoluta tal. Även mycket små negativa brottskattningar riskerar att leda till negativa värden för de länkade tidsserierna vissa månader. Brottskattningar för aggregerade tidsserier efter kön, ålder eller födelseland skapas genom att summera brottskattningarna för motsvarande undergrupper. Eftersom brottskattningarna för vissa undergrupper har satts till noll, så förändrar denna metodik även brottskattningarna för aggregerade tidsserier. Justeringarna i absoluta tal är emellertid små och uppgår som högst till ett tusental personer (exempelvis 800 personer för sysselsatta, båda könen 15–74 år).

Tabell 7.3. Skattningar av tidsseriebrott till följd av byte av hjälpinformation (ny minus gammal hjälpinformation) för samtliga undersökta variabler för personer utom för sysselsatta, arbetslösa och befolkningen, för åldersgruppen 15–74 år uppdelat på kön, mätt i tusentals personer. Signifikanta brott på 5 %-nivån indikeras med en asterisk.

Tidsserie	Män	Kvinnor	Totalt
Arbetslösa, heltidsstuderande	-0.8 (±0.6)*	-0.7 (±0.5)*	-1.5 (±0.9)*
I arbetskraften	-18.0 (±3.1)*	-18.1 (±3.8)*	-36.0 (±6.6)*
Ej i arbetskraften	18.0 (±3.1)*	18.1 (±3.8)*	36.0 (±6.6)*
Sysselsatta, i arbete	-19.4 (±3.5)*	-29.3 (±4.0)*	-48.7 (±7.1)*
Sysselsatta, frånvarande	3.9 (±1.1)*	14.0 (±1.4)*	18.0 (±2.1)*
Fast anställda	-6.7 (±3.2)*	2.5 (±3.8)	-4.3 (±6.5)
Tidsbegränsat anställda	-12.6 (±1.5)*	-16.4 (±1.7)*	-28.9 (±2.8)*
Anställda	-19.3 (±3.7)*	-13.9 (±4.4)*	-33.2 (±7.7)*
Företagare	3.8 (±1.6)*	-1.3 (±0.9)*	2.5 (±2.1)*
Privat anställda	-30.1 (±4.4)*	-26.3 (±4.2)*	-56.5 (±8.3)*
Statligt anställda	5.0 (±1.2)*	2.2 (±0.8)*	7.3 (±1.9)*
Kommunalt anställda	4.0 (±2.1)*	9.3 (±4.3)*	13.3 (±6.2)*
Primärkommunalt anställda	2.3 (±1.6)*	5.7 (±3.2)*	8.0 (±4.6)*
Landstingskommunalt anställda	1.7 (±0.6)*	3.6 (±1.3)*	5.3 (±1.7)*
Undersysselsatta	-3.2 (±0.8)*	-5.0 (±0.8)*	-8.1 (±1.2)*
Latent arbetssökande	0.7 (±0.5)*	0.2 (±0.5)	0.9 (±0.8)*
Ej i arbetskraften, heltidsstud.	2.4 (±1.4)*	1.3 (±1.4)	3.7 (±2.7)*
Ej i arbetskraften, pensionärer	9.5 (±1.1)*	8.6 (±1.1)*	18.1 (±1.9)*
Ej i arbetskraften, sjuka	4.7 (±1.1)*	4.3 (±1.5)*	8.9 (±2.5)*
Ej i arbetskraften, övriga	1.3 (±0.8)*	3.9 (±1.0)*	5.2 (±1.5)*

Tabell 7.4 Skattningar av tidsseriebrott till följd av ny hjälpinformation (ny minus gammal hjälpinformation) för de variabler för personer som studerats uppdelat på födelseland för åldersgruppen 15–74 år uppdelat på födelseland, mätt i tusentals personer. Signifikanta brott på 5 %-nivån indikeras med en asterisk.

Tidsserie	Inrikes födda	Utrikes födda	Totalt
I arbetskraften	-27.5 (±5.1)*	-8.5 (±1.8)*	-36.0 (±6.6)*
Ej i arbetskraften	27.5 (±5.1)*	8.5 (±1.8)*	36.0 (±6.6)*
Fast anställda	-8.6 (±4.8)*	4.4 (±2.1)*	-4.3 (±6.5)
Tidsbegränsat anställda	-17.3 (±2.2)*	-11.7 (±1.1)*	-28.9 (±2.8)*
Anställda	-25.9 (±5.6)*	-7.3 (±2.5)*	-33.2 (±7.7)*
Företagare	1.1 (±1.8)	1.4 (±0.6)*	2.5 (±2.1)*
Undersysselsatta	-4.4 (±0.9)*	-3.7 (±0.6)*	-8.1 (±1.2)*
Latent arbetssökande	0.8 (±0.6)*	0.0 (±0.4)	0.9 (±0.8)*

Brottskattningarna för latent arbetssökande har satts till noll i länken trots att brottskattningen för gruppen är svagt signifikant för 15–74 år. Detta val motiveras med att brottskattningarna för bytet av hjälpinformation för latent arbetssökande är mycket små i absoluta tal och de är dessutom obetydliga i förhållande till de avsevärt större brottskattningarna på grund av den nya ramlagen (se avsnitt 7.2).

7.2 Ny ramlag

Vi redovisar här brottskattningarna från dubbelmätningen. Tabell 7.5 visar, uppdelat på kön, brottskattningarna för åldersgruppen 15–74 år för samtliga tidsserier för antal personer. Tabell 7.6 visar, uppdelat på födelseland, brottskattningarna för åldersgruppen 15–74 år för de tidsserier för antal personer som har studerats efter födelseland.

Tabell 7.5. Skattningar av tidsseriebrott till följd av ny ramlag (ny minus gammal ramlag) för samtliga undersökta variabler för personer för åldersgruppen 15–74 år uppdelat på kön, mätt i tusentals personer. Signifikantabrott på 5 %-nivån indikeras med en asterisk.

Tidsserie	Män	Kvinnor	Totalt
Sysselsatta	-14.1 (±48.0)	9.6 (±48.0)	-4.5 (±55.7)
Arbetslösa	25.8 (±24.4)*	-8.5 (±25.8)	17.3 (±31.1)
Arbetslösa, heltidsstuderande	6.0 (±15.2)	-1.5 (±17.5)	4.5 (±22.8)
I arbetskraften	11.7 (±45.4)	1.1 (±46.8)	12.8 (±58.7)
Befolkningen	-26.3 (±5.8)*	-15.0 (±4.5)*	-41.3 (±7.3)*
Ej i arbetskraften	-38.0 (±45.8)	-16.0 (±47.0)	-54.1 (±59.2)
Sysselsatta, i arbete	2.2 (±45.6)	14.6 (±44.6)	16.8 (±56.3)
Sysselsatta, frånvarande	-16.3 (±28.3)	-5.0 (±29.5)	-21.3 (±40.6)
Fast anställda	-43.6 (±57.2)	-6.0 (±54.9)	-49.6 (±71.0)
Tidsbegränsat anställda	6.2 (±34.3)	27.9 (±36.7)	34.1 (±49.7)
Anställda	-37.4 (±53.5)	21.9 (±51.2)	-15.5 (±63.2)
Företagare	23.4 (±41.4)	-12.3 (±26.9)	11.0 (±49.4)
Privat anställda	-40.0 (±59.4)	52.6 (±56.8)	12.5 (±78.0)
Statligt anställda	-0.3 (±25.0)	-5.5 (±27.6)	-5.7 (±37.1)
Kommunalt anställda	5.3 (±36.2)	-24.1 (±51.7)	-18.8 (±62.3)
Primärkommunalt anställda	14.1 (±31.9)	-1.0 (±46.3)	13.1 (±55.8)
Landstingskommunalt anställda	-8.8 (±18.8)	-23.1 (±32.1)	-31.9 (±37.1)
Undersysselsatta	16.9 (±25.8)	22.0 (±25.8)	39.0 (±36.3)*
Latent arbetssökande	69.5 (±20.6)*	59.4 (±22.0)*	128.9 (±30.0)*
Ej i arbetskraften, heltidsstud.	-2.3 (±28.6)	-7.6 (±31.9)	-10.0 (±41.4)
Ej i arbetskraften, pensionärer	-17.3 (±28.8)	6.8 (±24.9)	-10.5 (±37.4)
Ej i arbetskraften, sjuka	7.2 (±22.0)	5.4 (±26.1)	12.6 (±33.1)
Ej i arbetskraften, övriga	-25.6 (±25.2)*	-20.6 (±29.8)	-46.3 (±38.5)*

Tabell 7.6. Skattningar av tidsseriebrott till följd av ny ramlag (ny minus gammal ramlag) för de variabler för personer som studerats uppdelat på födelseland för åldersgruppen 15–74 år uppdelat på födelseland, mätt i tusentals personer. Signifikanta brott på 5 %-nivån indikeras med en asterisk.

Tidsserie	Inrikes födda	Utrikes födda	Totalt
Sysselsatta	-5.2 (±58.1)	0.7 (±45.7)	-4.5 (±55.7)
Arbetslösa	32.6 (±23.6)*	-15.3 (±25.8)	17.3 (±31.1)
I arbetskraften	27.3 (±56.9)	-14.6 (±44.5)	12.8 (±58.7)
Befolkningen	-33.1 (±6.6)*	-8.2 (±4.1)	-41.3 (±7.3)*
Ej i arbetskraften	-60.4 (±57.3)*	6.4 (±44.7)	-54.1 (±59.2)
Fast anställda	-40.0 (±68.1)	-9.6 (±49.0)	-49.6 (±71.0)
Tidsbegränsat anställda	11.7 (±40.4)	22.3 (±30.3)	34.1 (±49.7)
Anställda	-28.3 (±62.9)	12.8 (±47.9)	-15.5 (±63.2)
Företagare	23.1 (±43.1)	-12.0 (±24.7)	11.0 (±49.4)
Undersysselsatta	9.6 (±29.0)	29.4 (±22.1)*	39.0 (±36.3)*
Latent arbetssökande	105.4 (±24.2)*	23.5 (±18.4)*	128.9 (±30.0)*

I stort sett samtliga redovisade brottskattningar från dubbelmätningen är icke-signifikanta. På totalnivå är det bara brottskattningarna för befolkningen, undersysselsatta, latent arbetssökande och ej i arbetskraften övriga som är signifikanta. Vidare är brottskattningarna för män och kvinnor generellt väldigt olika, både till storlek och tecken. Samma brist på mönster kan ses även för brottskattningar för inrikes födda och utrikes födda samt för brottskattningar för olika åldersgrupper (redovisas ej här). Att det inte går att urskilja ett tydligt mönster efter ålder, kön eller födelseland indikerar att brottskattningarna från dubbelmätningen i hög utsträckning beror på slumpmässiga effekter, såsom urvalsfel.

I detta sammanhang bör det noteras att dubbelmätningens skattning av populationsförändringen (-41,3 tusen) inte sammanfaller med den registerbaserade skattningen av populationsförändringen (-35,9 tusen). Denna avvikelse kan bero på att dubbelmätningens skattning är behäftad med ett visst urvalsfel och att det finns viss osäkerhet i registeruppgifternas täckningsgrad och aktualitet. Den främsta orsaken till avvikelserna är dock troligen att dubbelmätningens skattning påverkas av en felaktig hantering under 2021 av de urvalspersoner som baserat på sina svar i AKU klassats som ej boende i privata hushåll. Dessa personer räknades bort från befolkningen under innevarande AKU-omgång, men även under alla framtida AKU-omgångar. Eftersom det finns en viss omsättning mellan att bo i privat hushåll och inte, så har hanteringen inneburit att skattningen av hur många personer som inte bor i privat hushåll har ökat för varje kvartal. En följd av detta är att dubbelmätningen överskattar antalet personer som påverkas av

populationsförändringen. Denna effekt förklarar förmodligen också varför gruppen ej i arbetskraften övriga, till vilken många av de personer som påverkas av populationsförändringen hör, var signifikant i dubbelmätningen.

Inga brottskattningar från dubbelmätningen har använts direkt i länkningsperioden, men brottskattningarna för latent arbetssökande (se avsnitt 7.6) baseras delvis på dubbelmätningen eftersom effekterna för denna variabel är så starka i dubbelmätningen. Även flöden och tidsserieanalys har dock tagits i beaktande i detta fall.

7.3 Definitionsförändring

Vi redovisar här brottskattningar för definitionsförändringen av sysselsatta. Tabell 7.7 visar, uppdelat på kön, brottskattningarna för definitionsförändringen. Brottskattningarna baseras på medelvärden över 2020 av skattningar baserade på imputeringsmodellen för definitionsförändringen (se bilaga A). Notera att medelvärden över tidigare år i länkningsperioden generellt ger avsevärt mindre brottskattningar (se tabell A.6). Eftersom vi saknar explicita mått på osäkerheten för dessa brottskattningar, så kan vi inte redovisa vilka brottskattningar som är signifikanta eller ej i tabell 7.7.

Generellt används alla brottskattningar för definitionsförändringen i länkningsperioden. Ett undantag utgörs av åldersgruppen 15 år för vilken brottskattningarna är mycket små i absoluta tal. Även för variablerna sysselsatta i arbete, befolkningen, latent arbetssökande och ej i arbetskraften sjuka används brottskattningarna av olika skäl inte i länkningsperioden. För sysselsatta i arbete och befolkningen är orsaken helt enkelt att dessa variabler inte påverkas av definitionsförändringen. För latent arbetssökande är brottskattningarna för definitionsförändringen satta till noll eftersom vi har konstruerat en separat imputeringsmodell för latent arbetssökande. För ej i arbetskraften sjuka är det ett försumbart antal av de som påverkas av definitionsförändringen som hamnar i denna grupp. Gruppen är därför undantagen från imputeringsmodellen, vilket gör att brottskattningarna blir noll.

7.4 Populationsförändring

Vi redovisar här brottskattningar för populationsförändringen. Tabell 7.8 visar, uppdelat på ålder och kön respektive ålder och födelseland, brottskattningar för det totala antalet personer som påverkas av förändringen av målpopulation. Brottskattningarna baseras på medelvärden över 2020 av skattningar baserade på den modell som tagits fram utgående från registeruppgifter under länkningsperioden (se bilaga B). Eftersom vi saknar explicita mått på osäkerheten för dessa brottskattningar, så kan vi inte redovisa vilka brottskattningar som är signifikanta eller ej i tabell 7.8. Brottskattningarna baseras dock på

registeruppgifter, vilket gör att osäkerheten i skattningarna generellt torde vara liten.

Tabell 7.7. Skattningar av tidsseriebrott till följd av definitionsförändring (ny minus gammal definition) för alla tidsserier för personer utom befolkningen, sysselsatta i arbete och latent arbetssökande för åldersgruppen 15–74 år uppdelat på kön respektive födelseland, mätt i tusentals personer. Fält i tabellen som är markerade med streck innebär att variabeln i fråga inte har undersökts uppdelat på födelseland.

Tidsserie	Inr. födda	Utr. födda	Män	Kvinnor	Totalt
Sysselsatta	-22.3	-5.7	-12.0	-16.3	-28.3
Arbetslösa	3.2	1.1	1.7	2.5	4.2
Arbetslösa, stud.	–	–	0.7	1.4	2.1
I arbetskraften	-19.4	-4.6	-10.2	-13.8	-24.0
Ej i arbetskraften	19.4	4.6	10.2	13.8	24.0
Sysselsatta, frånv.	–	–	-12.0	-16.3	-28.3
Fast anställda	-17.2	-4.3	-8.3	-13.2	-21.5
Tidsbegr. anst.	-2.6	-0.5	-1.1	-2.0	-3.1
Anställda	-19.9	-4.7	-9.4	-15.2	-24.6
Företagare	-2.7	-1.0	-2.6	-1.1	-3.7
Privat anställda	–	–	-7.3	-9.1	-16.5
Statligt anställda	–	–	-0.3	-0.4	-0.7
Kommunalt anst.	–	–	-1.2	-5.6	-6.8
Primärkom. anst.	–	–	-1.1	-4.8	-5.9
Landst.kom. anst.	–	–	-0.1	-0.8	-0.9
Undersysselsatta	-1.7	-0.8	-1.2	-1.2	-2.5
Ej i arbetskr. stud.	–	–	4.1	8.6	12.6
Ej i arbetskr. pens.	–	–	1.0	0.9	2.0
Ej i arbetskr. sjuka	–	–	0.0	0.0	0.0
Ej i arbetskr. övriga	–	–	5.1	4.3	9.5

Med hjälp av svarsdata i AKU för personer som enligt registeruppgifter påverkas av förändringen i målpopulation kan vi även skatta påverkan av denna förändring på övriga variabler i AKU (se bilaga B för detaljer). Tabell 7.9 visar resultatet av sådana skattningar för 2020. Skattningarna som redovisas i tabell 7.9 varierar avsevärt under länkningsperioden främst på grund av att de personer som genomför värnplikt eller militär grundutbildning har varierat stort i antal, men även på grund av att dessa personer har klassificerats olika under perioden. Eftersom mycket få personer av de som enligt registeruppgifter inte bor i privata hushåll har klassats som undersysselsatta i AKU, så har denna variabel försumrats i modellen för populationsförändringen och variabeln saknas därför i tabell 7.9.

Tabell 7.8. Skattningar av tidsseriebrott till följd av byte av målpopulation, uppdelat på ålder och kön respektive på ålder och födelseland, mätt i tusentals personer.

Population	Inr. födda	Utr. födda	Män	Kvinnor	Totalt
16–19 år	-2,8	-0,4	-2,6	-0,6	-3,2
20–24 år	-2,6	-0,8	-2,6	-0,8	-3,3
25–34 år	-4,5	-1,7	-4,3	-1,9	-6,2
35–44 år	-3,0	-1,2	-2,8	-1,5	-4,2
45–54 år	-3,3	-0,9	-2,5	-1,7	-4,2
55–59 år	-2,0	-0,4	-1,4	-1,0	-2,4
60–64 år	-2,1	-0,5	-1,5	-1,1	-2,6
65–74 år	-8,1	-1,6	-5,0	-4,7	-9,7
15–74 år	-28,4	-7,5	-22,6	-13,3	-35,9

I avsnitt 7.2 redovisades den troliga orsaken till att dubbelmätningen överskattar storleken på populationsförändringen. En regressionsanalys av svarsdata visar att antalet personer som flyttar mellan privata hushåll och icke-privata hushåll är ca 11,5 tusen personer per kvartal. Detta motsvarar en överskattning av populationsförändringen i data från dubbelmätningen med ca 14 tusen personer. Att skillnaden mellan den registerbaserade skattningen av populationsförändringen och skattningen från dubbelmätningen bara uppgår till ca 5,5 tusen personer beror på att de båda skattningarna baseras på olika data och mätmetoder. Med registerbaserade metoder fångar vi förmodligen upp fler personer som inte bor i privata hushåll än vad vi fångar upp med AKU, eftersom bortfallet förmodligen är större för personer som inte bor i privata hushåll än för befolkningen i stort.

7.5 Extra blanketteffekter

Vi redovisar här de extra blanketteffekter som införts för vissa tidsserier. Vi börjar med latent arbetssökande och undersysselsatta för att därefter redovisa övriga variabler.

Latent arbetssökande är den variabel som uppvisar de i särklass största relativa förändringarna i och med införandet av den nya ramlagen. De införda korrektionerna baseras främst på tidsserieanalys (se bilaga C för detaljer), men även på dubbelmätningen och flödesdata. Tabell 7.10 visar de korrektioner som har använts i länkningen. För samtliga åldersgrupper har fyra femtedelar av den totala korrektionen lagts i gruppen inrikes födda och resterande del i gruppen utrikes födda. Det förefaller inte finnas några väsentliga könsskillnader i effekten för latent arbetssökande och samma korrektioner har därför införts för båda könen.

Tabell 7.9. Skattningar av tidsseriebrott till följd av byte av målpopulation för samtliga undersökta variabler för personer utom för befolkningen och undersysselsatta för åldersgruppen 15–74 år uppdelat på kön respektive födelseland, mätt i tusentals personer. Fält i tabellen som är markerade med streck innebär att variabeln i fråga inte har undersökts uppdelat på födelseland.

Tidsserie	Inr. födda	Utr. födda	Män	Kvinnor	Totalt
Sysselsatta	-4.9	-1.0	-3.9	-2.1	-6.0
Arbetslösa	-0.1	-0.1	-0.1	-0.1	-0.2
Arbetslösa, stud.	–	–	-0.1	-0.1	-0.1
I arbetskraften	-5.1	-1.1	-4.0	-2.2	-6.2
Ej i arbetskraften	-23.3	-6.3	-18.6	-11.1	-29.7
Sysselsatta, i arb.	–	–	-3.1	-1.7	-4.8
Sysselsatta, frånv.	–	–	-0.8	-0.4	-1.2
Fast anställda	-2.9	-0.7	-2.1	-1.5	-3.7
Tidsbegr. anst.	-2.0	-0.3	-1.7	-0.6	-2.3
Anställda	-4.9	-1.0	-3.9	-2.1	-6.0
Företagare	0.0	0.0	0.0	0.0	0.0
Privat anställda	–	–	-1.7	-1.2	-3.0
Statligt anställda	–	–	-1.4	-0.3	-1.6
Kommunalt anst.	–	–	-0.8	-0.6	-1.4
Primärkom. anst.	–	–	-0.8	-0.6	-1.4
Landst.kom. anst.	–	–	0.0	0.0	0.0
Latent arbetssök.	0.0	0.0	0.0	0.0	-0.1
Ej i arbetskr. stud.	–	–	-0.5	-0.4	-0.9
Ej i arbetskr. pens.	–	–	-3.5	-3.4	-6.9
Ej i arbetskr. sjuka	–	–	-7.1	-5.7	-12.8
Ej i arbetskr. övriga	–	–	-7.5	-1.7	-9.2

För utrikes födda undersysselsatta är skattningen av den totala definitionsförändringen på -0,4 tusen personer inte kompatibel med konfidensintervallet från dubbelmätningen som är 29,4±22,1, mätt i tusentals personer. En extra blanketteffekt på 9 tusen personer har därför införts för denna grupp. Effekten har fördelats jämnt mellan könen och de tre åldersgrupperna 25–34 år, 35–44 år och 45–54 år.

Utöver de extra blanketteffekterna för latent arbetssökande och undersysselsatta, så har extra blanketteffekter tagits med i länkningen för vissa variabler för åldersgrupperna 25–34 år, 60–64 år och 65–74 år. Vi redovisar dessa extra blanketteffekter nedan.

Tabell 7.10. Extra blanketteffekter för latent arbetssökande till följd av ny ramlag (ny minus gammal ramlag), uppdelat på ålder och födelseland, mätt i tusentals personer.

Ålder	Inrikes födda	Utrikes födda	Totalt
15 år	7.2	1.8	9.0
16–19 år	19.2	4.8	24.0
20–24 år	5.6	1.4	7.0
25–34 år	6.4	1.6	8.0
35–44 år	4.0	1.0	5.0
45–54 år	6.4	1.6	8.0
55–59 år	3.2	0.8	4.0
60–64 år	5.6	1.4	7.0
65–74 år	23.2	5.8	29.0
15–74 år	80.8	20.2	101.0

För åldersgruppen 25–34 år är skattningen av definitionsförändringen för antalet arbetslösa på 1,2 tusen personer inte kompatibel med konfidensintervallet från dubbelmätningen som är $22,4 \pm 18,0$, mätt i tusentals personer. Avvikelsen förefaller enligt dubbelmätningen vara jämnt fördelad mellan könen. Även för undergruppen arbetslösa heltidsstuderande är dubbelmätningen inte konsistent med skattningen av definitionsförändringen i denna åldersgrupp. En motsvarande avvikelse kan även ses för inrikes födda arbetslösa 15–74 år, där skattningen av definitionsförändringen på 3,3 tusen personer inte är kompatibel med konfidensintervallet från dubbelmätningen som är $32,6 \pm 23,6$, mätt i tusentals personer.

Ingen motsvarande avvikelse för antalet sysselsatta kan dock ses i åldersgruppen 25–34 år, varför tillskottet av arbetslösa torde komma från personer som tidigare bedömts vara utanför arbetskraften. Baserat på flöden och dubbelmätningen bedöms ökningen för arbetslösa härstamma från gruppen ej i arbetskraften övriga. En möjlig förklaring till förändringen kan vara den nya formuleringen av frågorna angående arbetssökande. Tabell 7.11 sammanfattar de extra blanketteffekter som införts för åldersgruppen 25–34 år i länken. Merparten av effekterna har lagts i gruppen inrikes födda till följd av avvikelsen för inrikes födda arbetslösa 15–74 år.

För åldersgruppen 60–64 år visar tidsserieanalys av länkade tidsserier som saknar en extra blanketteffekt att det finns ett signifikant nivåskifte i gruppen sysselsatta. En ökning med 2 tusen inrikes födda sysselsatta, jämnt fördelat mellan könen har därför lagts till. Ökningen påverkar även grupperna i arbetskraften, anställda, fast anställda, sysselsatta i arbete och anställda i privat sektor. En motsvarande minskning har införts i ej i arbetskraften pensionärer. Vidare har 5

tusen personer flyttats inom ej i arbetskraften, från övriga till pensionärer. Även denna förändring är jämnt fördelad mellan könen.

Tabell 7.11. Extra blanketteffekter för åldersgruppen 25–34 år, mätt i tusentals personer.

Tidsserie	Inr. födda	Utr. födda	Män	Kvinnor	Totalt
Arbetslösa	5.0	2.0	3.5	3.5	7.0
Arbetslösa, stud.	–	–	1.0	1.0	2.0
I arbetskraften	5.0	2.0	3.5	3.5	7.0
Ej i arbetskraften	-5.0	-2.0	-3.5	-3.5	-7.0
Ej i arbetsk. övriga	–	–	-3.5	-3.5	-7.0

För åldersgruppen 65–74 år indikerar såväl flödesanalysen som tids-serieanalysen ett flöde från pensionärer till sysselsatta på drygt 10 tusen personer, vilket kan ha orsakats av frånvaron av beroende intervjuer i den nya blanketten. Vidare gäller, för utrikes födda sysselsatta i åldersgruppen, att skattningen av definitions- och populationsförändringarna på -0,5 tusen personer inte är kompatibel med konfidensintervallet från dubbelmätningen som är $10,1 \pm 9,2$, mätt i tusental.

Om inte en extra (positiv) blanketteffekt för antalet sysselsatta i åldersgruppen 65–74 år införs, så blir den procentuella förändringen av antalet sysselsatta alltför stor i början av länkingsperioden till följd av en negativ effekt av bytet av hjälpinformation. Vi har lagt till en extra blanketteffekt för sysselsatta på 12 tusen personer och den har fördelats mellan personer i arbete och frånvarande så att den totala blanketteffekten för frånvarande blir nära noll. Detta val har gjorts eftersom tidsserien för antalet frånvarande i åldersgruppen 65–74 år är mycket liten i absoluta tal i början av länkingsperioden.

Tabell 7.12 sammanfattar de extra blanketteffekter som införts för åldersgruppen 65–74 år i länkningen. Som synes har vi även infört några extra blanketteffekter inom gruppen ej i arbetskraften. Gruppen sjuka har fått ett tillskott på 10 tusen personer från gruppen pensionärer. Merparten av de extra blanketteffekterna i åldersgruppen antas gälla inrikes födda, men en mindre del har tilldelats utrikes födda för att kompensera för inkonsistensen för utrikes födda sysselsatta.

Vi noterar avslutningsvis att för gruppen ej i arbetskraften övriga, så indikerar dubbelmätningen en negativ blanketteffekt i åldersgruppen 65–74 år. Vi har dock inte tagit med en sådan effekt i länkningen eftersom denna grupp är mycket liten i början av länkingsperioden. En negativ extra blanketteffekt skulle därför ge negativa värden på den länkade tidsserien vissa månader.

Tabell 7.12. Extra blanketteffekter för åldersgruppen 65–74 år, mätt i tusentals personer.

Tidsserie	Inr. födda	Utr. födda	Män	Kvinnor	Totalt
Sysselsatta	10.0	2.0	6.0	6.0	12.0
I arbetskraften, stud.	10.0	2.0	6.0	6.0	12.0
Ej i arbetskraften	-10.0	-2.0	-6.0	-6.0	-12.0
Sysselsatta, i arbete	–	–	3.5	3.5	7.0
Sysselsatta, frånv.	–	–	2.5	2.5	5.0
Fast anställda	2.0	2.0	2.0	2.0	4.0
Tidsbegr. anst.	8.0	0.0	4.0	4.0	8.0
Anställda	10.0	2.0	6.0	6.0	12.0
Privat anställda	–	–	6.0	6.0	12.0
Ej i arbetskr. pens.	–	–	-11.0	-11.0	-22.0
Ej i arbetskr. sjuka	–	–	5.0	5.0	10.0

7.6 Förändring av säsongsmönster

För att undersöka om säsongsmönstret förändrades i samband med bytet av hjälpinformation har metodiken i avsnitt 5.6 använts. Vi har tagit fram två uppsättningar länkade tidsserier, en som inkluderar skattningen av förändringen i säsongsmönster och en som inte inkluderar denna skattning. För båda dessa uppsättningar har vi undersökt hur mycket säsongseffekterna förändrats mellan en månad och samma månad föregående år. Filosofin bakom analysen är att den uppsättning av länkade tidsserier som bäst fångar säsongsmönstret kommer att ha mer stabila säsongseffekter över tid.

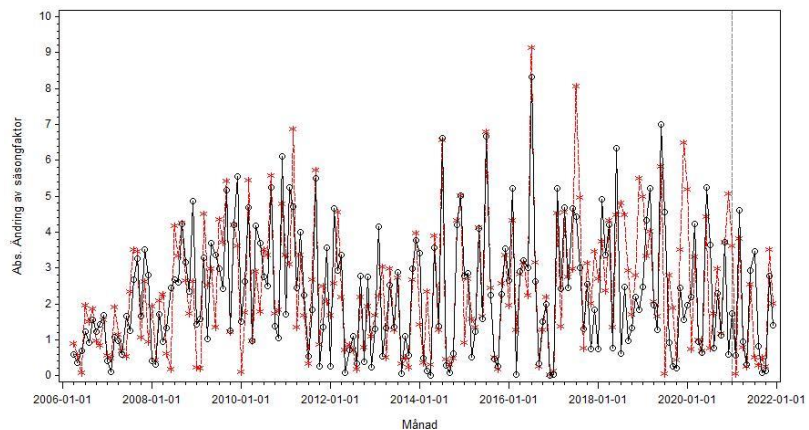
Figur 7.1 visar resultatet från analysen av säsongseffekter för det totala antalet sysselsatta 15–74 år (för denna variabel gick det att se ett ändrat säsongsmönster vid byte av hjälpinformation i den gamla blanketten). Som synes finns ingen avsevärd skillnad mellan stabiliteten hos säsongsmönstret för de båda uppsättningarna av tidsserier. Detta gäller för de flesta undersökta tidsserier. Eftersom kvaliteten hos de länkade tidsserierna inte förefaller öka nämnvärt om vi justerar för en förändring i säsongsmönstret, så har vi inte tagit med någon justering för nytt säsongsmönster i de länkade tidsserierna.

7.7 Arbetade timmar och outnyttjat arbetskraftsutbud

Tidsserierna för arbetade timmar påverkas inte av förändringen av definitionen av sysselsatta, eftersom denna förändring bara påverkar personer som var frånvarande från arbetet hela referensveckan och därmed inte arbetade några timmar. Även populationsförändringen och de extra blanketteffekterna förefaller ha liten inverkan på antalet

arbetade timmar, vilket också bekräftas av dubbelmätningen. Dessa effekter har dessutom olika tecken för sysselsatta och tar därför delvis ut varandra. Det följer därmed att tidsserierna för arbetade timmar främst påverkas av bytet av hjälpinformation.

Figur 7.1. Absolutbelopp av förändringen i säsongeffekt mellan en månad och samma månad föregående år för länkade tidsserier för antalet sysselsatta, båda könen 15–74 år, mätt i tusentals personer. Den ena länkade serien (svart linje med cirklar) tar hänsyn till en skattning av förändringen i säsongeffekt till följd av bytet av hjälpinformation och den andra länkade serien (röd linje med kryss) tar inte hänsyn till denna skattning.



I den länkning av AKU:s tidsserier som utfördes på makronivå 2022 togs länkade versioner av de viktigaste tidsserierna för antalet arbetade timmar fram. I mikroansatsen används dock bara styrvillkor som baseras på tidsserier för antalet personer och inte för antalet arbetade timmar. Baserat på de mikrolänkade vikterna kan vi likväl ta fram alla tidsserier för antalet arbetade timmar.

Tidsserierna för antalet potentiella arbetade timmar hos personer som tillhör det outnyttjade arbetskraftsutbudet, dvs personer som är arbetslösa, undersysselsatta eller latent arbetssökande, kan påverkas av både definitionsförändringen och populationsförändringen, men även av andra blanketteffekter. Till skillnad från för tidsserierna för arbetade timmar, så räcker det därför inte att bara betrakta brottskattningarna till följd av bytet av hjälpinformation för dessa tidsserier. Men precis som för tidsserierna för antalet arbetade timmar, så kan vi ta fram länkade tidsserier för dessa variabler med hjälp av mikroansatsen. I avsnitt 8.2.4 beskriver vi hur kvaliteten hos de mikrolänkade tidsserierna för arbetade timmar och det outnyttjade arbetskraftsutbudet har utvärderats.

7.8 Sammanfattning av brottskattningar

Tabell 7.13 sammanfattar alla brottskattningar som införts för tidsserier för antal personer i AKU, för båda könen, 15–74 år. Brottskattningarna ger upphov till en ökning av det relativa arbetslöshetstalet med 0,2 procentenheter, en minskning av det relativa arbetskraftstalet med 0,2 procentenheter och en minskning av sysselsättningsgraden med 0,4

procentenheter. Vi hänvisar till de publicerade tabellerna för värden på de länkade tidsserierna.

Tabell 7.13. Brottsskattningar som använts i länkningen av AKU:s tidsserier för personer för båda könen, 15–74 år, mätt i tusentals personer. Separata skattningar för bytet av hjälpinformation ($\hat{\Delta}^h$), definitionsförändringen ($\hat{\Delta}^d$), populationsförändringen ($\hat{\Delta}^p$) och resterande effekter av blankettbytet ($\hat{\Delta}^e$). Siffrorna för $\hat{\Delta}^d$ och $\hat{\Delta}^p$ utgörs av medelvärden över skattningarna för samtliga månader under 2020.

Tidsserie	$\hat{\Delta}^h$	$\hat{\Delta}^d$	$\hat{\Delta}^p$	$\hat{\Delta}^e$	Totalt
Sysselsatta	-30.0	-28.3	-6.0	14.0	-50.2
Arbetslösa	-4.0	4.2	-0.2	7.0	7.0
Arbetslösa, heltidsstud.	-1.1	2.1	-0.1	2.0	2.9
I arbetskraften	-34.0	-24.0	-6.2	21.0	-43.2
Befolkningen	0.0	0.0	-35.9	0.0	-35.9
Ej i arbetskraften	34.0	24.0	-29.7	-21.0	7.3
Sysselsatta, i arbete	-48.2	0.0	-4.8	9.0	-44.0
Sysselsatta, frånv.	18.2	-28.3	-1.2	5.0	-6.3
Fast anställda	-5.4	-21.5	-3.7	6.0	-24.6
Tidsbegränsat anst.	-26.9	-3.1	-2.3	8.0	-24.3
Anställda	-32.3	-24.6	-6.0	14.0	-48.9
Företagare	2.3	-3.7	0.0	0.0	-1.4
Privat anställda	-55.7	-17.4	-3.0	14.0	-62.1
Statligt anställda	7.3	-0.7	-1.6	0.0	5.0
Kommunalt anställda	13.3	-5.8	-1.4	0.0	6.1
Primärkom. anst.	8.0	-4.9	-1.4	0.0	1.7
Landstingskom. anst.	5.3	-0.9	0.0	0.0	4.4
Undersysselsatta	-6.9	-2.5	0.0	9.0	-0.4
Latent arbetssökande	0.0	0.0	-0.1	101	100.9
Ej i arbetskr. heltidsstud.	3.3	12.6	-0.9	0.0	15.0
Ej i arbetskr. pensionärer	18.5	1.9	-6.9	-19.0	-5.5
Ej i arbetskr. sjuka	8.9	0.0	-12.8	10.0	6.1
Ej i arbetskr. övriga	3.4	9.5	-9.2	-12.0	-8.3

8 Länkning med mikroansats

I avsnitt 6.2 beskrevs den metodik som har använts för att genomföra en länkning med mikroansats. Målsättningen har varit att justera uppräkningsvikterna för de personer som har svarat på AKU under länkingsperioden på ett sådant sätt att de justerade (mikrolänkade) vikterna blir konsistenta med de tidsserier som har länkats på makronivå. Givet de makrolänkade tidsserierna, så har arbetet med mikroansatsen bestått av tre komponenter: härledning av imputeringsmodeller för latent arbetssökande och frånvarande, val av styrvillkor till viktjusteringen samt kvalitetssäkring av de mikrolänkade tidsserierna, dvs av de tidsserier som erhålls vid uppräkningsperioden med de mikrolänkade vikterna. Imputeringsmodellerna finns beskrivna i bilagorna A och C. I avsnitt 8.1 beskriver vi valet av styrvillkor och i avsnitt 8.2 kvalitets-säkringen av de mikrolänkade tidsserierna. Avsnitt 8.3 beskriver de variabler för vilka de mikrolänkade tidsserierna inte uppfyller kvalitetskraven om jämförbarhet över tid och som därför har prickats vid publiceringen av länkade tidsserier.

8.1 Styrvillkor

Arbetet med att välja styrvillkor utfördes iterativt på följande sätt. Först valdes en startmängd bestående av de viktigaste tidsserierna i AKU och av de tidsserier som uppvisat störst tidsseriebrott i länkningen på makronivå. Baserat på denna startmängd härleddes mikrolänkade vikter och dessa vikter användes för att skapa mikrolänkade versioner av samtliga tidsserier som makrolänkats i etapp 1 och 2. Vid härledningen av de mikrolänkade tidsserierna så användes imputeringsmodellerna för latent arbetssökande och frånvarande för att justera svarsdata. De personer som baserat på sina svar i AKU bedömdes vara antingen latent arbetssökande eller påverkade av definitionsförändringen fick ändrat svarsdata. Dessutom togs de personer som svarat att de var värnpliktiga eller kontraktsanställda soldater bort från svarsmängden.

Vi jämförde de mikrolänkade tidsserierna med de makrolänkade tidsserierna för samtliga månader 2020 och adderade därefter de tidsserier där avvikelsen var som störst till styrvillkoren. Proceduren upprepades sedan och nya tidsserier adderades till styrvillkoren. Efter ett antal iterationer bedömdes avvikelserna mellan de mikrolänkade och de makrolänkade tidsserierna vara acceptabla. Avvikelserna var då som högst ungefär fem tusen personer enskilda månader och avvikelserna hade i dessa fall olika tecken olika månader, vilket tyder på att avvikelserna var slumpmässiga. Konsekventa avvikelser uppgick då till som högst ett par tusen personer.

De slutliga styrvillkoren består av sex styrvektorer som var och en för sig delar in hela målpopulationen i ömsesidigt disjunkta grupper. För varje grupp beräknas för varje månad under länkningsperioden hur stor andel den gruppen utgör av den makrolänkade populationen 15–74 år.

Styrvektor 1 består i huvudsak av åldersklasser korsat med kön och arbetskraftsstatus (i arbete, frånvarande, arbetslösa, ej i arbetskraften) samt uppdelning för latent arbetssökande. För åldersgruppen 15 år görs ingen uppdelning efter latent arbetssökande och för den äldsta åldersklassen är arbetskraftstatus reducerat till endast tre grupper. Tabell 8.1 visar de totalt 84 grupperna i styrvektor 1.

Tabell 8.1. Styrvektor 1 som i huvudsak ges av åldersklasser korsat med kön och arbetskraftsstatus. Numreringen i cellerna motsvarar platsen i styrvektorn i estimeringsprogrammet.

Ålder	I arbete		Frånvarande		Arbetslösa		Ej i arbetskraften	
	Man	Kvinna	Man	Kvinna	Man	Kvinna	Man	Kvinna
15 år	1	40	9	48	17	56	25	64

Ålder	I arbete		Frånvarande		Arbetslösa		Ej latent arbetssökande		Latent arbetssökande	
	Man	Kv.	Man	Kv.	Man	Kv.	Man	Kv.	Man	Kv.
16–19 år	2	41	10	49	18	57	26	65	33	72
20–24 år	3	42	11	50	19	58	27	66	34	73
25–34 år	4	43	12	51	20	59	28	67	35	74
35–44 år	5	44	13	52	21	60	29	68	36	75
45–54 år	6	45	14	53	22	61	30	69	37	76
55–59 år	7	46	15	54	23	62	31	70	38	77
60–64 år	8	47	16	55	24	63	32	71	39	78

Ålder	I arbetskraften		Ej latent arbetssökande		Latent arbetssökande	
	Man	Kvinna	Man	Man	Kvinna	Man
65–74 år	79	82	80	83	81	84

Styrvektor 2 består av en sektorsindelning om 7 grupper. Personer som saknar sektorstillhörighet i svarsdata betraktas som privat anställda i estimationen. Styrvektor 2 redovisas i tabell 8.2.

Tabell 8.2. Styrvektor 2 som ges av sektorstillhörighet korsat med kön. Numreringen i cellerna motsvarar platsen i styrvektorn i estimeringsprogrammet.

Sysselessatt utomlands	Privat anställd		Statligt anställd	Kommun. anställd	Inrikes företagare	Övriga
	Man	Kvinna				
1	2	3	4	5	6	7

Styrvektor 3 består av åldersklasser korsat med indelning av sysselsatta i fast anställda, tidsbegränsat anställda och företagare, indelning av ej i arbetskraften i undergrupperna heltidsstuderande, pensionärer, sjuka och övriga samt indelning av arbetslösa i heltidsstuderande och övriga. Vissa av kategorierna är även uppdelade efter kön. Tabell 8.3 visar de totalt 48 grupperna i styrvektor 3.

Tabell 8.3. Styrvektor 3 som ges av åldersklasser korsat med utvalda undergrupper till sysselsatta, arbetslösa respektive ej i arbetskraften, i vissa fall även korsat med kön. Numreringen i cellerna motsvarar platsen i styrvektorn i estimeringsprogrammet.

Ålder		Fast anställd		Tidsbegränsat anställd		Företagare		
25–34 år		1		2		3		
35–44 år		4		5		6		
65–74 år		7		8		9		
Ålder	Ej i arbetskraften heltidsstuderande		Ej i arbetskraften pensionärer		Ej i arbetskraften sjuka		Ej i arbetskraften övriga	
	Män	Kv.	Män	Kv.	Män	Kv.	Män	Kv.
16–19 år	29		30		31		32	
20–24 år	33		34		35		36	
25–34 år	10	14	13	17	11	15	12	16
35–44 år	37		38		39		40	
45–54 år	41		42		43		44	
55–59 år	21		18		19		20	
65–74 år	22	26	23	27	24	28	25	
Arbetslösa, 25–34 år, heltidsstuderande		Resterande arbetslösa		Resterande sysselsatta		Resterande ej i arbetskraften		
45		47		46		48		

Styrvektor 4 består av arbetskraftstatus uppdelat efter sysselsatta som är undersysselsatta, sysselsatta som inte är undersysselsatta, arbetslösa, ej i arbetskraften som är latent arbetssökande samt ej i arbetskraften som inte är latent arbetssökande korsat med födelseland i kategorierna inrikes födda respektive utrikes födda. Tabell 8.4 visar de 10 grupperna i styrvektor 4.

Tabell 8.4. Styrvektor 4 som ges av arbetskraftstatus korsat med födelseland. Numreringen i cellerna motsvarar platsen i styrvektorn i estimeringsprogrammet.

Födelseland	Under- sysselsatt	Sysselsatt, ej undersysselsatt	Arbetslös	Latent arbets- sökande	Ej i arbets- kraften, ej latent arbets- sökande
Inrikes född	1	2	3	5	4
Utrikes född	6	7	8	10	9

Styrvektor 5 består av åldersklasser korsat med födelseland och arbetskraftstatus. Tabell 8.5 visar de totalt 54 grupperna i styrvektor 5.

Tabell 8.5. Styrvektor 5 som ges av åldersklasser korsat med födelseland och arbetskraftstatus. Numreringen i cellerna motsvarar platsen i styrvektorn i estimeringsprogrammet.

Ålder	Sysselsatta		Arbetslösa		Ej i arbetskraften	
	Inrikes födda	Utrikes födda	Inrikes födda	Utrikes födda	Inrikes födda	Utrikes födda
15 år	1	10	19	28	37	46
16–19 år	2	11	20	29	38	47
20–24 år	3	12	21	30	39	48
25–34 år	4	13	22	31	40	49
35–44 år	5	14	23	32	41	50
45–54 år	6	15	24	33	42	51
55–59 år	7	16	25	34	43	52
60–64 år	8	17	26	35	44	53
65–74 år	9	18	27	36	45	54

Styrvektor 6 består av den hjälpvektor med regionstillhörighet som ingår i den ordinarie hjälpinformationen i AKU. Till skillnad mot de andra fem styrvektorena, så består denna vektor alltså inte av makrolänkade tidsserier. Styrvektor 6 redovisas i tabell 8.6.

Tabell 8.6. Styrvektor 6 som ges av region. Numreringen i cellerna motsvarar platsen i styrvektorn i estimeringsprogrammet.

Stockholms län exkl. Stockholms kommun	1	Västmanlands län	14
Uppsala län	2	Dalarnas län	15
Södermanlands län	3	Gävleborgs län	16
Östergötlands län	4	Västernorrlands län	17
Jönköpings län	5	Jämtlands län	18
Kronobergs län	6	Västerbottens län	19
Kalmar län	7	Norrbottnens län	20
Gotlands län	8	Stockholms kommun	21
Blekinge län	9	Malmö kommun	22
Hallands län	10	Göteborgs kommun	23
Västra Götalands län exkl. Göteborgs kommun	11	Delar av Skåne län 1	24
Värmlands län	12	Delar av Skåne län 2	25
Örebro län	13	Delar av Skåne län 3	26

Ett antal av styrvillkoren i de sex styrvektorerna går att skriva som linjärkombinationer av andra styrvillkor. Exempelvis är summan av kategorierna 1, 9, 17, 25, 40, 48, 56 och 64 i styrvektor 1 lika med summan av kategorierna 1, 10, 19, 28, 37 och 46 i styrvektor 5 eftersom båda dessa summor är lika med andelen 15-åringar i målpopulationen. Totalt kan 45 styrvillkor tas bort på grund av linjära samband mellan styrvillkoren. Det totala antalet styrvillkor är därmed $84 + 7 + 48 + 10 + 54 + 26 - 45 = 184$. Av dessa anger 25 styrvillkor regions-tillhörighet, så det totala antalet styrvillkor som bestäms av makrolänkade tidsserier uppgår till 159. Detta är ungefär en åttondel av det totala antalet tidsserier som makrolänkats i etapp 1 och 2.

8.2 Kvalitetssäkring av mikrolänkade tidsserier

Kvalitetssäkringsarbetet av de mikrolänkade tidsserierna har bestått av ett antal olika moment som beskrivs i detalj i detta avsnitt.

8.2.1 Kvalitetssäkring av vikter

Ett nödvändigt villkor för att bevara summakonsistensen mellan alla AKU:s tidsserier är att samtliga mikrolänkade vikter är icke-negativa. Redan i samband med kalibreringen analyserades de mikrolänkade vikterna i detta avseende. När negativa vikter uppkom så analyserades datamaterialet för att upptäcka bakomliggande orsaker till de negativa vikterna. Efter att orsakerna åtgärdats kördes nya vikter fram för hela länkningsperioden.

En orsak till negativa vikter var att styrvillkoren var alltför detaljerade i förhållande till antalet svarande i mikrodata som hade den kombination av egenskaper som angavs av styrvillkoret. Detta inträffade exempelvis vissa månader när åldersgruppen 65–74 år var indelad i alltför många kategorier efter arbetskraftstatus och kön. Problemet åtgärdades genom att välja en grövre indelning i styrvektorn för just den åldersgruppen (se tabell 8.1). En annan orsak till negativa vikter var att styrvillkoren sattes till att räkna upp en grupp som helt saknade observationer vissa månader. Detta inträffade bland annat för yngre latent arbetssökande och löstes med hjälp av imputering för gruppen i fråga.

För att säkerställa en hög kvalitet hos de mikrolänkade tidsserier som inte ingår i styrvillkoren, så krävs det att de mikrolänkade vikterna har en rimlig spridning och följer en fördelning som saknar extrema värden. Mycket stora värden på vikterna innebär att enskilda observationer får stor påverkan på statistiken samtidigt som många andra observationer får alltför liten vikt. För att undvika gränsdragningar gällande vad som kan anses vara en alltför stor vikt så analyserades kvoten mellan de mikrolänkade vikterna och originalvikterna, båda normerade mot respektive totalpopulation på månadsbasis. Normeringen är nödvändig eftersom originalvikterna och de mikrolänkade vikterna räknas upp mot två olika populationer till följd av populationsförändringen. Om vi låter

v_i^o beteckna originalvikten för individ i under månad t och v_i^L motsvarande mikrolänkade vikt, så ges kvoten k_i av

$$k_i = \frac{\left(v_i^L / \sum_{i \in s_t} v_i^L \right)}{\left(v_i^o / \sum_{i \in s_t} v_i^o \right)},$$

där s_t är svarsmängden under månad t . Värden på k_i som understeg 0,5 eller översteg 2 valdes sedan ut för ytterligare kontroll. Med andra ord analyserades de fall då mikrolänkningen resulterade i uppräkningsvikter som mer än fördubblades eller halverades jämfört med de ursprungliga uppräkningsvikterna.

För de styrvektorer som beskrevs i avsnitt 8.1, så var det totalt 493 svarande som hade ett värde på k_i som understeg 0,5 och totalt 211 svarande som hade ett värde på k_i som översteg 2. Det totala antalet svarande i AKU under hela länkingsperioden 2005–2020 uppgick samtidigt till över 3 150 000 personer, så det var förhållandevis mycket få personer som uppvisade starkt avvikande värden på k_i .

De svarande som uppfyllde villkoret $k_i < 0,5$ bestod till största delen av inrikes födda personer som var utanför arbetskraften i gruppen övriga. Dessa vikter bedömdes som oproblematisks, dels eftersom den absoluta skillnaden mellan originalvikterna och de mikrolänkade vikterna var liten, dels eftersom det var få personer per månad som uppfyllde villkoret $k_i < 0,5$.

Av de 211 personer som uppfyllde villkoret $k_i > 2$ så var 153 imputerade latent arbetssökande. Det faktum att denna grupp har blivit mycket större i den nya blanketten kan förklara varför deras mikrolänkade vikter skiljde sig åt så pass mycket från originalvikterna. De övriga personerna som uppfyllde villkoret $k_i > 2$ var framför allt personer i åldersgruppen 65–74 år som befann sig utanför arbetskraften och som inte var latent arbetssökande. Ingen av dessa 211 personer hade så pass hög mikrolänkad vikt i absoluta tal att det föranledde vidare åtgärd.

8.2.2. Kvalitetssäkring av mikrolänkade tidsserier som länkats på makronivå

Ett drygt tusental makrolänkade tidsserier publicerades i två etapper under våren/sommaren 2022. Av dessa har 159 tidsserier använts som styrvillkor i mikrolänkningen och de mikrolänkade versionerna av dessa tidsserier förväntas därför överensstämja helt med motsvarande makrolänkade tidsserier. Vi har säkerställt att så också är fallet.

För övriga tidsserier som tidigare länkats på makronivå förväntas vissa avvikelser förekomma eftersom de mikrolänkade vikterna inte är kalibrerade mot dessa tidsserier. Kvalitetssäkring av dessa tidsserier har

fokuserat på rimlighetsbedömningar av genomsnittliga nivåmässiga och procentuella differenser mellan uppräknade mikrolänkade tidsserier och motsvarande tidigare makrolänkade tidsserier. Utöver genomsnittliga nivåskillnader har även rimlighetsbedömningar gjorts av utvecklingen över tid under länkningsperioden 2005–2020. Vidare omfattar den kvalitetssäkring som beskrivs i avsnitt 8.2.3 även tidsserier som tidigare makrolänkats.

8.2.3. Kvalitetssäkring av mikrolänkade tidsserier som inte länkats på makronivå

Genom uppräkning med mikrolänkade vikter är det i princip möjligt att ta fram alla tänkbara tidsserier i AKU. Kvalitetssäkring av tidsserier som inte tidigare makrolänkats har fokuserat på det urval av publicerade tidsserier som omfattas av AKU:s grundtabeller och Statistikdatabasen (SSD). Bland dessa tidsserier har i sin tur ett urval av centrala tidsserier valts ut för särskilt prioriterad kvalitetssäkring.

Kvalitetssäkring av mikrolänkade tidsserier har genomförts med flera olika ansatser, mestadels utifrån ett gemensamt framtaget datamaterial av tidsserier uppräknade med originalvikter respektive mikrolänkade vikter. Rimlighetsbedömningar av genomsnittliga länkningskorrektioner, nivåmässigt och procentuellt, har gjorts med utgångspunkt i de skattningar av tidsseriebrott som gjorts vid tidigare makrolänkning. Exempelvis har det bedömts vara rimligt att mikrolänkade tidsserier relaterade till latent arbetslösa och frånvarande, samt tidsserier för vissa åldersgrupper, korrigeras i högre utsträckning än övriga tidsserier.

En stor del av kvalitetssäkringen har baserats på ansatser som innebär anpassning av tidsseriemodeller i syfte att identifiera ett nivåskifte i trendserien vid tidpunkten januari 2021. Skattningar av tidsseriebrott (nivåskiften) har gjorts för cirka 7 700 uppräknade mikrolänkade tidsserier med tidsseriemodeller av typen X13-ARIMA. Kvalitetssäkringen med tidsseriemodeller har skett i två steg där det första steget utgjorts av en automatiserad analys med programvaran Proc X13 i SAS.

Den automatiserade ansatsen togs fram dels för att få en god överblick över brottsskattningarnas storlek och signifikansnivå, dels för att på ett tidseffektivt sätt kunna återupprepa kvalitetssäkringsarbetet när nya mikrolänkade vikter togs fram under den iterativa processen att bestämma styrvillkor. Problematiken med "falska positiva" tidsseriebrott i enskilda tidsserier vid multipla skattningar har hanterats genom att identifiera gemensamma mönster över tabeller, huvudserier (sysselsatta, arbetslösa, etc.) och demografiska grupper (kön, åldersklasser samt inrikes respektive utrikes födda).

Tabeller där avvikande mönster identifierats i den automatiserade kvalitetssäkringen har därefter analyserats vidare visuellt i

programvaran JDemetra. Snarlika modellspecifikationer har använts i X13 och JDemetra.

Även andra ansatser har använts i kvalitetssäkringsarbetet. Exempelvis har vi jämfört skattade kvartalsflöden för att analysera storleken på tidsseriebrott till följd av definitionsförändringen för vissa grupper av långtidsfrånvarande sysselsatta, med respektive utan mikrolänkade vikter, uppdelat på olika frånvaroorsaker och frånvarotidens längd.

8.2.4. Kvalitetssäkring av tidsserier för arbetade timmar och outnyttjat arbetskraftsutbud

Som nämndes i avsnitt 7.7 så har inga styrvillkor avseende antalet arbetade timmar eller antalet potentiella arbetade timmar för det outnyttjade arbetskraftsutbudet använts i kalibreringen av de mikrolänkade vikterna. Förhoppningen var att vi med en väl utförd mikrolänkning av antalet personer i olika grupper även skulle erhålla länkade tidsserier av hög kvalitet för antalet arbetade timmar och det outnyttjade arbetskraftsutbudet. Kvalitetssäkring av dessa tidsserier var därför av särskild vikt.

Tidsserier avseende antalet arbetade timmar, med uppdelningar på näringsgren och kön, samt medelarbetstid uppdelat på kön och åldersklasser, ingår i de tabellpaket som kvalitetssäkrats enligt den metod som beskrivs i avsnitt 8.2.1–8.2.2 och hör till de tidsserier som prioriterats särskilt i kvalitetssäkringsarbetet. I dessa tabellpaket ingår även tidsserier avseende det outnyttjade arbetskraftsutbudet uppdelat på arbetslösa, latent arbetslösa och undersysselsatta, samt på kön och åldersklasser.

En annan del av kvalitetssäkringen av tidsserier för antalet arbetade timmar har bestått av analys av tidsserier för antalet sysselsatta personer per näringsgren och arbetstid. Detta beror både på att antalet arbetade timmar redovisas per näringsgren och på att länkningen av antalet arbetade timmar i hög utsträckning förväntas bero på länkningen av antalet heltids- respektive deltidssysselsatta personer inom respektive näringsgren. För antalet arbetade timmar var det därför av särskild vikt att kvalitetssäkra utvecklingen av antalet sysselsatta per näringsgren samt arbetstid. Tidsserier för antalet sysselsatta per näringsgren har kvalitetssäkrats med SNI-indelning på tvåsiffernivå. Vidare har SNI-indelningar på bokstavs nivå kvalitetssäkrats korsat med ytterligare uppdelningar, efter anknytningsgrad till arbetsmarknaden (anställda, fast anställda och sysselsatta totalt) samt efter arbetstid (kort deltid, lång deltid och heltid). Utöver skattningar av tidsseriebrott i januari 2021 har de mikrolänkade tidsserierna även inspekterats visuellt avseende kvarstående tidsseriebrott samt utvecklingen över tid av länkingskorrektionernas storlek.

8.3 Tidsserier som ej är jämförbara över tid

Inom ramen för arbetet med att kvalitetssäkra mikrolänkade tidsserier har ett stort antal tidsserier analyserats. Det har då konstaterats att för ett mindre antal variabler är tidsseriebrottet av sådan karaktär att det inte kan hanteras med mikrolänkning och de mikrolänkade tidsserierna för dessa variabler är därför inte jämförbara över tid. Till denna kategori hör tidsserier för överenskommen arbetstid och tidsserier för övertid. För dessa variabler har mikrolänkade tidsserier därför inte publicerats för perioden 2005–2020.

Utöver dessa tidsserier konstaterades vid kvalitetssäkringsarbetet av mikrolänkningen även en bristande jämförbarhet över tid för tidsserier avseende arbetslöshetstidens längd, framför allt för redovisning av längre arbetslöshetstid. Denna bristande jämförbarhet hanteras genom imputeringar av insamlade uppgifter enligt nya ramlagen, se SCB (2023) för en mer detaljerad beskrivning.

8.3.1 Överenskommen arbetstid

Överenskommen arbetstid skapas utifrån antal timmar som uppgiftslämnaren har som överenskommen arbetstid. Redovisningen sker efter kort deltid (1–19 timmar), lång deltid (20–34 timmar), heltid (35–timmar) samt genomsnittlig överenskommen arbetstid.

Det finns två huvudanledningar till att dessa tidsserier inte är jämförbara över tid. Den ena huvudanledningen är att information om överenskommen arbetstid under perioden 2021–2022 enbart samlades in för huvudsysslan och inte för bisysslan, vilket skedde både till och med 2020 och från och med 2023. Den andra huvudanledningen är det ökade bortfall som uppkommit i frågan om överenskommen arbetstid i och med att det i den nya ramlagen är tillåtet med partiellt bortfall i form av svarsalternativen ”vet ej” och ”vill ej svara”. Detta bortfall har främst påverkat tidsbegränsat anställda.

Båda dessa förändringar har bidragit till att antalet personer med kort respektive lång deltid ligger på en lägre nivå i och med den nya ramlagen. På samma sätt ligger tidsserien för genomsnittlig överenskommen arbetstid på en högre nivå i och med den nya ramlagen. Som tidigare nämnts hanteras denna brist i jämförbarhet genom att mikrolänkade tidsserier avseende överenskommen arbetstid ej publiceras för perioden 2005–2020.

8.3.2 Övertid

Uppgifter om övertid skapas utifrån antalet timmar som uppgiftslämnaren har svarat att denna har arbetat mer än vanligt under referensveckan och om dessa timmar var övertid med eller utan ersättning. Redovisningen sker utifrån antalet personer som har arbetat övertid med uppdelat efter om övertiden har skett med eller utan

ersättning samt antalet övertidstimmar uppdelat efter om övertiden har skett med eller utan ersättning.

Den främsta förklaringen till att ett tidsseriebrott har uppkommit i dessa tidsserier är att frågorna om övertid ställs på ett annat sätt än tidigare. Nu efterfrågas vilken typ av timmar som arbetats utöver den vanliga arbetstiden. Tidigare ställdes separata frågor avseende övertid.

Tidserierna för övertid har i och med den nya ramlagen hamnat på en lägre nivå än tidigare och detta gäller både övertid totalt och efter uppdelning i grupperna ersättning respektive ej ersättning. Då det inte har varit möjligt att länka för detta tidsseriebrott, så publiceras mikrolänkade tidsserier avseende övertid inte för perioden 2005–2020.

Referenser

EU (2019), Regulation 2019/1700 of the European Parliament and of the Council of 10 October 2019 establishing a common framework for European statistics relating to persons and households. Tillgänglig från: <http://data.europa.eu/eli/reg/2019/1700/oj>

SCB (2023), *Förändringar vid ompublicering av AKU, år 2021-*.

Statistics Sweden (2012), Research and Development – Methodology reports from Statistics Sweden, *Linking a system of time series. Recalculation of Labour Force Surveys 1987–1992*.

Statistics Sweden (2013a), Background Facts, Population and Welfare Statistics 2013:2, *Revising Surveys – Linking Old and New Data*.

Statistics Sweden (2013b), Background Facts, Labour and Education Statistics 2013:5, *Linking of central series in Labour Force Surveys (LFS) 1970–1986*.

Bilagor

A. Modell för skattning av definitionsförändring

Vi beskriver i denna bilaga den modell som har använts för att skatta definitionsförändringen för sysselsatta och som även har legat till grund för imputering av personer som med stor sannolikhet påverkas av definitionsförändringen.

A.1. Analys av personer som påverkas av definitionsförändringen

Som utgångspunkt i analysen använder vi data om de 401 personer som svarade på ett sådant sätt på extrafrågorna i AKU under perioden februari–december 2020 att de efter införandet av den nya ramlagen hade påverkats av definitionsförändringen. Målsättningen här är att hitta kriterier som särskiljer dessa personer från hela mängden med 12 077 personer som var frånvarande från arbetet hela referensveckan under någon av dessa elva månader.

Vi undersöker först frånvaroskäl i tabell A.1 och ser att 383 av de 401 personerna som påverkades av definitionsförändringen har angivit något av de fem frånvaroskälen ”Ledig för studier utan lön”, ”Annan tjänstledighet”, ”Permitterad”, ”Arbetsbrist” eller ”Övrigt”. Ytterligare fyra andra frånvaroskäl uppvisar också en relativt hög andel av personer som påverkas av definitionsförändringen (25–40%). Totalt är det dock under 11 månader bara 10 personer som har angivit något av dessa frånvaroskäl och som påverkas av definitionsförändringen och vi har därför försummat dessa frånvaroskäl i modellen.

Totalt var det 1 330 personer som angav något av de fem frånvaroskäl som räknades upp i föregående stycke. I genomsnitt 29% av dessa personer påverkades av definitionsförändringen. Vi utför ett antal homogenitetstest för att undersöka vilka bakgrundsvariabler som har signifikant påverkan på benägenheten att påverkas av definitionsförändringen. De undersökta variablerna är intervjumånad, kön, ålder (med samma indelning som i länkningen), födelseland, region, näringsgren, utbildningsnivå, frånvarotidens längd och anknytningsgrad till arbetsmarknaden. Baserat på p-värdena (signifikansnivå 5%) för homogenitetstesten väljer vi att inte undersöka variablerna intervjumånad, födelseland och region vidare.

Vi undersöker härnäst de fem frånvaroskälen var för sig. Vi gör denna uppdelning eftersom sannolikheten att påverkas av definitionsförändringen varierar mellan 17% och 54% för de fem frånvaroskälen och det är därför olämpligt att använda samma modell för alla fem frånvaroskälen. För samtliga fem frånvaroskäl har vi utfört

homogenitetstest med avseende på de sex variablerna kön, ålder, näringsgren, utbildningsnivå, frånvarotidens längd och anknytningsgrad till arbetsmarknaden.

Tabell A.1. Antal personer med olika frånvaroskäl som påverkas (Def) respektive inte påverkas (Ej def) av definitionsförändringen under perioden februari – december 2020.

Frånvaroskäl	Def	Ej def	Frånvaroskäl	Def	Ej def
Sjuk/läkarbesök/tandläkarbesök	1	2188	Arbetskonflikt	2	3
Semester/lov för skolpersonal	0	6153	Olämpligt väder	2	6
Värnplikt/civilplikt, i Sverige eller utomlands	2	3	Friårsledig	0	1
Repövning	0	1	Flexledig	0	26
Betald föräldraledighet vid födsel/adoption	4	1481	Kompleddig/ledig mot inarbetning	0	62
Annan betald föräldraledighet	2	256	Arbetstidens förläggning	1	357
Obetald tjänstledighet för vård av barn	4	9	Helgdag/helgdagsafton eller annan arbetsfri dag	0	174
Ledig för studier utan lön	173	146	Arbete i bisyssla	0	1
Annan tjänstledighet	37	67	Vet ej	0	6
Permitterad	46	227	Övrigt	58	232
Arbetsbrist	69	275	Partiellt bortfall	0	2

För frånvaroskäl ”Ledig för studier utan lön” är det homogenitetstestet för anknytningsgrad till arbetsmarknaden som har lägst p-värde och den variabeln bör alltså ha störst påverkan på benägenheten att påverkas av definitionsförändringen. Vi delar därför upp mängden av personer med detta frånvaroskäl (319 personer) i en mängd med fast anställda (254 personer) och en mängd med tidsbegränsat anställda och företagare (63 personer).

För fast anställda med frånvaroskäl ”Ledig för studier utan lön” är näringsgren och utbildningsnivå ej längre signifikanta och undantas därför från den fortsatta analysen. Tabell A.2 visar vidare att grupperna med 3–4 veckors frånvaro, 5–12 veckors frånvaro och minst 13 veckors frånvaro har liknande sannolikhet att påverkas av definitionsförändringen (67, 70 respektive 72 procent) och vi slår därför ihop dessa grupper. Grupperna med 1 respektive 2 veckors frånvaro har också liknande sannolikhet att påverkas av definitionsförändringen (17 respektive 30 procent och vi slår ihop dessa grupper. Uppdelat på frånvarotidens längd har ålder och kön inte längre en signifikant påverkan och vi använder därför bara frånvarotidens längd i modellen för fast anställda med frånvaroskäl ”Ledig för studier utan lön”. Den slutliga modellen redovisas i tabell A.5.

För tidsbegränsat anställda och företagare med frånvaroskäl "Ledig för studier utan lön" är det bara frånvarotidens längd som har signifikant påverkan och används i den fortsatta analysen. På grund av få personer i grupperna (se tabell A.2), så slår vi ihop grupperna med 5–12 veckors frånvaro och minst 13 veckors frånvaro samt även grupperna med 1, 2 och 3–4 veckors frånvaro. Modellen för tidsbegränsat anställda och företagare med frånvaroskäl "Ledig för studier utan lön" redovisas i tabell A.5).

Även för frånvaroskäl "Annan tjänstledighet" är det bara frånvarotidens längd som har signifikant påverkan på benägenheten att påverkas av definitionsförändringen och som används i den fortsatta analysen. På grund av få personer i grupperna (se tabell A.2), så slår vi ihop grupperna med 1, 2 och 3–4 veckors frånvaro. Modellen för frånvaroskäl "Annan tjänstledighet" redovisas i tabell A.5).

Tabell A.2. Antal personer med frånvaroskäl "Ledig för studier utan lön" eller "Annan tjänstledighet" under perioden februari–december 2020 som påverkas (Def) respektive inte påverkas (Ej def) av definitionsförändringen uppdelat på frånvaroskäl, frånvarotidens längd samt, delvis, anknytningsgrad till arbetsmarknaden.

Frånvarotidens längd	Ledig för studier utan lön				Annan tjänstledighet	
	Fast anställda		Tidsbegränsat anställda och företagare			
	Def	Ej def	Def	Ej def	Def	Ej def
1 vecka	4 (17%)	19 (83%)	0 (0%)	16 (100%)	2 (8%)	22 (92%)
2 veckor	3 (30%)	7 (70%)	1 (10%)	9 (90%)	1 (10%)	9 (90%)
3–4 veckor	16 (67%)	8 (33%)	0 (0%)	15 (100%)	2 (20%)	8 (80%)
5–12 veckor	33 (70%)	14 (30%)	3 (21%)	11 (79%)	8 (32%)	17 (68%)
Minst 13 veckor	108 (72%)	42 (28%)	4 (50%)	4 (50%)	24 (69%)	11 (31%)
Totalt	164 (65%)	90 (35%)	8 (13%)	55 (87%)	37 (37%)	64 (63%)

De tre frånvaroskälen "Permitterad", "Arbetsbrist" och "Övrigt" har alla ungefär samma beroende på antalet frånvaroveckor (se tabell A.3) och ett helt annat beroende på antalet frånvaroveckor än de båda frånvaroskälen "Ledig för studier utan lön" och "Annan tjänstledighet". Vi analyserar därför de tre frånvaroskälen "Permitterad", "Arbetsbrist" och "Övrigt" tillsammans. Baserat på tabell A.3 så slår vi även ihop grupperna med 1 och 2 veckors frånvaro och grupperna med 5–12 veckors frånvaro och minst 13 veckors frånvaro.

För de tre frånvaroskälen "Permitterad", "Arbetsbrist" och "Övrigt" är kön, näringsgren och utbildningsnivå inte längre signifikanta, men däremot ålder, frånvarotidens längd och anknytningsgrad till arbetsmarknaden. Om vi delar upp gruppen efter anknytningsgrad till

arbetsmarknaden, så är ålder inte längre signifikant. Tabell A.4 visar hur benägenheten att påverkas av definitionsförändringen beror på frånvarotidens längd och anknytningsgrad till arbetsmarknaden. Modellen baseras på dessa båda variabler och redovisas i tabell A.5).

Tabell A.3. Antal personer med frånvaroskäl "Permitterad", "Arbetsbrist" eller "Övrigt" under perioden februari–december 2020 som påverkas (Def) respektive inte påverkas (Ej def) av definitionsförändringen uppdelat på frånvaroskäl och frånvarotidens längd.

Frånvarotidens längd	Permitterad		Arbetsbrist		Övrigt	
	Def	Ej def	Def	Ej def	Def	Ej def
1 vecka	3 (5%)	52 (95%)	6 (6%)	92 (94%)	6 (7%)	83 (93%)
2 veckor	4 (9%)	41 (91%)	4 (11%)	34 (89%)	2 (7%)	26 (93%)
3–4 veckor	8 (15%)	46 (85%)	12 (19%)	50 (81%)	8 (17%)	38 (83%)
5–12 veckor	21 (26%)	61 (74%)	19 (26%)	54 (74%)	22 (32%)	47 (68%)
Minst 13 veckor	10 (27%)	27 (73%)	28 (38%)	45 (72%)	20 (34%)	38 (66%)
Totalt	46 (17%)	227 (83%)	69 (20%)	275 (80%)	58 (20%)	232 (80%)

Tabell A.4. Antal personer med frånvaroskäl "Permitterad", "Arbetsbrist" eller "Övrigt" under perioden februari–december 2020 som påverkas (Def) respektive inte påverkas (Ej def) av definitionsförändringen uppdelat på anknytningsgrad till arbetsmarknaden och frånvarotidens längd.

Frånvarotidens längd	Permitterad, arbetsbrist eller övrigt					
	Fast anställda		Tidsbegränsat anställda		Företagare	
	Def	Ej def	Def	Ej def	Def	Ej def
1–2 veckor	13 (7%)	173 (93%)	6 (6%)	88 (94%)	6 (8%)	67 (92%)
3–4 veckor	14 (16%)	72 (84%)	4 (13%)	27 (87%)	10 (22%)	35 (78%)
Minst 5 veckor	54 (30%)	126 (70%)	17 (20%)	68 (80%)	49 (39%)	78 (61%)
Totalt	81 (18%)	371 (82%)	27 (13%)	183 (87%)	65 (26%)	180 (74%)

Tabell A.5 sammanfattar modellen som vi använder för att bestämma antalet personer som påverkas av definitionsförändringen. Som synes i tabellen så ger modellen för varje kombination av frånvaroskäl, frånvarotidens längd och anknytningsgrad till arbetsmarknaden en sannolikhet att påverkas av definitionsförändringen. Det finns 15 olika kombinationer av dessa variabler som motsvarar en positiv sannolikhet (en för varje fält i tabellen). För varje månad under länkningsperioden har vi, baserat på den svarsdata som har samlats in i AKU den månaden, delat in mängden frånvarande personer i 16 olika grupper (de 15 grupperna med nollskild sannolikhet i tabell A.5 och en grupp bestående av resterande personer). För var en av dessa grupper har vi

slumpmässigt valt ut en så stor andel personer som tabell A.5 anger. Den sammanlagda vikten för dessa personer utgör skattningen av antalet personer som påverkas av definitionsförändringen den månaden. Baserat på de utvalda personernas AKU-svar kan vi även skatta definitionsförändringens påverkan på samtliga undergrupper till sysselsatta, såsom anknytningsgrad till arbetsmarknaden, sektor etc.

Tabell A.5. Sammanfattning av modell för definitionsförändringen för sysselsatta. Tabellen visar modellens skattningar av andelen personer, mätt i procent, i olika kategorier av frånvarande som påverkas av definitionsförändringen.

Frånvaroskäl	Anknytnings- grad	Frånvarotidens längd			
		1–2 veckor	3–4 veckor	5–12 veckor	Minst 13 veckor
Ledig för studier utan lön	Fast anställda	20	70		
	Tidsbegränsat anställda och företagare	0		30	
Annan tjänstledighet	Samtliga	10		30	70
Permitterad, Arbetsbrist eller Övrigt	Fast anställda	5	15	30	
	Tidsbegränsat anställda	5	10	20	
	Företagare	10	20	40	

En person som påverkades av definitionsförändringen ett kvartal påverkas med relativt stor sannolikhet av definitionsförändringen även nästföljande kvartal. Vi har undersökt om detta gäller för de 401 personer som påverkades av definitionsförändringen under perioden februari–december 2020. Av de som svarat på AKU två kvartal på rad och som haft samma frånvaroskäl båda kvartalen, så påverkas fyra femtedelar av de som påverkades av definitionsförändringen vid första intervjun och som hade minst 13 veckors frånvaro vid andra intervjun av definitionsförändringen även vid andra intervjun. Baserat på denna information låter vi 80% av de personer som modellen valt ut som påverkade av definitionsförändringen ett kvartal och som har svarat på AKU nästa kvartal och då har minst 13 veckors frånvaro ingå i mängden av personer som påverkas av definitionsförändringen även nästa kvartal.

Tabell A.6 visar skattningar av det genomsnittliga antalet personer som påverkas av definitionsförändringen varje månad under åren 2005–2020. Modellen bygger på antagandet att sannolikheten att påverkas av definitionsförändringen är konstant över tid givet en viss kombination av frånvaroskäl, frånvarotidens längd och anknytningsgrad till arbetsmarknaden. Vi kan tyvärr inte verifiera att detta antagande är korrekt, men noterar att skattningarna som fås med denna metod är betydligt mindre under början av länkingsperioden än de skattningar som

erhålls genom att skatta definitionsförändringen under 2020 och sedan vikta denna brottskattning med populationens storlek. Den metod som vi använder här bygger inte bara på svarsdata från 2020 utan på svarsdata från hela länkningsperioden. Vi ser i tabell A.6 att det finns en väsentlig coroneffekt i antalet personer som påverkas av definitionsförändringen eftersom skattningen för 2020 ligger ungefär 10 tusen personer högre än övriga år under länkningsperioden.

Tabell A.6. Skattning av genomsnittligt antal personer per månad som påverkas av definitionsförändringen under perioden 2005–2020, uppdelat efter år, mätt i tusental.

År	2005	2006	2007	2008	2009	2010	2011	2012
Antal	17.6	15.3	14.4	13.9	15.6	16.4	14.0	13.7
År	2013	2014	2015	2016	2017	2018	2019	2020
Antal	14.9	14.8	15.8	17.8	17.5	18.1	18.6	28.2

A.2. Klassificering i nya blanketten av de som påverkas av definitionsförändringen

Som utgångspunkt i denna analys använder vi svarsdata från 2021. Totalt 384 personer svarade på ett sådant sätt på AKU 2021 att de på grund av definitionsförändringen hade fått en annan arbetskraftsstatus i den gamla blanketten än vad de fick i den nya blanketten. Med hjälp av denna datamängd kan vi analysera vilken gruppstillhörighet personer som påverkas av definitionsförändringen får i den nya blanketten.

Ungefär en tiondel av de som påverkas av definitionsförändringen klassificeras som latent arbetssökande. Eftersom vi har en separat modell för imputering av latent arbetssökande (se bilaga C), så väljer vi att försumma denna grupp här och inte imputera någon av de långtidsfrånvarande som latent arbetssökande.

Vi undersöker först vilka av de som påverkas av definitionsförändringen som klassas som arbetslösa och därefter hur resterande personer fördelas mellan de fyra undergrupperna till ej i arbetskraften. Precis som i avsnitt A.1, så fokuserar vi här på de personer som har angivit något av de fem frånvaroskälerna "Ledig för studier utan lön", "Annan tjänstledighet", "Permitterad", "Arbetsbrist" eller "Övrigt".

Av de 318 personer som angivit något av de fem frånvaroskälerna i föregående avsnitt, så är 64 personer arbetslösa. Om vi undersöker hur benägenheten att vara arbetslös beror på ålder, så ser vi ett signifikant samband. Framför allt utmärker sig den äldsta åldersgruppen där endast 3% av de som påverkas av definitionsförändringen är arbetslösa. Vi antar därför att arbetslösa långtidsfrånvarande bara finns i åldrarna 16–64 år (och försummar därmed även åldersgruppen 15 år).

Vi undersöker därefter sambandet mellan frånvarotidens längd och benägenheten att vara arbetslös. Bland personerna med 1–12 veckors frånvaro är 12% arbetslösa, men bland personerna med minst 13 veckors frånvaro är 28% arbetslösa (se tabell A.7). Vi undersöker dessa båda grupper var för sig och finner att för personer med kortare frånvarotid är benägenheten att vara arbetslös oberoende av frånvaroskäl. I modellen låter vi därför 10% av personerna med kortare frånvarotid vara arbetslösa.

För personer med minst 13 veckors frånvaro är andelen arbetslösa beroende av både frånvaroskäl och ålder men beroendet på frånvaroskäl är starkare. Vi använder här modellen att 15% av de med frånvaroskäl "Studier utan lön" eller "Annan tjänstledighet" är arbetslösa, att 35% av de med frånvaroskäl "Övrigt" är arbetslösa och att 60% av de med frånvaroskäl "Permitterad" eller "Arbetsbrist" är arbetslösa (se tabell A.7).

Tabell A.7. Antal personer med frånvaroskäl "Studier utan lön", "Annan tjänstledighet", "Permitterad", "Arbetsbrist" eller "Övrigt" under 2021 som påverkas av definitionsförändringen och som är arbetslösa respektive ej i arbetskraften uppdelat på frånvarotidens längd och frånvaroskäl.

Frånvarotidens längd	Frånvaroskäl	Arbetslös	Ej i arbetskraften
1–12 veckor	Samtliga fem nedanstående	11 (12%)	83 (88%)
Minst 13 veckor	Studier utan lön	15 (17%)	72 (83%)
	Annan tjänstledighet	7 (18%)	33 (82%)
	Permitterad	6 (60%)	4 (40%)
	Arbetsbrist	14 (58%)	10 (42%)
	Övrigt	9 (35%)	17 (65%)

Vi övergår nu till att analysera de kvarvarande 254 personerna som alla tillhör någon av de fyra undergrupperna till ej i arbetskraften. Tabell A.8 visar hur fördelningen på de fyra undergrupperna till ej i arbetskraften beror på ålder. Vi noterar att endast 2% tillhör undergruppen sjuka och försummar därför denna grupp i modellen. Eftersom frågor om studier ställs oavsett arbetskraftsstatus, så kan vi från svaren på dessa frågor avgöra vilka personer som tillhör undergruppen heltidsstuderande. Samtliga personer i åldrarna 15–59 år som inte tillhör undergruppen heltidsstuderande placeras i modellen i undergruppen övriga.

Det kvarstår att undersöka hur personer i åldrarna 60–74 år ska fördelas mellan undergrupperna pensionärer och övriga. Som synes i tabell A.9 beror denna fördelning på frånvaroskäl. Baserat på data från de 52 personerna i åldrarna 60–74 år, så formulerar vi modellen enligt följande. I åldersgruppen 60–64 år låter vi 75% av personerna med

frånvaroskäl "Övrigt" vara pensionärer och resterande personer förs till undergruppen övriga. Samtliga personer med något av de andra fyra frånvaroskälen förs till undergruppen övriga. I åldersgruppen 65–74 år låter vi 80% av samtliga personer vara pensionärer och resterande personer förs till undergruppen övriga, oberoende av frånvaroskäl.

Tabell A.8. Antal personer med frånvaroskäl "Studier utan lön", "Annan tjänstledighet", "Permitterad", "Arbetsbrist" eller "Övrigt" under 2021 som påverkas av definitionsförändringen och som tillhör olika undergrupper till ej i arbetskraften uppdelat på åldersgrupper.

Ålder	Heltidsstuderande	Pensionärer	Sjuka	Övriga
16–19 år	5 (83%)	0 (0%)	0 (0%)	1 (16%)
20–24 år	23 (96%)	0 (0%)	0 (0%)	1 (4%)
25–34 år	72 (90%)	0 (0%)	0 (0%)	8 (10%)
35–44 år	34 (74%)	0 (0%)	0 (0%)	12 (26%)
45–54 år	14 (44%)	0 (0%)	3 (9%)	15 (47%)
55–59 år	1 (8%)	1 (8%)	1 (8%)	9 (75%)
60–64 år	0 (0%)	8 (44%)	0 (0%)	10 (56%)
65–74 år	0 (0%)	27 (75%)	2 (6%)	7 (19%)

Tabell A.9. Antal personer i åldersgrupperna 60–64 år och 65–74 år som påverkas av definitionsförändringen under 2021 och som tillhör någon av undergrupperna pensionärer respektive övriga till ej i arbetskraften, uppdelat efter frånvaroskäl.

Ålder	Frånvaroskäl	Pensionärer	Övriga
60-64 år	Studier utan lön, annan tjänstledighet, permitterad eller arbetsbrist	0 (0%)	7 (100%)
	Övrigt	8 (73%)	3 (27%)
65-74 år	Studier utan lön, annan tjänstledighet, permitterad eller arbetsbrist	10 (71%)	4 (29%)
	Övrigt	17 (85%)	3 (15%)

A.3. Imputeringsmodell för definitionsförändringen

I avsnitt A.1 beskrev vi kortfattat hur urvalet av personer som påverkas av definitionsförändringen har utförts baserat på procentsatserna i tabell A.5. Varje månad delar vi in mängden frånvarande personer i 16 olika imputeringsgrupper (de 15 grupperna med nollskilda procentsatser i tabell A.5 och en grupp med resterande personer). Låt N beteckna antalet personer i en av dessa imputeringsgrupper och låt p beteckna procentsats i tabell A.5 som hör till imputeringsgruppen. Vi sorterar först ut de personer i denna grupp som har minst 13 veckors frånvaro, som besvarade AKU föregående kvartal och som då angav samma frånvaroskäl som innevarande månad och betecknar denna mängd personer med H . Låt nu M beteckna 80 % av antalet personer i H ,

avrundat till närmaste heltal. Vi väljer slumpmässigt ut M personer ur H och tar sedan bort resterande personer i H från imputeringsgruppen. Från resten av imputeringsgruppen väljer vi sedan slumpmässigt ut $\max(Np - M, 0)$ personer.

Svarsdata och vikter för de personer som valts ut från någon av imputeringsgrupperna används för att, för varje månad, ta fram skattningar av alla variabler som påverkas i negativ riktning av definitionsförändringen.

Med hjälp av modellen som beskrivs i avsnitt A.2 kan vi sedan avgöra om personer som påverkas av definitionsförändringen är arbetslösa eller ej i arbetskraften. I det förstnämnda fallet kan vi avgöra om personerna är heltidsstuderande arbetslösa eller ej via deras svar på frågor om studier. I det sistnämnda fallet kan vi använda modellen som beskrivs i avsnitt A.2 för att avgöra vilken undergrupp till ej i arbetskraften som personerna tillhör. Med hjälp av dessa modeller och personernas vikter kan vi, för varje månad, ta fram skattningar av alla variabler som påverkas i positiv riktning av definitionsförändringen.

Ovanstående metodik har använts för att ta fram skattningar av definitionsförändringen. Ursprungstanken med att ta fram en imputeringsmodell var att klassificera om alla personer som enligt modellen påverkas av definitionsförändringen och ge dessa personer den arbetskraftsstatus som de hade haft i den nya blanketten. Detta omöjliggörs dock av att personer som klassificeras som arbetslösa får svara på ett antal frågor om sitt arbetssökande som de av naturliga skäl inte har fått besvara i den gamla blanketten eftersom de tidigare klassificerades som sysselsatta. Vi har därför valt att bara imputera värden för de personer som påverkas av definitionsförändringen och som enligt modellen klassificeras som ej i arbetskraften. I genomsnitt utgör dessa personer ungefär 80% av det totala antalet personer som påverkas av definitionsförändringen.

B. Modell för skattning av populationsförändring

Vid införandet av den nya ramlagen i AKU vid årsskiftet 2020/2021 togs personer som inte bor i privata hushåll bort från målpopulationen i AKU. Som nämndes i avsnitt 4.4, så tillhör personer som inte bor i privata hushåll i allmänhet en av följande tre huvudgrupper: (a) personer som bor i särskilda boenden, såsom äldreboenden, servicehus eller gruppboendestäder, (b) värnpliktiga eller kontraktsanställda soldater eller (c) fängelseinterner. Eftersom övriga grupper är avsevärt mindre till antalet än de tre huvudgrupperna och vi dessutom saknar registeruppgifter för dem, så har de övriga grupperna försumats här.

Antalet personer i de tre huvudgrupperna kan till viss del skattas med hjälp av befintliga register. Vi beskriver i denna bilaga hur uppgifter från register och modeller har kombinerats för att för varje månad under länkingsperioden 2005–2020 och för var och en av de makrolänkade tidsserierna i etapp 1 och 2 skatta antalet personer i var och en av de tre huvudgrupperna ovan.

Vid modelleringen har vi utgått från följande. Publicerade AKU-data över antalet personer i befolkningen uppdelat på kön, ålder och födelseland baseras på register och har därför hög precision. För att behålla denna höga precision i länkningen på makronivå är det önskvärt att ta fram detaljerade skattningar av antalet personer som inte bor i privata hushåll uppdelat på kön, ålder och födelseland. Publicerade AKU-data över antalet personer i olika undergrupper till befolkningen (exempelvis sysselsatt, frånvarande m.m.) är däremot urvalsskattningar som är behäftade med osäkerhet, varför det inte är nödvändigt med en lika detaljerad modell för undergrupperna av befolkningen.

I avsnitt B.1 beskriver vi skattningar av antalet personer i särskilda boenden, i avsnitt B.2 skattningar av antalet värnpliktiga och i avsnitt B.3. skattningar av antalet fängelseinterner.

B.1. Särskilda boenden

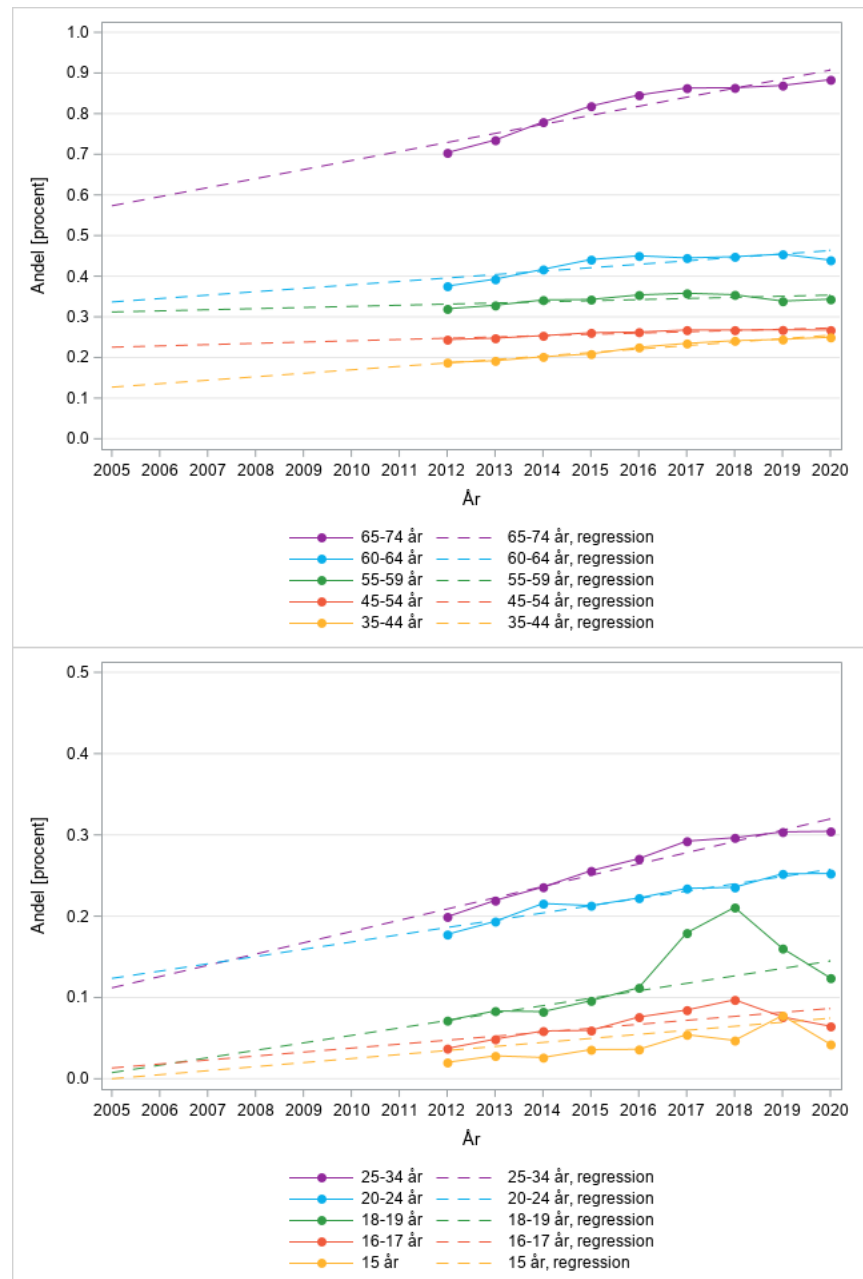
Från och med 2012 innehåller Fastighetsregistret adressuppgifter för särskilda boenden. Genom att samköra dessa adressuppgifter med folkbokföringsadressen i RTB, så kan vi sortera ut alla personer som var folkbokförda på ett särskilt boende den 31/12 ett visst år. Med hjälp av informationen i RTB kan vi också bryta upp antalet personer som är folkbokförda på ett särskilt boende efter kön, ålder och födelseland.

B.1.1 Skattning av totaler och åldersfördelning

Registerinformationen från 2012–2020 kan vi sedan använda för att modellera antalet personer som var folkbokförda på ett särskilt boende under åren 2005–2011. Figur B.1 visar andelen av RTB-befolkningen som var folkbokförda på ett särskilt boende under åren 2012–2020 uppbrutet på den åldersindelning som används i AKU. Som synes

uppvisar andelarna nästintill linjära ökningar över tid, men med olika lutningar för olika åldersgrupper.

Figur B.1. Andel av befolkningen i RTB som var folkbokförda på ett särskilt boende under perioden 2012–2020 uppbrutet på åldersgrupper i spannet 35–74 år (övre bilden) och 15–34 år (undre bilden). För hela perioden 2005–2020 redovisas linjära regressionsmodeller för andelen personer som är folkbokförda på ett särskilt boende.



För varje åldersgrupp har vi anpassat en linjär regressionsmodell med årtal som kovariat för andelen personer som var folkbokförda på ett särskilt boende. För två av åldersgrupperna har vi gjort särskilda anpassningar av modellerna. För åldersgruppen 15 år har vi använt en

regressionsmodell utan konstantterm för att undvika negativa värden på den skattade andelen 15-åringar på särskilda boenden i början av länkningsperioden. För åldersgruppen 18–19 år har åren 2017–2018 uteslutits från dataunderlaget till regressionsmodellen eftersom dessa år uppvisar ett avvikande mönster, förmodligen på grund av en hög andel ensamkommande flyktingar i denna åldersgrupp (se figur B.1). Figur B.1 redovisar dessa regressionsanpassningar för hela perioden 2005–2020. Antalet personer som är folkbokförda på särskilda boenden kan antas vara relativt konstant över ett enskilt år, varför det är rimligt att använda årsskattningen som månadsskattning för alla månader under året.

B.1.2 Skattning av könsfördelning

För att kunna inkludera kön i modellen för åren 2005–2011 har vi undersökt hur de som var folkbokförda på särskilda boenden åren 2012–2020 fördelas efter kön. Denna fördelning återfinns i tabell B.1. Precis som i regressionsmodellen, så har vi här uteslutit åren 2017–2018 för åldersgruppen 18–19 år eftersom dessa år avviker kraftigt i figur B.1.

Tabell B.1. Könsfördelning för personer som var folkbokförda på särskilt boende under åren 2012–2020, uppdelat på åldersgrupper.

Åldersgrupp	15 år	16–17 år	18–19 år	20–24 år	25–34 år
Andel män	50 %	63 %	61 %	57 %	58 %
Åldersgrupp	35–44 år	45–54 år	55–59 år	60–64 år	65–74 år
Andel män	57 %	55 %	57 %	56 %	50 %

Tabell B.1 visar att könsfördelningen är helt jämn i den yngsta och den äldsta åldersgruppen, att andelen män är markant högre (61–63 %) i åldersgrupperna i spannet 16–19 år och att andelen män är något högre (55–58 %) i alla andra åldersgrupper. I modellen för åren 2005–2011 använder vi faktorer för överrepresentation av män respektive kvinnor. För män använder vi faktorn 1.23 för åldersgrupperna i spannet 16–19 år och faktorn 1.13 för åldersgrupperna i spannet 20–64 år. För kvinnor använder vi faktorerna 0.77 för åldersgrupperna i spannet 16–19 år och 0.87 för åldersgrupperna i spannet 20–64 år. För den yngsta och den äldsta åldersgruppen används faktorn 1.00 för såväl män som kvinnor.

B.1.3 Skattning av fördelning efter födelseland

Vi undersöker härnäst födelseland och redovisar i tabell B.2 andelen inrikes födda bland personer som var folkbokförda på ett särskilt boende åren 2012–2015 och som jämförelse andelen inrikes födda i RTB samma år. Anledningen till att vi här bara har använt data från en delmängd av tidsintervallet 2012–2020 är att överrepresentationen av utrikes födda ökar kraftigt under perioden. Det är därför troligt att åren 2012–2015 är mer representativa för åren 2005–2011 än vad åren 2012–2020 är. Som synes är utrikes födda kraftigt överrepresenterade i de tre yngsta

åldersgrupperna, något överrepresenterade i åldersgruppen 20–24 år, men något underrepresenterade i alla andra åldersgrupper utom den äldsta.

De båda kolumnerna längst till höger i tabell B.2 visar en faktor f_i för överrepresentation av inrikes födda på särskilda boenden och motsvarande faktor f_u för utrikes födda. Dessa båda faktorer har beräknats enligt följande. Låt N beteckna storleken på befolkningen, b andelen av befolkningen som är folkbokförd på ett särskilt boende och c_i andelen inrikes födda i befolkningen. Då gäller att andelen inrikes födda på särskilda boenden, här betecknad b_i , ges av

$$b_i = \frac{\text{Antal inrikes födda på särskilda boenden}}{\text{Antal totalt på särskilda boenden}} = \frac{Nbc_i f_i}{Nb} = c_i f_i.$$

Notera att vi kan bestämma f_i genom att avläsa b_i och c_i från de två kolumnerna längst till vänster i tabell B.2. Vidare vet vi att andelen utrikes födda på särskilda boenden, b_u , uppfyller de båda ekvationerna $b_u = (1 - c_i)f_u$ och $b_i + b_u = 1$, vilka kan kombineras till

$$f_u = \frac{1 - b_i}{1 - c_i}.$$

Tabell B.2. Andel inrikes födda bland personer som var folkbokförda på särskilt boende under åren 2012–2016 (första kolumnen) och andelen inrikes födda i RTB samma period (andra kolumnen). Faktor för överrepresentation av andel folkbokförda på särskilda boende för inrikes födda respektive utrikes födda (tredje respektive fjärde kolumnen). Alla variabler är uppdelade efter åldersgrupper.

Ålder	Andel inrikes födda på särskilt boende	Andel inrikes födda i befolkningen	Faktor för överrepresentation av inrikes födda	Faktor för överrepresentation av utrikes födda
15 år	50 %	87 %	0.57	3.96
16–17 år	33 %	87 %	0.39	4.93
18–19 år	57 %	87 %	0.66	3.28
20–24 år	83 %	86 %	0.96	1.22
25–34 år	80 %	75 %	1.07	0.79
35–44 år	83 %	77 %	1.08	0.74
45–54 år	86 %	80 %	1.07	0.73
55–59 år	87 %	82 %	1.06	0.75
60–64 år	87 %	85 %	1.03	0.81
65–74 år	87 %	87 %	1.00	1.03

I modellen för åren 2005–2011 använder vi för utrikes födda faktorerna 4, 5 och 3 för de tre yngsta åldersgrupperna, 1.22 för åldersgruppen 20–24 år, 0.77 för åldersgrupperna i spannet 25–64 år och 1.00 för den äldsta åldersgruppen. På motsvarande sätt använder vi för inrikes födda

faktorerna 0.56, 0.38 och 0.70 för de tre yngsta åldersgrupperna, 0.96 för åldersgruppen 20–24 år, 1.06 för åldersgrupperna i spannet 25–64 år och 1.00 för den äldsta åldersgruppen.

Skattningen av antalet personer som är folkbokförda på särskilda boenden under något av åren 2005–2011 för en given uppdelning efter kön, ålder och födelseland utgår från antalet personer i RTB det året för samma uppdelning efter kön, ålder och födelseland. Vi multiplicerar detta antal med värdet på den skattade regressionslinjen i den åldersgruppen det året och multiplicerar sedan med de båda faktorerna för överrepresentation för kön respektive födelseland som hör till den givna uppdelningen. För att säkerställa summakonsistens, så tar vi sedan fram skattningar på totalnivå genom att aggregera ihop skattningar uppbrutna på kön, ålder och födelseland.

B.1.4 Skattning av undergrupper till befolkningen

Med hjälp av förfarandet som beskrivits ovan kan vi skatta det totala antalet personer som var folkbokförda på särskilda boenden en viss månad under länkningsperioden och som därför ska tas bort från målpopulationen. Detta är dock inte tillräckligt för att genomföra länkningen på makronivå av AKU. Vi behöver också veta vilka undergrupper till befolkningen som personer som är folkbokförda på särskilda boenden skulle ha klassificerats till att tillhöra i AKU för att kunna ta bort dem också från dessa undergrupper.

För att kunna fördela ut antalet personer folkbokförda på särskilda boenden på olika undergrupper till befolkningen, så har vi samkört registerdata över dessa personer under åren 2012–2020 med svarsdata från AKU under samma tidsperiod. Vi har sedan använt den viktade fördelningen av dessa svar som grund för skattningar av fördelningen på undergrupper till befolkningen för hela länkningsperioden. Det är möjligt att bortfallet är systematiskt i denna grupp, så att exempelvis personer som står längre från arbetsmarknaden är mindre benägna att svara på AKU och att detta kan påverka den skattade fördelningen på undergrupper till befolkningen. Eftersom vi inte kan avgöra om ett sådant systematiskt bortfall föreligger och populationsförändringen är relativt liten i absoluta tal, så väljer vi att bortse från denna potentiella skevhet.

Tabell B.3–B.5 visar hur personer folkbokförda på särskilda boenden fördelas på olika undergrupper till befolkningen. Eftersom data saknas för 15-åringar, antar vi att alla i den åldersgruppen tillhör gruppen ej i arbetskraften heltidsstuderande och att 20 % av personerna är latent arbetssökande. Variablerna i tabell B.5 länkas inte med avseende på åldersgrupper varför tabell B.5 bara innehåller skattningar för hela åldersspannet 15–74 år.

I länknings av populationen utgår vi från tabell B.3–B.5, men vissa små grupper försummas och skattningarna för andra grupper avrundas. Till att börja med försummar vi alla skattningar för företagare samt skattningarna för undersysselsatta, sysselsatta utomlands, statligt anställda och landstingskommunalt anställda i alla åldersgrupper.

Tabell B.3. För personer som både medverkat i AKU och som varit folkbokförda på ett särskilt boende under samma år under tidsperioden 2012–2020 redovisas andelarna, mätta i procent, som tillhör olika undergrupper till befolkningen, uppdelat på åldersgrupper. För åldersgruppen 15 år saknas data. Redovisade undergrupper är i arbete (Iarb), frånvarande (FV), fast anställda (FA), tidsbegränsat anställda (TA), företagare (Ftg), undersysselsatta (USyss), arbetslösa (Alos) respektive arbetslösa heltidsstuderande (AlSt). Notera att Iarb och FV summerar till sysselsatta, att FA och TA summerar till anställda, att FA, TA och Ftg summerar till sysselsatta samt att FA, TA, Ftg och Alos summerar till arbetskraften.

Ålder	Iarb	FV	FA	TA	Ftg	USyss	Alos	AlSt
16–17 år	11	4	4	13	0	0	17	9
18–19 år	6	5	3	7	0	0	24	18
20–24 år	26	5	14	13	3	2	9	4
25–34 år	26	4	18	10	3	3	1	1
35–44 år	18	2	15	4	1	3	4	3
45–54 år	14	3	14	1	1	1	3	1
55–59 år	11	4	13	2	0	2	2	0
60–64 år	15	2	14	1	2	0	2	0
65–74 år	5	1	3	3	1	0	0	0

Tabell B.4. För personer som både medverkat i AKU och som varit folkbokförda på ett särskilt boende under samma år under tidsperioden 2012–2020 redovisas andelarna, mätta i procent, som tillhör olika undergrupper till befolkningen, uppdelat på åldersgrupper. För åldersgruppen 15 år saknas data. Redovisade undergrupper är latent arbetssökande (Lat), ej i arbetskraften heltidsstuderande (EjSt), ej i arbetskraften pensionärer (EjPe), ej i arbetskraften sjuka (EjSj) respektive ej i arbetskraften övriga (EjÖv). Notera att EjSt, EjPe, EjSj och EjÖv summerar till ej i arbetskraften.

Ålder	Lat	EjSt	EjPe	EjSj	EjÖv
16–17 år	14	65	0	0	1
18–19 år	10	55	0	0	12
20–24 år	1	15	0	40	6
25–34 år	1	5	0	51	12
35–44 år	1	4	0	57	13
45–54 år	1	0	1	73	7
55–59 år	1	0	1	72	10
60–64 år	1	0	3	68	10
65–74 år	0	0	71	19	4

Tabell B.5. För personer som både medverkat i AKU och som varit folkbokförda på ett särskilt boende under samma år under tidsperioden 2012–2020 redovisas andelarna, mätta i procent, som tillhör olika undergrupper till befolkningen. Redovisade undergrupper är sysselsatta utomlands (SyUtr), privat anställda i Sverige (PrivSv), statligt anställda i Sverige (StatSv), kommunalt anställda i Sverige (KomSv), företagare och medhjälpande hushållsmedlemmar i Sverige (FtgSv), företagare i aktieföretag i Sverige (FtgAB), primärkommunalt anställda i Sverige (PriSv) respektive landstingskommunalt anställda i Sverige (LanSv). Notera att SyUtr, PrivSv, StatSv, KomSv och FtgSv summerar till sysselsatta samt att PriSv och LanSv summerar till kommunalt anställda i Sverige.

Ålder	SyUtr	PrivSv	StatSv	KomSv	FtgSv	FtgAB	PriSv	LanSv
15–74 år	0	10	1	4	2	1	4	0

Modellerna för åldersgrupperna i spannet 16–24 år redovisas i tabell B.6. Alla kategorier som saknas i tabellen försummas.

Tabell B.6. Modellfördelning för personer i åldersgrupperna i spannet 16–24 år som är folkbokförda på ett särskilt boende, redovisad i procent.

Ålder	Iarb	FV	FA	TA	Alos	AISt	EjSt	EjSj	EjÖv	Lat
16–17 år	10	5	5	10	15	10	65	0	5	15
18–19 år	10	5	5	10	20	15	55	0	10	10
20–24 år	25	5	15	15	10	5	15	40	5	0

Modellerna för åldersgrupperna i spannet 25–74 år redovisas i tabell B.7. Alla kategorier som inte finns med i tabellen, såsom arbetslösa, försummas för åldersgrupperna. För variablerna i tabell B.5 som bara länkas utan uppdelning på åldersgrupper, låter vi i modellen två tredjedelar av de sysselsatta vara privat anställda och resterande tredjedel vara primärkommunalt anställda och därmed även kommunalt anställda.

Tabell B.7. Modellfördelning för personer i åldersgrupper i spannet 25–74 år som är folkbokförda på ett särskilt boende, redovisad i procent.

Ålder	Iarb	FV	FA	TA	EjSt	EjSj	EjPe	EjÖv
25–34 år	25	5	20	10	5	50	0	15
35–44 år	15	5	15	5	5	60	0	15
45–54 år	10	5	15	0	0	75	0	10
55–59 år	10	5	15	0	0	75	0	10
60–64 år	10	5	15	0	0	70	5	10
65–74 år	5	0	5	0	0	20	70	5

B.2 Värnpliktiga

I Inkomst- och taxeringsregistret (IoT) finns under perioden 2003–2010 variabeln IDAGERS, som för varje person och år redovisar det totala

utbetalade beloppet i dagersättning för värnpliktiga. Under perioden juli 2011 till juni 2018 ersattes värnplikten av grundläggande militär utbildning (GMU) och under perioden 2011–2018 finns därför i IoT i stället variabeln IGMUERS, som för varje person och år redovisar det utbetalade beloppet i dagersättning för GMU. SCB har bestämt att personer som genomgår GMU (och även kontraktsanställda soldater) ska likställas med värnpliktiga i AKU och därför tillhör ingen av dessa grupper längre målpopulationen. I juli 2018 återupptogs värnplikten och från och med den månaden finns även ett månatligt register över personer som erhållit dagersättning för värnpliktiga.

B.2.1 Skattning av totaler

Tabell B.8 visar antalet personer under 45 år som uppfyller något av villkoren $IDAGERS > 0$ eller $IGMUERS > 0$ under åren 2003–2018 samt antalet personer under 45 år som har fått ersättning någon av årets månader under åren 2019–2021. Åldersgränsen på 45 år sorterar bara bort enstaka personer, men tillämpas eftersom den övre åldersgränsen för värnplikt är 45 år. Genom att samköra registren för två på varandra följande år har vi även kunnat ta fram antalet personer som fått dagersättning ett år utan att ha fått det året före. Antalet sådana nya personer med dagersättning under åren 2003–2010 och 2018–2021 stämmer väl överens med Försvarmaktens egna siffror över antal inryckta⁵.

För perioden fram till och med 2010 noteras att ungefär 80 % av de personer som var nyinskrivna ett år finns kvar i registret året därpå. Detta är konsistent med att värnpliktiga rycker in i början av hösten och påbörjar en utbildning som pågår en bit in på kommande år. För åren då värnplikten var avskaffad är mönstret ett annat. Under dessa år var endast 400–700 personer årligen kvar i registret till nästkommande år. Detta är konsistent med att GMU var en tre månader lång utbildning med flera intag per år. De flesta personer som påbörjade utbildningen under ett givet år färdigställde därför utbildningen under samma år och fick därför bara dagersättning under ett kalenderår. Från och med 2016 är det återigen en andel på runt 80 % av de som är nyinskrivna ett år som får dagersättning även nästkommande år, vilket är konsistent med en återgång till längre utbildningstider.

För att skatta antalet värnpliktiga varje månad under perioden januari 2005–december 2010 använder vi mängden personer som fick dagersättning för första gången ett visst år. Denna mängd används till skattningarna från och med juli det året till och med juni året därpå. Antalet värnpliktiga en given månad fås genom att multiplicera det totala antalet nya personer med dagersättning med en av de tolv procentsatserna som redovisas i tabell B.9. Tabellen baseras på data

⁵ Se artikeln "Värnplikt i Sverige" på wikipedia: https://sv.wikipedia.org/wiki/Värnplikt_i_Sverige (hämtad 2023-10-18).

över säsongsmönster för månatliga utbetalningar av dagersättning under åren 2018–2021.

Tabell B.8. Antal personer under 45 år som fick dagersättning per år under perioden 2003–2021 (Dagersättning) och antalet personer under 45 år som fick dagersättning för första gången ett givet år under perioden 2003–2021 (Ny dagersättning). För perioden 2003–2018 kommer informationen från IoT och för perioden 2019–2021 från månadsregistret över dagersättning. Som jämförelse visas data från Försvarmakten över antalet nyinryckta vämpliktiga per år under åren med vämplikt (Inryckta). För åren 2018–2020 är Försvarmaktens data avrundat till hela hundratal.

År	Dager-sättning	Ny dag-ersättn.	Inryckta	År	Dager-sättning	Ny dag-ersättn.	Inryckta
2003	28068		15529	2012	2754	2062	–
2004	27653	14842	14446	2013	4724	4245	–
2005	21158	9542	9225	2014	3289	2617	–
2006	17097	10246	10129	2015	3607	3014	–
2007	11309	4851	4730	2016	3437	2727	–
2008	10283	7942	7908	2017	4931	2706	–
2009	8621	7333	7345	2018	6464	4067	3700
2010	6955	1769	1644	2019	7727	4572	4500
2011	4165	3774	–	2020	8657	4862	4900

Tabell B.9. Andel av personerna som fick dagersättning för första gången ett visst år som uppskattas vara värnpliktiga en given månad under kommande år.

Månad	Juli	Augusti	Septemb.	Oktober	November	December
Andel	40 %	70 %	100 %	100 %	90 %	90 %
Månad	Januari	Februari	Mars	April	Maj	Juni
Andel	80 %	80 %	75 %	75 %	70 %	70 %

Under perioden januari 2011–juni 2016 antas alla nya personer med dagersättning ha genomgått en tre månaders GMU med flera olika starttillfällen under året. De flesta personer förefaller också ha kompletterat GMU med en kortare vidareutbildning. Om vi jämför typvärdet på årsutbetalningarna (ca 15 000 kr) med ersättningen per dag (72 kr), så kan vi dra slutsatsen att den typiska utbildningslängden var 7 månader. De flesta kan alltså antas ha fått dagersättning under ca 8 månader (2/3 av månaderna) och det totala antalet värnpliktiga skattas därför, för årets samtliga månader, till 2/3 av det totala antalet nya personer med dagersättning.

Perioden juli 2016–juni 2018 utgör en övergångsperiod med längre utbildningstider men ingen värnplikt. Under det första av dessa båda år skattas antalet värnpliktiga som två tredjedelar av de personer som var nyinskrivna 2016 multiplicerat med lämplig faktor i tabell B.9. Under

det andra året skattas antalet värnpliktiga som antalet nya personer med dagersättning 2017 multiplicerat med lämplig faktor i tabell B.9.

B.2.2 Skattning av fördelning efter ålder, kön och födelseland

Under hela perioden januari 2005–juni 2018 antas fördelningen av antalet värnpliktiga efter ålder, kön och födelseland en given månad vara densamma som för den mängd personer som ligger till grund för skattningen av det totala antalet värnpliktiga den månaden.

För perioden juli 2018–december 2020 används data från det månatliga registret över personer med dagersättning för att skatta det totala antalet värnpliktiga. Under denna period har vi månatlig information även om kön, ålder och födelseland. Antalet värnpliktiga med olika kombinationer av kön, ålder och födelseland kan därför erhållas direkt från registret.

B.2.3 Jämförelse mellan olika skattningar av totaler

Figur B.2 visar en jämförelse mellan fyra olika metoder för att skatta det genomsnittliga antalet värnpliktiga per månad. Utöver den IoT-baserade modell som beskrivits i avsnitt B.2.1 har vi även implementerat en modell som utgår från de totala årliga utbetalningsbeloppen i IoT och storleken på ersättningen per dag. Med denna information kan vi skatta antalet dagar varje person fått ersättning och därigenom skatta det genomsnittliga antalet värnpliktiga per månad. Vi har trunkerat skattningarna av antalet ersättningsdagar, så att detta belopp inte överstiger 365. De båda IoT-baserade modellerna ger resultat som är kvalitativt lika varandra, men särskilt under senare år ger den senare metoden högre skattningar av antalet värnpliktiga. Detta beror sannolikt på att totalbeloppen i IoT inkluderar även andra ersättningar än dagersättning, såsom ersättning för material, resor etc., vilket överdriver skattningarna för den andra IoT-baserade metoden något.

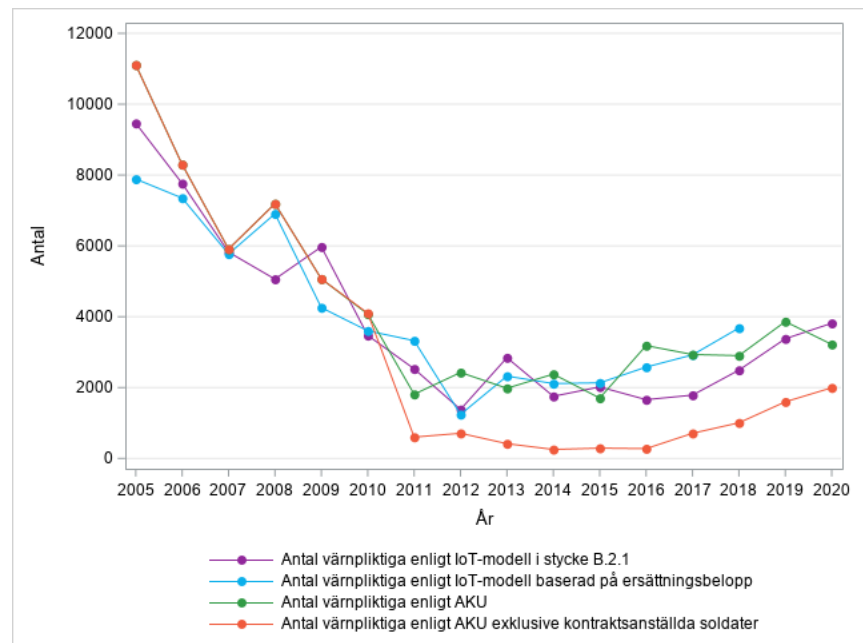
Figur B.2 visar även viktade skattningar av antalet värnpliktiga enligt AKU. Figuren visar både totalskattningar och skattningar som exkluderar kontraktsanställda soldater. Skattningarna av antalet kontraktsanställda soldater är mer volatila än skattningarna av övriga värnpliktiga. Vi ser att skattningarna för den första IoT-modellen och AKU-skattningen av de som inte är kontraktsanställda soldater följer samma mönster men med en konsekvent skillnad på cirka 2 000 personer från och med 2011. Generellt är överensstämmelsen mellan den första IoT-modellen och AKU-skattningen av det totala antalet värnpliktiga god.

B.2.4 Skattning av undergrupper till befolkningen

Precis som för de som är folkbokförda i särskilda boenden, så behöver vi skatta hur många av de värnpliktiga som skulle ha klassificerats till att tillhöra olika undergrupper i AKU. Vi använder här AKU-data för att

härleda en enkel modell för vilka undergrupper som de värnpliktiga har tillhört.

Figur B.2. Genomsnittligt antal värnpliktiga under perioden 2005–2020 enligt AKU (med respektive utan personer som klassats som kontraktsanställda soldater) och de två IoT-baserade modeller som beskrivits i avsnitt B.2.1 respektive B.2.3.



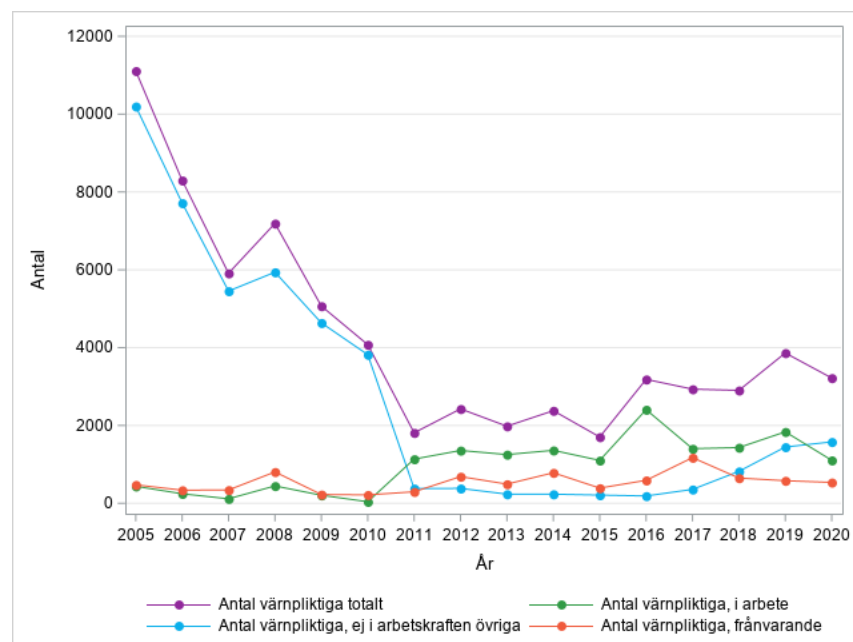
Figur B.3 visar hur många av de värnpliktiga i AKU som kategoriserats i någon av följande tre undergrupper: i arbete, frånvarande respektive ej i arbetskraften övriga. Som synes är kategorin ej i arbetskraften övriga, klart dominant (ca 90 %) under de inledande åren med värnplikt (2005–2010). Därefter tar kategorierna i arbete respektive frånvarande över nästan helt under åren utan värnplikt (2011–2017). Från och med 2018 ökar successivt antalet personer i kategorin ej i arbetskraften övriga.

I tabell B.10 har vi analyserat data mer utförligt efter åldersgrupper. Ett tydligt mönster är att antalet frånvarande och antalet fast anställda ligger nära varandra. Detsamma gäller för antalet i arbete och antalet tidsbegränsat anställda. Vi har också undersökt om det finns ett säsongsmönster i fördelningen på undergrupper. Under åren 2011–2020 är andelen frånvarande något högre i december och under sommar-månaderna, särskilt i juli. I modellen har vi därför låtit skattningarna för i arbete och frånvarande byta plats under juli månad.

Vi har också undersökt de andra undergrupper till befolkningen som vi har länkat på makronivå. De tre undergrupper till ej i arbetskraften som saknas i tabell B.10 samt arbetslösa är helt försumbara och detsamma gäller latent arbetssökande samt företagare. Antalet undersysselsatta är helt försumbart under åren 2005–2010 och uppgår till ca 5 % av de sysselsatta värnpliktiga under åren 2011–2020. De undersysselsatta

värnpliktiga är i absoluta tal dock inte fler än ca 150 personer per månad och har därför försumrats i modellen.

Figur B.3. Genomsnittligt antal värnpliktiga i undergrupperna i arbete, frånvarande respektive ej i arbetskraften enligt AKU under perioden 2005–2020.



Tabell B.10. Genomsnittligt antal värnpliktiga i undergrupper till befolkningen per månad enligt AKU under olika delar av perioden 2005–2020, uppdelat i åldersgrupper.

Period	Ålder	Ej i AK, övriga	Fast anställda	Tidsbegr. anställda	I arbete	Frånvarande
2005–2010	18–19 år	3860	200	140	130	230
	20–24 år	2400	200	80	120	160
	25–34 år	20	10	0	0	10
2011–2017	18–19 år	120	10	90	90	10
	20–24 år	130	160	730	670	230
	25–34 år	20	220	640	540	330
2018–2020	18–19 år	770	50	480	460	70
	20–24 år	420	130	460	450	140
	25–34 år	40	190	530	370	340

Tabell B.11 visar sektortillhörigheten för anställda värnpliktiga under olika delar av länkningsperioden. Som synes var i stort sett samtliga dessa personer privat anställda under åren 2005–2010, medan ca 10 %

var privat anställda under åren 2011–2020 och resten då var statligt anställda. De kommunalt anställda var försumbara under hela perioden.

Tabell B.11. Genomsnittligt antal vämpliktiga i olika sektorer enligt AKU under olika delar av perioden 2005–2020.

Period	Privat anställda	Statligt anställda	Primärkom. anställda	Landst.kom. anställda
2005–2010	570	10	50	20
2011–2017	170	1780	70	20
2018–2020	190	1770	30	50

I modellen för undergrupper har vi utgått från tabell B.10–B.11 ovan, försummat vissa skattningar och avrundat andra skattningar. Eftersom antalet värnpliktiga uppdelat på åldersgrupper varierar ganska mycket över året (främst på grund av värnpliktiga som fyller 20 år och byter åldersgrupp medan de gör sin militärtjänstgöring), så har vi omvandlat de absoluta talen i tabell B.10–B.11 till relativa tal.

Den resulterande modellen för åldersgrupperna i spannet 18–34 år sammanfattas i tabell B.12. För åldersgrupper med personer som fyllt 35 år antar vi att 70 % är tidsbegränsat anställda / i arbete och att 30 % är fast anställda / frånvarande. Notera att modellen är oberoende av kön och födelseland och att den endast beror på månad genom att skattningarna för i arbete och frånvarande byter plats under juli månad under åren 2011–2020.

Tabell B.12. Modellfördelning för värnpliktiga i åldersgrupperna i spannet 18–34 år, redovisad i procent.

Period	Ålder	Ej i arbets-kraften övriga	Frånvaran-de / Fast anställda	I arbete / Tidsbegr. anställda	Privat anställda	Statligt anställda
2005–2010	18–19 år	92	5	3	Samtliga anställda	–
	20–24 år	92	5	3		
	25–34 år	70	30	0		
2011–Juni 2018	18–19 år	50	0	50	10 % av de anställda	90 % av de anställda
	20–24 år	10	20	70		
	25–34 år	0	30	70		
Juli 2018–2020	18–19 år	60	5	35	10 % av de anställda	90 % av de anställda
	20–24 år	40	15	45		
	25–34 år	5	30	65		

B.3. Fängelseinterner

SCB har från Kriminalvården erhållit data över antalet pågående fängelseverkställigheter, vilket i stort motsvarar antalet intagna på

anstalter i Sverige. Leveransen innehåller data för den första oktober varje år under perioden 2006–2017 samt för den första i månaden varje månad från och med januari 2018.

Antalet pågående fängelseverkställigheter redovisas uppdelat efter kön korsat med den åldersindelning som används i AKU. Antalet finns också redovisat efter medborgarskapsland samt efter den regionala indelning som används i hjälpinformationen i AKU. Tabell B.13 ger ett exempel på ålders- och könsfördelningen för antalet intagna under några av åren under länkingsperioden.

Från den månatliga data som finns tillgänglig för åren 2018–2020, ser vi att det inte finns något tydligt säsongsmönster för antalet intagna. Vi använder därför under åren 2006–2017 uppgifterna från första oktober för årets samtliga månader. Dessutom använder vi också 2006 års data även för samtliga månader 2005, eftersom data från Kriminalvården saknas för 2005.

Tabell B.13. Antal pågående fängelseverkställigheter den första oktober åren 2006, 2010, 2015 och 2020, uppdelat på åldersgrupper och kön. Åldersgruppen 15 år har utelämnats eftersom intagna saknades i den åldersgruppen.

Ålder	2006		2010		2015		2020	
	Män	Kv.	Män	Kv.	Män	Kv.	Män	Kv.
16–19 år	75	2	58	1	47	1	66	2
20–24 år	712	23	708	27	544	14	899	30
25–34 år	1359	66	1300	52	1129	57	1590	85
35–44 år	1106	70	948	48	776	56	974	64
45–54 år	732	48	736	65	547	35	531	61
55–59 år	180	13	179	17	180	12	189	16
60–64 år	126	4	110	11	95	7	116	13
65–74 år	63	1	62	6	60	6	90	7
Totalt	4353	227	4101	227	3378	188	4455	278

Kriminalvårdens data innehåller uppgifter om antal intagna efter medborgarskapsland medan AKU använder uppgifter om födelseland. För att konvertera Kriminalvårdens uppgifter till att vara uppdelade efter födelseland har vi för varje år under perioden 2005–2020 använt RTB för att ta fram andelarna som är inrikes födda respektive utrikes födda av de personer i befolkningen som har svenskt medborgarskap. Vi har även tagit fram samma andelar för de personer som saknar svenskt medborgarskap (se tabell B.14 för ett exempel).

Tabell B.14. Andelar inrikes födda respektive utrikes födda för personer som ingick i RTB 2005 och som hade svenskt medborgarskap respektive utländskt eller inget medborgarskap.

Medborgarskap	Andel inrikes födda	Andel utrikes födda
Svenskt	90,4%	9,6%
Utländskt eller saknas	7,9%	92,1%

Med hjälp av tabeller såsom den i tabell B.14 kan vi skatta andelen intagna som är inrikes födda. För 2005, exempelvis, då 3629 intagna hade svenskt medborgarskap och 957 utländskt medborgarskap blir antalet inrikes födda $3629 * 0,904 + 957 * 0,079 = 3356$ personer, vilket motsvarar 73,2% av antalet intagna. I modellen för antalet intagna antar vi att samma fördelning mellan inrikes födda och utrikes födda gäller oberoende av kön eller ålderskategori, men däremot att fördelningen mellan inrikes födda och utrikes födda varierar från år till år.

Vi har slutligen antagit att samtliga intagna hör till undergruppen ej i arbetskraften, övriga.

C. Modeller för latent arbetssökande

Skattningarna av det totala antalet latent arbetssökande är efter införandet av den nya ramlagen ungefär dubbelt så stora som tidigare. I de yngsta och äldsta åldersgrupperna är de procentuella förändringarna ännu större. Dessa förändringar av antalet latent arbetssökande har justerats i länken på makronivå med hjälp av extra blanketteffekter.

Som nämndes i avsnitt 6.2 har vi även tagit fram en imputeringsmodell för latent arbetssökande för att undvika två problem i mikrolänken. Utan imputering finns dels en risk för alltför stora viktjusteringar i mikrolänken, dels en risk för att det för vissa nedbrytningar saknas personer vars vikt vi kan justera vid mikrolänken för att säkerställa att de mikrolänkade och makrolänkade tidsserierna överensstämmer. Denna bilaga beskriver den framtagna imputeringsmodellen och även metoden som har använts för att härleda de extra blanketteffekterna för latent arbetssökande.

C.1. Brottskattningar för latent arbetssökande

Imputeringsmodellen bygger på en jämförande analys av AKU-data från sju år med det gamla förfarandet (2015–2021), där urvalet under 2021 är en femtedel av det ordinarie urvalet med ökad osäkerhet som följd, och AKU-data från två år med det nya förfarandet (2021–2022). Vi har även skattat tidsseriebrottet för latent arbetssökande med säsongrensning-programmet JDemetra och tagit in dessa skattningar i bedömningen (se tabell C.1). Tabell C.1 innehåller också skattningar från dubbelmätningen 2021 som jämförelse. Som synes är de båda uppsättningarna av brottskattningar relativt samstämmiga.

Tabell C.1. Skattningar av tidsseriebrott för latent arbetssökande baserat på dubbelmätningen respektive tidsserieanalys (JDemetra), uppdelat på åldersgrupper och mätt i tusental. Notera att brottskattningarna som bygger på tidsserieanalys är optimerade för varje tidsserie var för sig och de är därför inte summakonsistenta.

Ålder	Dubbelmätning	Tidsserieanalys
15 år	15.3	15.8
16–19 år	31.5	26.6
20–24 år	12.8	3.2
25–34 år	11.7	11.9
35–44 år	7.0	6.7
45–54 år	8.7	8.2
55–59 år	3.7	3.5
60–64 år	2.4	5.7
65–74 år	35.9	44.7
15–74 år	128.9	118.1

Målsättningen med imputeringsmodellen är att det imputerade mikro-datat ska överensstämma väl med den mikrodata som hade erhållits om vi hade använt det nya förfarandet under perioden 2005–2020. För att åstadkomma en god överensstämmelse bör de personer som imputeras till att ha varit latent arbetssökande vara lika, med avseende på relevanta variabler, de som verkligen hade klassificerats som latent arbetssökande med det nya förfarandet. Vi undersöker därför mönster gällande de som klassificerades som latent arbetssökande före och efter omläggningen och baserar modellen på denna information.

Latent arbetssökande är en undergrupp till gruppen ej i arbetskraften och denna grupp kan, oberoende av klassificering som latent arbets-sökande, delas in i fyra undergrupper: Heltidsstuderande, pensionärer, sjuka respektive övriga. Andra variabler som är relevanta för modellen är undersökningsmånad samt de variabler som under åren 2005–2020 användes i hjälpinformationen, nämligen ålder, kön, födelseland och region.

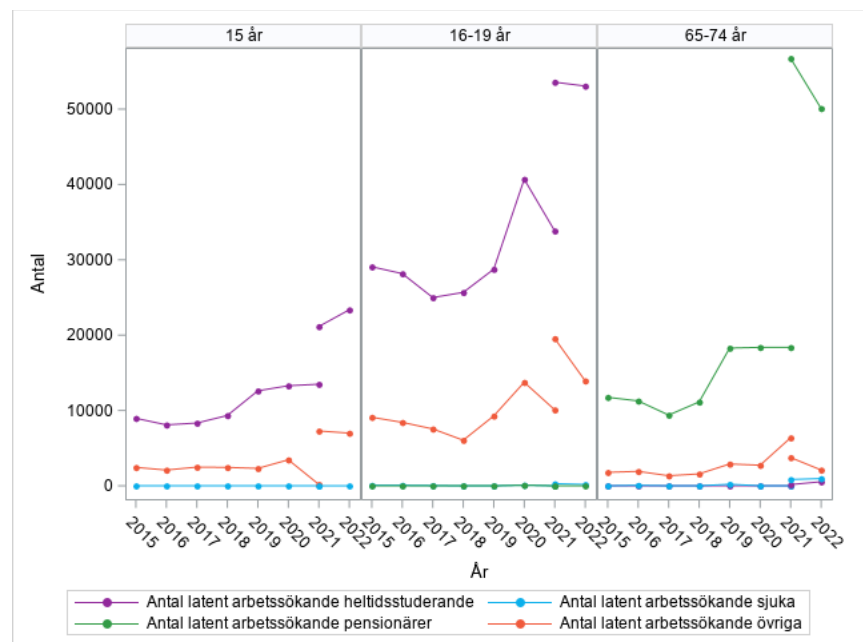
Eftersom skattningarna av tidsseriebrottet för latent arbetssökande är liknande för män och kvinnor har vi valt att inte ha med kön som variabel i modellen. Vidare är brottskattningarna för inrikes födda respektive utrikes födda proportionella mot dessa båda grupperns andel av befolkningen, varför bakgrund inte heller tas med som variabel i modellen. Vi försummar även region eftersom blanketteffekten för antalet latent arbetssökande rimligen bör vara oberoende av region. Det kvarstår att studera hur förändringen av latent arbetssökande beror på undersökningsmånad, ålder och de fyra undergrupperna till ej i arbetskraften.

Vi undersöker först hur antalet latent arbetssökande beror på ålder och de fyra undergrupperna till ej i arbetskraften. Figur C.1 visar antalet i de två yngsta åldersgrupperna och i den äldsta åldersgruppen, dvs i de åldersgrupper där förändringarna i och med det nya förfarandet var som störst, både procentuellt och i absoluta tal. Vi noterar att i dessa åldersgrupper skedde en ökning av antalet latent arbetssökande redan före införandet av den nya ramlagen. Åtminstone för åldersgruppen 16–19 år kan detta misstänkas vara en följd av coronapandemin. Brottskattningar som baseras på data från 2021 riskerar därför att bli oproportionerligt stora under större delen av länkningsperioden. För åldersgrupperna 15 år, 16–19 år och 65–74 år har vi därför valt att vara konservativa vid val av extra blanketteffekter. De effekter som införts i länkningen på makronivå är därför 20% lägre än de skulle ha varit om vi hade bedömt att 2021 hade varit fullt representativt för hela länkningsperioden.

För 15-åringarna är ökningen ca nio tusen personer heltidsstuderande och ca tre tusen personer övriga. I åldersgruppen 65–74 år är ökningen ca 36 tusen personer och hela ökningen sker för pensionärer.

Utvecklingen i åldersgruppen 16–19 år är däremot mer svårtolkad eftersom tidsserierna för antalet latent arbetssökande i det gamla förfarandet under åren 2020–2021 är relativt volatila och skattningen för 2020 ligger på en mycket högre nivå än tidigare år. För att reducera coroneffekten och effekten av det begränsade urvalet, så jämför vi skattningarna från 2019 och 2022 och drar slutsatsen att ökningen i gruppen heltidsstuderande är ca 24 tusen personer och i gruppen övriga ca fem tusen personer. Samtliga brottskattningar i dessa åldersgrupper har sedan reducerats med 20% (se tabell C.2).

Figur C.1. Antalet latent arbetssökande i olika undergrupper till ej i arbetskraften för åldersgrupperna 15 år, 16–19 år och 65–74 år under sju år med det gamla förfarandet (2015–2021) respektive två år med det nya förfarandet (2021–2022).

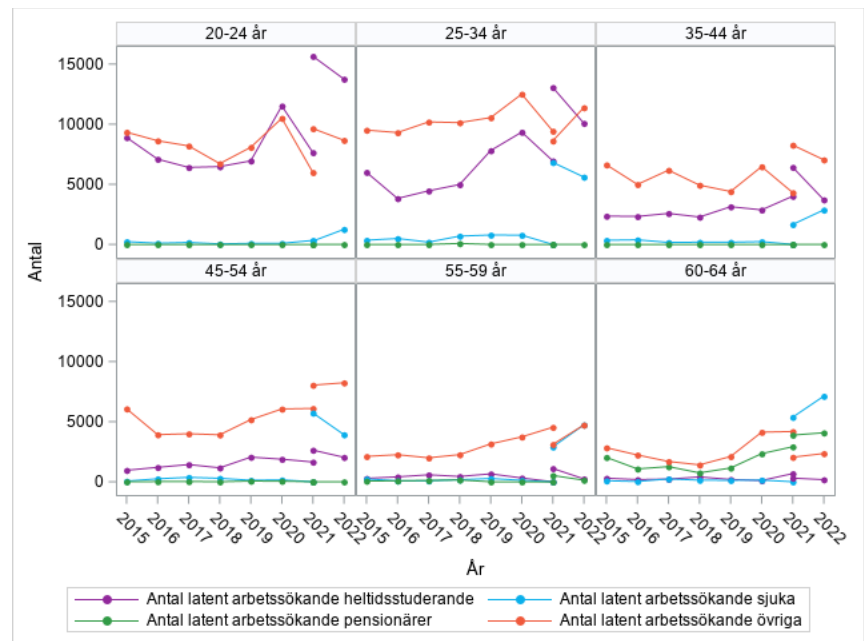


Figur C.2 visar antalet latent arbetssökande efter ålder och undergrupp till ej i arbetskraften för de åldersgrupper som inte fanns med i figur C.1. Precis som för åldersgruppen 16–19 år, så är utvecklingen för de heltidsstuderande, men även i viss mån för de övriga, här svårtolkad. Om vi även för åldersgrupperna i figur C.2 jämför skattningarna från 2019 och 2022, så ser vi en ökning på ca sju tusen personer i åldersgruppen 20–24 år och en ökning på ca två tusen personer i åldersgruppen 25–34 år, men obefintliga ökningar i övriga åldersgrupper. För gruppen övriga ser vi en ökning på ca tre tusen personer vardera för åldersgrupperna 35–44 år och 45–54 år, men i övrigt inga tydliga förändringar.

För gruppen sjuka är resultatet mer säkert, främst eftersom antalet latent arbetssökande sjuka var marginellt i det gamla förfarandet. Ökningarna i det nya förfarandet skattas till ca sex tusen personer i åldersgruppen 25–34 år, ca två tusen personer i åldersgruppen 35–44 år,

ca fem tusen personer i åldersgruppen 45–54 år, ca fyra tusen personer i åldersgruppen 55–59 år och ca sex tusen personer i åldersgruppen 60–64 år. I den sistnämnda åldersgruppen ser vi dessutom en ökning på ca ett tusen personer i gruppen pensionärer.

Figur C.2. Antalet latent arbetssökande i olika undergrupper till ej i arbetskraften för åldersspannet 20–64 år under sju år med det gamla förfarandet (2015–2021) respektive två år med det nya förfarandet (2021–2022).

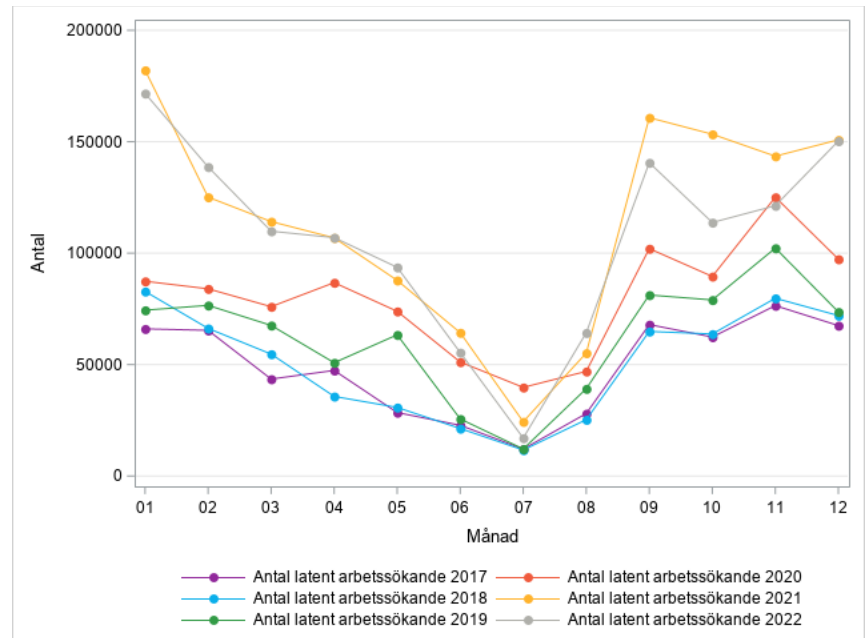


Vi undersöker härnäst om det finns en säsongseffekt i förändringen av andelen latent arbetssökande och plottar därför andelarna latent arbetssökande respektive ej latent arbetssökande som en funktion av månad och åldersgrupp. För sjuka och pensionärer finns inget synbart säsongsmönster och dessa plottar redovisas därför ej.

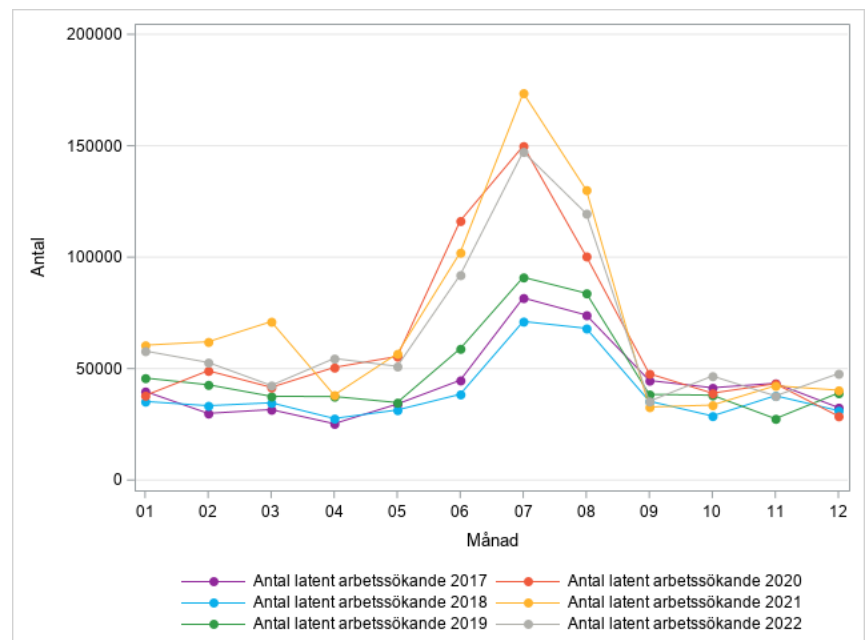
Som synes i figur C.3 och C.4, så är ökningen för heltidsstuderande betydligt större under terminerna än under sommaren och i juli går det inte att se någon ökning alls. Det motsatta gäller för gruppen övriga, där ökningen är betydligt större under sommarmånaderna och minimal under perioden september–maj. Om vi jämför ökningarna i de båda grupperna, så ser vi att under perioden juni–augusti så är ökningen i gruppen övriga störst, medan under resten av året är ökningen i gruppen heltidsstuderande störst.

Tabell C.2 sammanfattar extra blanketteffekter för latent arbetssökande uppdelat på åldersgrupper och undergrupper till ej i arbetskraften som vi har använt i länkningen på makronivå.

Figur C.3. Antalet latent arbetssökande heltidsstuderande för olika månader under fyra år med det gamla förfarandet (2017–2020) och två år med det nya förfarandet (2021–2022).



Figur C.4. Antalet latent arbetssökande övriga för olika månader under fyra år med det gamla förfarandet (2017–2020) och två år med det nya förfarandet (2021–2022).



Tabell C.2. Brottskattningar för latent arbetssökande, uppdelat på åldersgrupper och undergrupper till ej i arbetskraften, mätt i tusental personer.

Ålder	Heltids-studerande	Sjuka	Pensionärer	Övriga	Totalt
15 år	7			2	9
16–19 år	19			5	24
20–24 år	7				7
25–34 år	2	6			8
35–44 år		2		3	5
45–54 år		5		3	8
55–59 år		4			4
60–64 år		6	1		7
65–74 år			29		29
15–74 år	35	23	30	13	101

C.2. Imputeringsmodell för latent arbetssökande

Baserat på analysen i avsnitt C.1 härleder vi nu en imputeringsmodell för latent arbetssökande. För enkelhets skull så kommer vi för varje månad och åldersgrupp bara att imputera personer från en eller två av undergrupperna till ej i arbetskraften. I modellen har vi valt att försumma följande brottskattningar från tabell C.2: heltidsstuderande i åldersgruppen 25–34 år samt pensionärer i åldersgruppen 60–64 år.

För varje månad under länkningsperioden, så definierar vi nio olika imputeringsgrupper med urvalspersoner. Varje imputeringsgrupp innehåller personer som inte är latent arbetssökande och som tillhör en viss ålderskategori och en, eller i några undantagsfall två, undergrupper till ej i arbetskraften. Ur varje imputeringsgrupp kommer ett visst antal urvalspersoner att väljas ut slumpmässigt och klassificeras som latent arbetssökande. Hur imputeringsgrupperna definieras framgår av tabell C.3.

För var och en av de nio imputeringsgrupperna beräknar vi först summan av de kalibrerade (original)vikterna för alla urvalspersoner i imputeringsgruppen, här betecknad W . De makrolänkade tidsserierna baseras vidare på brottskattningarna i tabell C.2. Vi har härlett sådana skattningar, här betecknade L , för varje åldersgrupp och månad. Antalet personer N i imputeringsgruppen som vi väljer att klassificera om till latent arbetssökande ges sedan av kvoten L/W avrundat till närmaste heltal.

Om de N personerna som ska imputeras i varje imputeringsgrupp dras helt slumpmässigt ur imputeringsgruppen, så kommer korrelationen mellan personer som klassificeras som latent arbetssökande vid två på

varandra följande kvartal med stor sannolikhet att bli alltför låg. För att undersöka hur stor denna korrelation bör vara har vi studerat data från 2021–2022. Betrakta mängden personer som var latent arbetssökande en given månad under perioden januari 2021–september 2022. AKU-data visar att i snitt 40 % av dessa personer var latent arbetssökande även vid nästkommande intervjutillfälle tre månader senare. Dock inkluderas här exempelvis personer som var sysselsatta vid det senare intervjutillfället och dessa personer är inte av intresse vid imputering av latent arbetssökande. Om vi i stället bara betraktar personer som tillhör samma undergrupp till ej i arbetskraften vid båda intervjutillfällena, så är sannolikheten i snitt 60 % att en latent arbetssökande person är latent arbetssökande tre månader senare.

Tabell C.3. Definition av imputeringsgrupper för latent arbetssökande.

Ålder	September–Maj	Juni och augusti	Juli
15 år	Heltidsstuderande	Heltidsstuderande & Övriga	Övriga
16–19 år	Heltidsstuderande	Heltidsstuderande & Övriga	Övriga
20–24 år	Heltidsstuderande	Övriga	
25–34 år	Sjuka		
35–44 år	Sjuka & Övriga		
45–54 år	Sjuka & Övriga		
55–59 år	Sjuka		
60–64 år	Sjuka		
65–74 år	Pensionärer		

Vid imputering en given månad implementerar vi därför följande metodik. Vi sorterar först ut de personer i imputeringsgruppen som imputerades till att vara latent arbetssökande föregående kvartal och betecknar denna mängd med H . Låt nu M beteckna 60 % av antalet personer i H , avrundat till närmaste heltal. Vi väljer slumpmässigt ut M personer ur H och imputerar till att vara latent arbetssökande och tar sedan bort resterande personer i H från imputeringsgruppen. Från kvarstoden av imputeringsgruppen väljer vi slumpmässigt ut $N - M$ personer som vi imputerar till att vara latent arbetssökande.

C.3. Imputeringsmodell för önskat antal arbetade timmar för latent arbetssökande

För att kunna beräkna det möjliga antalet arbetade timmar för personer som tillhör det outnyttjade arbetskraftsutbudet, dit latent arbetssökande räknas, så behöver vi skatta hur många timmar de som har imputerats till att vara latent arbetssökande skulle kunna och vilja arbeta. Denna timvariabel har generellt inte samlats in för de som

imputerats och vi behöver därför imputera värden även för denna variabel.

För att konstruera en imputeringsmodell för timvariabeln har vi undersökt värdena på timvariabeln för de som har klassats som latent arbetssökande under länkningsperioden. Med homogenitetstest har vi undersökt om timvariabeln beror på intervjumånad, år, ålder och vilken undergrupp till ej i arbetskraften som urvalspersonen tillhör.

Homogenitetstesten visar att det finns ett starkt beroende på månad (längre önskad arbetstid under sommarmånaderna), ålder (längre önskad arbetstid för medelålders) och undergrupp (längre önskad arbetstid för övriga än för de andra undergrupperna), men att beroendet på år är relativt svagt även om den önskade arbetstiden generellt har blivit något kortare under senare år.

För var och en av årets tolv månader och var och en av de nio olika imputeringsgrupperna den månaden (se tabell C.3) har vi skapat en mängd som består av samtliga insamlade värden på timvariabeln för latent arbetssökande urvalspersoner i AKU som uppfyllde kriterierna för imputeringsgruppen under den givna månaden under något av de 16 åren 2005–2020. Från denna mängd har vi sedan, för varje person som imputerats till att vara latent arbetssökande i den imputeringsgruppen den månaden, dragit ett slumpmässigt valt värde på timvariabeln.