

Fortsatt diskussion av vägningstalens utformning med hänsyn till coronakrisen

Coronapandemins effekter på SCB:s inflationsmätningar har diskuterats i två tidigare pm till nämnden; Carlsson och Ståhl (2020) och Ståhl (2020). Föreliggande pm ger ytterligare underlag för diskussion kring en specifik fråga, nämligen vägningstalens utformning i månadslänken under 2021 och 2022. Nämnden inbjuds att komma med synpunkter.

1. Syfte

Syftet med promemorian är att utgöra underlag för diskussion kring utformningen av vägningstal inom KPI, KPIF och HIKP för år 2021 och 2022. För KPI/KPIF:s del ligger fokus på utformningen av de så kallade månadslänkarna, medan årslänkarna kommer att diskuteras på ett framtida möte.

2. Vägningstal i HIKP och KPI

KPI, KPIF och HIKP beräknas alla som kedjeindex. Nya indextal skapas varje månad genom att en indexlänk som sträcker sig fram till och med den sista månaden kedjas/länkas på en tidigare serie. För HIKP gäller att denna sista länk sträcker sig från december föregående år till och med aktuell månad (y,m) :

$$HIKP_{2015}^{y,m} = HIKP_{2015}^{y-1,12} \times I_{y-1,12}^{y,m} \quad (1)$$

För KPI och KPIF består den sista länken egentligen av två delar - en ”årslänk” och en ”månadslänk” – och sträcker sig hela vägen bak till helåret ($y-3$). För de frågor som diskuteras i denna pm gäller att utformning är identisk mellan KPI och KPIF och därmed kommer enbart KPI att diskuteras. Fokus ligger också, som redan nämnts, på den allra sista delen av kedjan d.v.s. på månadslänken. Denna går från helåret ($y-2$) till och med aktuell månad:

$$KPI_{1980}^{y,m} = KPI_{1980}^{y-3} \times I_{y-3}^{y-2} \times I_{y-2}^{y,m} \quad (2)$$

Med ”vägningstal” avses, slutligen, de vikter som används i samband med att delindex för enskilda produktgrupper aggregeras till en sammansatt totallänk. Den fråga som diskuteras i föreliggande pm är med andra ord vilka vikter som ska användas inom de två länkarna $I_{y-1,12}^{y,m}$ och $I_{y-2}^{y,m}$, i formel (1) respektive (2), under år 2021 och 2022.

2.1 Vägningstal i HIKP

För HIKP:s del sker aggregeringen till den sammansatta länken $I_{y-1,12}^{y,m}$ enligt följande:

$$I_{y-1,12}^{y,m} = \sum_{\forall g} \{w^y(g) \cdot I_{y-1,12}^{y,m}(g)\}$$

där $I_{y-1,12}^{y,m}(g)$ betecknar ett delindex som beskriver prisutvecklingen från december föregående år till och med aktuell månad, för produktgrupp g , $w^y(g)$ är produktgruppens vägningstal, och summeringen görs över alla produktgrupper som ingår i HIKP-korgen. Vägningstalen ges vidare av:

$$w^y(g) = \frac{U^{y-2}(g) \cdot I_{y-2}^{y-1,12}(g)}{\sum_{\forall g} \{U^{y-2}(g) \cdot I_{y-2}^{y-1,12}(g)\}} \quad (3)$$

där $U^{y-2}(g)$ betecknar värdet av den totala privata inhemska hushållskonsumtionen under år ($y-2$) inom produktgrupp g .

Formel (3) beskriver SCB:s normala förfarande vid beräkningen av HIKP-vägningstal.¹ Implicit bygger detta förfarande på ett antagande om att konsumtionens fördelning mellan olika produktgrupper under år ($y-2$) även är representativ för fördelningen under år ($y-1$); detta eftersom målvariabeln för HIKP egentligen är ett Laspeyres-typ index med

¹ I praktiken görs beräkningen av vägningtalsbelopp på KPI-produktgruppsnivå, och fördelas sedan på i HIKP ingående elementära produktgrupper enligt COICOP-klassificeringen. För enkelhets skull bortser vi från denna detalj i promemorian.

helårsvikter från ($y-1$). I HIKP-manualen (Eurostat, 2018, s.30) preciseras detta enligt följande:

”Member States shall produce HICPs using sub-index weights which [...] aim to be as representative as possible for consumers’ expenditure patterns in the previous calendar year“

Enligt HIKP-förordningen (Eurostat, 2020a, artikel 3), ska dock HIKP:s vägningstal i praktiken baseras på data från år ($y-2$) men samtidigt ”göras representativa” för just helåret ($y-1$):

“Member States shall derive the sub-index and elementary aggregate weights used in the index for year t as follows: [...]

[National accounts data] for year $t-2$ and any available and relevant information from household budget surveys and other data sources shall be used to obtain subclass expenditure shares and divide them among the elementary aggregates of the subclass [...]

The expenditure shares for year $t-2$ shall be reviewed and updated to make them representative of year $t-1$.“

Som nämnades i tidigare pm till nämnden har det emellertid varit oklart exakt hur olika länder i praktiken tolkat formuleringen om att ”göra vikterna representativa för år ($y-1$)”, något som blivit extra tydligt i och med diskussionerna under coronakrisen. Mot bakgrund av detta arbetar Eurostat just nu med en särskild vägledning angående viktberäkningarna för 2021 och 2022. Vägledningen kommer att diskuteras på ett möte i slutet av november men ett utkast har redan skickats ut till medlemsländerna. Här tydliggörs att den metod som normalt sett tillämpas av bland annat Sverige är fullt tillräcklig under ”normala tider” men att detta inte gäller för 2021 och 2022:

“In normal times, structural changes between $t-2$ and $t-1$ are limited so that $t-2$ data can be used to estimate $t-1$ (either by price-updating or not price-updating between $t-2$ and $t-1$). Clearly, this is not the case when consumption expenditure changes significantly, both in level and structure, between $t-2$ and $t-1$, such as in the current COVID-19 pandemic. “

I utkastet rekommenderas länderna att inför 2021 och 2022 använda ”preliminary national accounts data and other sources” för att se till att vikterna blir representativa för året innan (d.v.s. för 2020 respektive 2021).

SCB avvaktar Eurostats formella vägledning innan slutligt beslut tas, men har börjat planera för den händelse att särskilda åtgärder kan behöva vidtas i årets viktberäkningar. Om det visar sig att de rekommendationer som

föreslås i utkastet kvarstår, så planerar SCB att inför 2021 (och 2022) applicera en framskrivningsmetod för att göra HIKP-vägningstalen mer representativa för föregående års konsumtionsmönster. Rent tekniskt så innebär detta att HIKP-vägningstalen under 2021 och 2022 då inte skulle räknas enligt formel (3) utan istället med följande justerade formel:

$$w^y(g) = \frac{U^{y-2}(g) \cdot F_{y-2}^{y-1}(g) \cdot I_{y-1}^{y-1,12}(g)}{\sum_g \{U^{y-2}(g) \cdot F_{y-2}^{y-1}(g) \cdot I_{y-1}^{y-1,12}(g)\}}$$

där $\hat{F}_{y-2}^{y-1}(g)$ betecknar en skattad ”framskrivningsfaktor” vilken beskriver hushållskonsumtionens utveckling, i löpande priser, mellan helåren ($y-2$) och ($y-1$) inom den aktuella produktgruppen, g , och $I_{y-1}^{y-1,12}(g)$ är ett index som beskriver prisskillnaden mellan helåret ($y-1$) och december samma år.

Delindexet ovan är av en typ som normalt sett inte används i HIKP eller KPI men kan förslagsvis skapas utifrån reviderade elementära produktgruppsindex, $I_{y-1,12}^{y,m}(g)$, enligt följande:²

$$I_{y-1}^{y-1,12}(g) = \frac{I_{y-2,12}^{y-1,12}(g)}{\frac{1}{12} \sum_{m=1}^{12} I_{y-2,12}^{y-1,m}(g)}$$

I avsnitt 3 återkommer vi till hur framskrivningsfaktorerna, $\hat{F}_{y-2}^{y-1}(g)$, är tänkta att tas fram i praktiken.

Flera skäl finns till varför en justerad beräkning av HIKP-vägningstalen planeras att baseras på framskrivningsfaktorer, istället för på en mer direkt metod. För det första så finns underlag gällande hushållskonsumtionen under 2019 redan tillgängligt, och SCB har därmed redan påbörjat beräkningen av beloppen $U^{2019}(g)$. (Se bilagan för en beskrivning av vilka underlag som används för vägningstalsberäkningarna.) Kvartalsuppgifter från Nationalräkenskaperna (NR) avseende kvartal 3 2020 blir dock tillgängliga först i slutet av november, och Eurostats formella vägledning ser ju dessutom ut att dröja till början av december. Det har alltså bedömts nödvändigt att påbörja beräkningarna innan den slutliga vägledningen finns på plats, vilket möjliggörs av tillvägagångssättet med separata framskrivningsfaktorer. En fördel med det förslagna beräkningssättet är också att ”vanliga vägningstal” (vägningstal vilka *inte* justerats med framskrivningsfaktorer) skapas som en biprodukt av kalkylen; dessa kan längre fram användas för att beräkna effekten av den särskilda justeringen under 2021 och 2022. Slutligen finns även vissa fördelar när det kommer till granskning; faktorerna kommer att behöva baseras på ett mer osäkert underlag än det som vanligen används för att räkna konsumtionsbelopp inom varje produktgrupp, och kan därmed behöva justeras för extremvärden eller misstänkta felaktigheter. Det kan då vara en pedagogisk

² För detaljer kring dessa elementärindex, se SCB (2020a).

fördel om 2019-beloppen är konsistenta med Nationalräkenskapernas värden medan eventuella justeringar först och främst görs i framskrivningsfaktorerna.

2.2 Vägningstal i KPI

För KPI:s månadslänk sker aggregeringen till en totallänk enligt följande:

$$I_{y-2}^{y,m} = \sum_{\forall g} \{w^y(g) \cdot I_{y-2}^{y,m}(g)\} \quad (4)$$

där delindex $I_{y-2}^{y,m}(g)$ beskriver prisutvecklingen mellan helåret ($y-2$) och aktuell månad (y,m) för produktgrupp g , $w^y(g)$ är vägningstalet, och summeringen görs över alla produktgrupper som ingår i KPI-korgen. KPI-vägningstalen beräknas vidare normalt sett enligt:

$$w^y(g) = \frac{U^{y-2}(g)}{\sum_{\forall g} U^{y-2}(g)} \quad (5)$$

Mot bakgrund av SCB:s preliminära planer för beräkningen av HIKP-vägningstal inför 2021 och 2022 ligger det nära till hands att diskutera följande tre alternativa beräkningsmetoder för KPI-vägningstalen under samma period:

- (i) Räkna vägningstal med standardmetod d.v.s. enligt formel (5).
- (ii) Räkna vägningstal baserat på uppskattade konsumtionsbelopp för år ($y-1$), i praktiken med hjälp av framskrivningsfaktorerna;

$$w^y(g) = \frac{U^{y-2}(g) \cdot \bar{F}_{y-2}^{y-1}(g) / I_{y-2}^{y-1}(g)}{\sum_{\forall g} \{U^{y-2}(g) \cdot \bar{F}_{y-2}^{y-1}(g) / I_{y-2}^{y-1}(g)\}}$$

där $I_{y-2}^{y-1}(g)$ är ett delindex som beskriver prisutvecklingen mellan helåren ($y-2$) och ($y-1$) för produktgrupp g . (Dessa delindex skulle av praktiska skäl behöva räknas i en satellitkalkyl utanför produktions-systemet, men konstrueras förslagsvis efter samma princip som de vanliga produktvisa årslänkarna; jmf SCB, 2020a.)

- (iii) Räkna vägningstal baserat på ett genomsnitt av konsumtionsbeloppen ($y-2$) och uppskattade konsumtionsbelopp för ($y-1$), återigen med hjälp av framskrivningsfaktorerna;

$$w^y(g) = \frac{U^{y-2}(g) \cdot \sqrt{\bar{F}_{y-2}^{y-1}(g) / I_{y-2}^{y-1}(g)}}{\sum_{\forall g} \{U^{y-2}(g) \cdot \sqrt{\bar{F}_{y-2}^{y-1}(g) / I_{y-2}^{y-1}(g)}\}}$$

Utöver metoderna (i) – (iii) så skulle ett tänkbart alternativ kunna vara att utöka kedjan (2) med ytterligare en årslänk, och komplettera detta med en kortare månadslänk:

$$KPI_{1980}^{y,m} = KPI_{1980}^{y-3} \times I_{y-3}^{y-2} \times \boxed{I_{y-2}^{y-1} \times I_{y-1}^{y,m}} \quad (6)$$

Vi kallar detta för alternativ (iv). Den extra årslänken skulle kunna räknas enligt standardmetod (jmf SCB, 2020a), men baserat på approximativa (framskrivna) viktuppgifter för år $(y-1)$:

$$I_{y-2}^{y-1} = \sum_{y,g} \{w(g) \cdot I_{y-2}^{y-1}(g)\}, \text{ där } w(g) = \frac{U^{y-2}(g) \cdot \sqrt{\hat{F}_{y-2}^{y-1}(g)/I_{y-2}^{y-1}(g)}}{\sum_{y,g} \{U^{y-2}(g) \cdot \sqrt{\hat{F}_{y-2}^{y-1}(g)/I_{y-2}^{y-1}(g)}\}}$$

Månadslänken skulle i sin tur precis som idag kunna räknas enligt Laspeyres formel, vilket i detta fall motsvarar:

$$I_{y-1}^{y,m} = \sum_{y,g} \{w^y(g) \cdot I_{y-1}^{y,m}(g)\}$$

med

$$w^y(g) = \frac{U^{y-2}(g) \cdot \hat{F}_{y-2}^{y-1}(g)}{\sum_{y,g} \{U^{y-2}(g) \cdot \hat{F}_{y-2}^{y-1}(g)\}}$$

och där själva delindexerna, $I_{y-1}^{y,m}(g)$, kan räknas enligt:³

$$I_{y-1}^{y,m}(g) = \frac{I_{y-1,12}^{y,m}(g) \cdot I_{y-2,12}^{y-1,12}(g)}{\frac{1}{12} \sum_{m=1}^{12} I_{y-2,12}^{y-1,m}(g)}$$

I nästa avsnitt diskuteras argument för och emot de olika alternativen.

Det bör också nämnas att vi på föregående möte diskuterade två ytterligare alternativ till formulering av månadslänken, vilka inte byggde på framskrivningsfaktorer. Anledningen till varför vi valt att inte följa upp dessa alternativ i föreliggande pm är att det nu står klart att vägledningen från Eurostat angående HIKP:s vikter med stor sannolikhet kommer att innehålla någon form av rekommendation för framskrivning. På föregående möte uppgav flera av nämndens ledamöter att någorlunda konsistens i hanteringen mellan HIKP och KPI/KPIF är att föredra, och SCB har därför bedömt att en eventuell justering av KPI:s vägningstal i så fall borde baseras på liknande framskrivningsfaktorer som inom HIKP, snarare än att göras med en helt separat metod.

³ Här representerar index $I_{y-1,12}^{y,m}(g)$ ett så kallat preliminärt elementärt produktgruppsindex medan övriga komponenter är reviderade produktgruppsindex; jmf SCB (2020a).

2.3 Argument för alternativ (i) – (iv)

Grunden till KPI:s nuvarande indexkonstruktion lades i samband med 1999 års offentliga utredning om översyn av konsumentprisindex (SOU 1999:124). I denna fastslås att målvariabeln för KPI är ett så kallat betingat levnadskostnadsindex, samt att ett sådant index bäst approximeras med hjälp av en "superlativ" indexformel.⁴ Resonemanget hade direkt bäring på valet av konstruktion för KPI:s årslänkar, vilka räknas i enlighet med teorin för ett Walshindex. Vägningstal för årslänkarna baseras därmed på geometriska genomsnitt av två intilliggande års konsumtionsbelopp (justerat för prisförändring).⁵ För månadslänken är det dock inte möjligt (och kanske inte heller optimalt) att använda en superlativ formel. I utredningen beskrivs istället en målsättning för månadslänken vara att i någon mening "prognosticera" framtida resultat enligt den superlativa årslänken (s. 56).⁶ Olika alternativa formuleringar för månadslänken diskuteras i utredningen, och jämförs även empiriskt. Den formel som till slut väljs är konstruerad i enlighet med teorin för Laspeyres indexformel. Huvudargumentet för att använda just Laspeyres formel verkar vara enkelhet (på sidan 219 noteras att "*simplicity speaks in favour of the Laspeyres index*"). Skillnaderna i resultat mot de mer komplicerade alternativ som också testas visar sig dessutom vara små.

Mot bakgrund av den diskussion som fördes i utredningen, samt i ljuset av de praktiska förutsättningar som föreligger, kan vi se följande argument för och emot de olika alternativ som listades i avsnitt 2.2:

- (i) För alternativ (i) talar bland annat enkelhet och tydlighet. Beräkningar görs då i enlighet med etablerad praxis och är därmed transparenta för användare. Risken för manuella fel minimeras också.

Ett mer teoretisk argument för alternativ (i) är att månadslänken helt enkelt är *definierad* som ett Laspeyresindex, vilket innebär att konsumtionsbelopp per definition bör tas från prisbasperioden (i detta fall helåret y-2). Till skillnad från i HIKP så finns ju ingen uttalad ambition för månadslänken att ha vikter som är "representativa för föregående helår". Medan den planerade hanteringen inom HIKP är i linje med HIKP:s legala ramverk så utgör därmed en eventuell

⁴ En superlativ indexformel är en indexformel som ger ett exakt (teoretiskt) levnadskostnadsindex givet en så kallad "flexibel" enhetskostnadsfunktion; jmf SOU 1999:124, s.174.

⁵ Alla superlativa indexformler har vikter som är symmetriska, d.v.s. som bygger på konsumtionsbelopp från såväl prisbasperioden som aktuell period.

⁶ En superlativ formel för månadslänken är inte praktiskt tillämplig av flera skäl. För det första på grund av att konsumtionsuppgifter blir tillgängliga först med viss eftersläpning, för det andra på grund av att dessa uppgifter vanligen enbart finns tillgängliga på helårsbasis medan en superlativ månadslänk skulle kräva månatliga konsumtionsuppgifter.

framskrivning av KPI:s vikter i större utsträckning ett avsteg från den grundtanke som framlades i samband med den offentliga utredningen. Det som talar *emot* alternativ (i) är att inflationstakten enligt KPI och KPIF under 2021 i så fall skulle baseras helt på vägningstal som inte är anpassade till coronasituationen. Detta kommer troligtvis även att leda till större diskrepanser än vanligt mellan de två måtten KPIF och HIKP, givet att justerade vägningstal används för det senare måttet.

- (ii) Alternativ (ii) skulle kunna betraktas som mest konsistent med idén om månadslänken som en prediktor för framtida årslänk. Om man betraktar månadslänkens huvuduppgift som att prognosticera en superlativ årslänk så innebär det att vikterna allra helst skulle vara baserade på konsumtionsbelopp från både år $(y-1)$ och y . Detta är förstås praktiskt omöjligt, men man kan argumentera för att alternativ (ii) är det som bäst approximerar ett sådant ideal. Dessutom är (ii) kanske det alternativ som mest liknar den approach som troligtvis kommer att väljas för HIKP.

Nackdelarna med alternativ (ii) är bland annat praktiska, då det innebär en specialhantering under 2021 och 2022. Det finns även en risk för att användare betraktar en eventuell specialjustering som godtycklig - särskilt om kriterierna för när en alternativ viktning ska användas framstår som otydliga.

- (iii) Alternativ (iii) kan möjligen betraktas som det mest optimala alternativet för månadslänken som sådan; då vikterna i detta fall baseras på konsumtionsbelopp från såväl länkens prisbasperiod $(y-2)$ som mellanperioden $(y-1)$ så ligger alternativet relativt nära en superlativ form av månadslänk – en sådan länk skulle i teorin ha vägningstal baserade på konsumtionsbelopp från år $(y-2)$ såväl som den senaste perioden (y,m) . Notera dock att utredningen snarare specificerar målsättningen för månadslänken som att prognosticera en framtida superlativ *årslänk*, vilket inte nödvändigtvis är detsamma som att approximera en superlativ månadslänk.

Nackdelarna med alternativ (iii) är i princip desamma som för alternativ (ii). Om en målsättning är att ha så hög konsistens som möjligt mellan KPI/KPIF och HIKP så kan även det faktum att konsumtionen under $(y-2)$ här tillåts påverka vikterna i större utsträckning ses som en nackdel. Å andra sidan lär detta förfarande kunna ge viss robusthet gentemot volatilitet i vikterna, särskilt med tanke på att ett mer osäkert underlag kommer att behöva användas för framskrivningen.

- (iv) Alternativ (iv), med en extra årslänk och en förkortad månadslänk, är tilltalande rent principiellt då det passar väl in i nuvarande indexkonstruktion. Tyvärr skulle detta alternativ innebära betydande beräkningstekniska svårigheter då SCB:s produktionssystem idag är

helt och hållet anpassade efter formel (4).⁷ Förutom att komplicera beräkningarna under 2021 och 2022 finns därmed en påtaglig risk för att alternativ (iv) skulle leda till tekniska komplikationer under ytterligare flera år framöver. SCB bedömer därmed inte detta alternativ som en framkomlig väg på kort sikt.

2.4 Kommentar angående korgeffekten

Valet av vikter under 2021 och 2022 kommer att påverka den så kallade ”korgeffekten” under perioden 2021-2023, då denna beror direkt av vägningstalens storlek. Om en annan konstruktion än vanligt används för månadslänken så kommer även korgeffekten att rent tekniskt få en annan form. Den mer intuitiva tolkningen av korgeffekten som ”den inflationstakt som skulle ha blivit resultatet om inga priser ändrats [sedan samma period föregående år / sedan föregående månad]” kvarstår dock.

Detta skulle eventuellt kunna betraktas som en nackdel jämfört mot att behålla ordinarie viktning, då korgeffekten redan idag uppfattas som svårtolkad av många användare. Det kan bli svårt att kommunicera skillnader mellan uppkommen korgeffekt och den korgeffekt som hade blivit fallet med ordinarie metod. I princip är det emellertid möjligt för SCB att beräkna båda dessa effekter och redovisa skillnaderna, något som skulle kunna vara av intresse åtminstone för en del mer avancerade användare.

3. Framskrivningsfaktorer

Beräkningen av de alternativa vägningstalen utgår alla från att framskrivningsfaktorer, $F_{y-2}^{y-1}(g)$, går att skatta på ett någorlunda korrekt sätt för samtliga produktgrupper. I själva verket finns betydande svårigheter relaterat till detta. Framskrivningsfaktorerna kommer till stor del att behöva baseras på ett mindre detaljerat underlag än vad som är fallet för ordinarie vägningstalsberäkningar. Dessutom blir preliminära uppgifter från NR avseende helårskonsumtionen 2020 inte tillgängliga förrän i slutet av februari 2021 – faktorerna kommer därmed att behöva skattas utifrån uppgifter för enbart de tre första kvartalen, kompletterat med någon form av modellberäkning (prognos) för det fjärde kvartalet. Nedan diskuterar vi dessa två frågor i mer detalj. Nämndens synpunkter kommer att vara viktiga för utformningen av det fortsatta arbetet.

⁷ Man kan visa att produkten av de två sista komponenterna i (6) inte kan skrivas på formen (4) med vikter som summerar till ett över alla produktgrupper, vilket är vad produktionssystemet kräver.

3.1 Detaljeringsnivå

Beräkningen av framskrivningsfaktorer kommer, åtminstone till största delen, att baseras på kvartalsunderlag från NR. Detta underlag är inte lika detaljerat som de uppgifter som normalt sett används för vägningstalsberäkningarna. Faktorer kommer därmed inte att kunna beräknas separat för varje enskild produktgrupp, g , utan istället på en högre nivå omfattande drygt 100 NR-aggregat enligt COICOP-klassificeringen.

Fördelningen inom varje NR-aggregat kommer i första hand att antas proportionell mot $(y-2)$ års fördelning. Aggregat som uppvisar en särskilt kraftig utveckling mellan 2019 och 2020 planeras dock att analyseras i mer detalj, och om möjligt så kommer ytterligare underlag att eftersökas för dessa. I tabell 1 listas de NR-aggregat som hittills identifierats som sådana där en finare fördelning kan behövas (listan är preliminär).

Tabell 1: Exempel på aggregat inom vilka ytterligare underlag kan behövas.

NR-aggregat (COICOP)	Motsvarande produktgrupper i KPI
07.3.3 (Lufttransporter, äv helikopter)	6305 (Inrikes flygresor) 6311 (Utrikes flygresor)
09.4.2.7 (Teater, konserthus, nöjeställn)	7804 (Teaterbiljett) 7805 (Inträdesbiljett, nöjen)
11.2 (Hotellservice, övernattningsservice)	9603 (Stugbyvistelse) 9604 (Hotellrum, vardag) 9605 (Hotellrum, veckoslut) 9606 (Camping)

Det är värt att kommentera att en del av de underlag som normalt sett används för att uppskatta fördelningar under $(y-2)$ inom NR-aggregaten eventuellt även skulle vara möjliga att få tillgång till för $(y-1)$. SCB har dock bedömt att det relativt tidskrävande utredningsarbete som skulle krävas för att uppskatta explicita fördelningsmatriser baserat på $(y-1)$ istället för $(y-2)$ i samtliga fall där detta vore möjligt troligtvis inte skulle ge särskilt stora effekter på slutresultatet. I praktiken används i många fall fördelningsmatriser baserade på äldre underlag än $(y-2)$ även i de ordinarie beräkningarna (se bilagan för detaljer).

Frågan om detaljeringsnivå diskuteras även i Eurostats utkast till vägledning, men de rekommendationer som ges är relativt generella. Man påpekar dock specifikt att vissa delar av konsumtionen är viktigare att detaljstudera än andra:

“As a minimum, the expenditures of the most heavily affected segments of consumption should be re-estimated. These are typically (but not exclusively) fuels, passenger transport (in particular by air), recreational and cultural services, package holidays, restaurants and hotels. This list may differ from country to country.”

I utkastet listas även ett antal möjliga källor i form av korttidsstatistik som skulle kunna vara användbar för att åstadkomma en finare fördelning. (Samma källor rekommenderas även som möjligt underlag för prognostisering av det fjärde kvartalet – jämför avsnitt 3.2 nedan.) Även här ges rum för relativt olika tillvägagångssätt inom olika länder. Bland annat skriver man:

“Countries’ NSIs may have more or more detailed data available [...] e.g. from high-frequency household budget surveys or ad-hoc COVID-oriented surveys. NSIs may also have access to other data sources, for example from credit card data or scanner data. Many national central banks have set-up high-frequency COVID dashboards that could be consulted.”

Sammanfattningsvis så reflekterar formuleringarna i utkastet Eurostats vilja att möjliggöra för de länder som har tillgång till detaljerade datakällor att faktiskt använda dessa, samtidigt som det finns en förståelse för att alla länder inte har samma möjligheter. Man trycker också på att *“it is important that countries use all available data sources to make the best estimates possible”*.

3.2 Modellberäkning för kvartal 4

Tolkningen av en framskrivningsfaktor, $F_{y-2}^{y-1}(g)$, är som kvoten mellan konsumtionsbeloppet under år $(y-1)$ och år $(y-2)$ inom produktgrupp g :

$$F_{y-2}^{y-1}(g) = \frac{U^{y-1}(g)}{U^{y-2}(g)}$$

Enligt diskussionen i avsnitt 3.1 planeras framskrivningsfaktorer i första hand att skattas på en högre nivå än för varje enskild produktgrupp, nämligen för de drygt 100 olika NR-aggregaten. Vi betecknar fortsättningsvis denna nivå med bokstaven a , och använder beteckningen $U^{y,k}(a)$ för att beskriva hushållskonsumtionen under kvartal k år y inom aggregatet a , vilket ger:

$$F_{y-2}^{y-1}(a) = \frac{\sum_{k=1}^4 U^{y-1,k}(a)}{\sum_{k=1}^4 U^{y-2,k}(a)}$$

Vid den tidpunkt då vägningstalen måste vara klara gäller att uppgifter för kvartal fyra $(y-1)$ inte finns tillgängliga, vilket innebär att någon form av modellberäkning av det fjärde kvartalets konsumtion kommer att krävas. Mer eller mindre sofistikerade metoder är förstås tänkbara och vid val av metod bör även transparens vägas mot andra kvalitetsaspekter. De allra enklaste metoderna är baserade uteslutande på en framskrivning av uppgifterna från de tre första kvartalen. Exempelvis så kan utvecklingen

under kvartal fyra, i nominella termer, antas motsvara den under kvartal ett till tre, vilket ger följande skattning:

$$(i) \quad \hat{F}_{y-2}^{y-1}(a) = \frac{\sum_{k=1}^3 U^{y-1,k}(a)}{\sum_{k=1}^3 U^{y-2,k}(a)}$$

Ett annat alternativ är att anta att volymutvecklingen (utvecklingen av hushållskonsumtionen i reala termer) under det fjärde kvartalet liknar genomsnittet under de tre första kvartalen:

$$(ii) \quad \hat{F}_{y-2}^{y-1}(a) = \frac{\sum_{k=1}^3 U^{y-1,k}(a) + U^{y-2,4}(a) \cdot I_{y-2,4}^{y-1,4}(a) \frac{\sum_{k=1}^3 (U^{y-1,k}(a)/I_{y-2,k}^{y-1,k}(a))}{\sum_{k=1}^3 U^{y-2,k}(a)}}{\sum_{k=1}^4 U^{y-2,k}(a)}$$

där $I_{y-2,k}^{y-1,k}(a)$ betecknar är ett index som beskriver prisutvecklingen mellan kvartal k år ($y-2$) och samma kvartal nästkommande år.⁸ Ett sådant index kan beräknas utifrån KPI:s elementära (reviderade) produktgruppsindex, förslagsvis enligt följande:

$$I_{y-2,k}^{y-1,k}(a) = \sum_{g \in a} \{w(g) \cdot I_{y-2,k}^{y-1,k}(g)\}, \text{ där } I_{y-2,k}^{y-1,k}(g) = \frac{I_{y-3,12}^{y-2,12}(g) \sum_{m \in k} I_{y-2,12}^{y-1,m}(g)}{\sum_{m \in k} I_{y-3,12}^{y-2,m}(g)}$$

och $w(g)$ ges av formel (4).

Ytterligare två alternativ är att utgå från antagande om att ingen värdeutveckling alls ägt rum sedan samma period föregående år;

$$(iii) \quad \hat{F}_{y-2}^{y-1}(a) = \frac{\sum_{k=1}^3 U^{y-1,k}(a) + U^{y-2,4}(a)}{\sum_{k=1}^4 U^{y-2,k}(a)}$$

alternativt att ingen volymutveckling ägt rum;

$$(iv) \quad \hat{F}_{y-2}^{y-1}(a) = \frac{\sum_{k=1}^3 U^{y-1,k}(a) + U^{y-2,4}(a) \cdot I_{y-2,4}^{y-1,4}(a)}{\sum_{k=1}^4 U^{y-2,k}(a)}$$

Istället för dessa enkla alternativ, vilka enbart utnyttjar information från de senaste sju kvartalen, skulle tidsseriemodellering av en längre historisk serie för varje aggregat kunna vara en framkomlig väg. Det är emellertid förstås oklart i vilken mån säsongsmönstret under 2020 kan antas likna det som gällt under tidigare år.

Som ett första steg i utredningsarbetet har de fyra enkla metoderna (i) – (iv) testats på historiska data för perioden 2015 – 2019. Detta har skett

⁸ Notera att formeln utgår från att enbart konsumtionsbelopp i löpande priser finns tillgängliga – vilket är fallet i skrivande stund – ett (troligtvis bättre) alternativ är förstås att även efterfråga fastprisberäknade belopp från Nationalräkenskaperna och använda dessa för att uppskatta volymutvecklingen för det fjärde kvartalet.

genom att de skattade framskrivningsfaktorerna jämförts med motsvarande faktorer räknade utifrån NR:s preliminära årsberäkningar (motsvarande den version som normalt sett används som underlag för KPI:s vägningstalsberäkningar). Rent generellt kan man förstås i någon mån ifrågasätta värdet av historiska jämförelser i det här fallet, med tanke på hur pass speciella konsumtionsmönstren varit under just 2020. Även för det historiska materialet gäller dock att mönstren inte är särskilt tydliga. En jämförelse på totalnivå (se tabell 2) indikerar att den i någon mening allra enklaste metoden – metod (i) ovan – fungerar väl så bra som övriga tre alternativ.⁹ Resultaten varierar dock, som väntat, mellan olika aggregat. I tabell 3 redovisas andelen aggregat för vilka respektive metod gav minst fel (definierat i termer av kvadratavvikelsen). För ungefär en tredjedel av de drygt 100 aggregaten så gav metod (i) det minsta kvadratfelet, medan en ungefär lika stor andel av aggregaten prognostiserades bäst med metod (ii).

Tabell 2: Relativt prognosfel jämfört med metod (i). Värden större än ett indikerar att metoden ger ett i genomsnitt större prognosfel än metod (i), medan värden mindre än ett indikerar att metoden ger ett i genomsnitt mindre prognosfel än metod (i).

Metod / implicit antagande	\hat{F}_{2015}^{2016}	\hat{F}_{2016}^{2017}	\hat{F}_{2017}^{2018}	\hat{F}_{2018}^{2019}
(i) Värdeutveckling likadan som under kvartal 1-3	1,00	1,00	1,00	1,00
(ii) Volymutveckling likadan som under kvartal 1-3	1,02	1,00	1,01	1,00
(iii) Ingen värdeutveckling	1,11	1,07	1,01	1,00
(iv) Ingen volymutveckling	1,09	1,05	1,01	0,99

Tabell 3: Andel aggregat för vilka respektive metod gav minst fel (i procent).

Metod / implicit antagande	\hat{F}_{2015}^{2016}	\hat{F}_{2016}^{2017}	\hat{F}_{2017}^{2018}	\hat{F}_{2018}^{2019}
(i) Värdeutveckling likadan som under kvartal 1-3	36,1	30,6	27,8	30,5
(ii) Volymutveckling likadan som under kvartal 1-3	25,9	33,3	28,7	30,6
(iii) Ingen värdeutveckling	21,3	19,4	27,8	20,4
(iv) Ingen volymutveckling	16,7	16,7	15,7	18,5
	100,0	100,0	100,0	100,0

⁹ Totalfelet för en viss period och metod har räknats som $\sqrt{\sum_{va} \left[w^y(a) \cdot \left(\hat{F}_{y-2}^{y-1}(a) - F_{y-2}^{y-1}(a) \right)^2 \right]}$, med vikter konstruerade i enlighet med formel (4), men summerat per aggregat, a . Beräkningen av $\hat{F}_{y-2}^{y-1}(a)$ har genomgående baserats på de kvartalsuppgifter som publicerades av NR i samband med första publiceringen av det tredje kvartalet år $(y-1)$, för att i så hög grad som möjligt efterlikna den verkliga situationen.

SCB kommer framöver att analysera resultaten i mer detalj – det är möjligt att mer tydliga slutsatser kan dras för enskilda aggregat. Resultaten planeras även att jämföras mot en explicit formulerad tidsseriemodell.

Givet att målet för framskrivningsfaktorerna är att uppnå ett så litet prognosfel som möjligt så torde det allra mest optimala vara att kombinera en ”automatisk” (statistisk) framskrivningsmetod med ytterligare källor för respektive aggregat. Exempel på sådana källor kan vara prognoser gjorda av andra myndigheter eller olika typer av preliminär månadsstatistik. SCB publicerar exempelvis sedan 2013 den månatliga så kallade ”HUKO-indikatorn” (*Hushållens konsumtionsindikator*) ca 40 dagar efter referensmånadens slut; novembersiffran kommer alltså att publiceras i början av januari 2021. HUKO-indikatorn är avsedd att prognosticera just nationalräkenskapernas kvartalsberäkningar, dock på relativt aggregerad nivå (ungefär motsvarande KPI:s huvudgrupper).¹⁰ Arbetet med att hitta lämpliga underlag för respektive aggregat bör idealt bedrivas i samarbete med experter från nationalräkenskaperna och primärstatistiken.

Ett alternativ är att betrakta transparens och enkelhet som ett mål i sig; det kan då finnas skäl att välja en mer generell metod även om en sådan skulle ge mindre optimala prognoser för varje enskilt aggregat.

4. Frågor till nämnden

- Vilka synpunkter har nämnden gällande vilken formel som är mest lämplig att använda för att räkna månadslänkar i KPI och KPIF under 2021 och 2022, mot bakgrund av Eurostats kommande vägledning för HIKP?
- Vilka synpunkter har nämnden på hur eventuella framskrivningsfaktorer lämpligen ska beräknas? Är transparens ett viktigt mål i sig eller bör varje aggregat behandlas separat?
- Har nämnden några generella synpunkter kring vilka kriterier som bör vara uppfyllda för att SCB ska överväga att justera sin ordinarie indexkonstruktion?
- Har nämnden några synpunkter på hur SCB bäst kommunicerar eventuella ändringar i beräkningsprinciperna till viktiga användare och till allmänheten?

¹⁰ Statistikunderlaget som används för beräkning av HUKO-indikatorn är i grunden branschuppdelat men fördelas på ändamål med hjälp av vikter från nationalräkenskapernas årsberäkningar. Tio olika så kallade ”analytiska aggregat” produceras vilka relativt väl följer huvudgruppsindelningen enligt COICOP (enbart nio av dessa publiceras dock i SSD). För närmare information se <https://www.scb.se/hitta-statistik/statistik-efter-amne/handel-med-varor-och-tjanster/>.

Referenser

Carlsson, E. och Ståhl, O. (2020), *Hantering av effekter av Corona-pandemin i KPI och HIKP*, PM inför möte 008 i Nämnden för KPI den 1 april 2020.

<https://www.scb.se/contentassets/1b48f2064ebd46a78eda4d68d51c0403/hantering-av-effekter-av-corona-pandemin-i-kpi-och-hikp---uppdaterad-2020-05-11.pdf>

Ståhl, O. (2020), *Vägningstal och indexkonstruktion i samband med coronakrisen*, PM inför möte 009 i Nämnden för KPI den 25 september 2020.

<https://www.scb.se/contentassets/1b48f2064ebd46a78eda4d68d51c0403/9/vagningstal-och-indexkonstruktion-i-samband-med-coronakrisen.pdf>

Eurostat (2018), *Harmonised Index of Consumer Prices (HICP) Methodological manual*, <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GO-17-015>

Eurostat (2020a), *Commission Implementing Regulation 2020/1148 of 31 July 2020*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2020.252.01.0012.01.ENG.

SCB (2020a), *Statistikens framställning, KPI 2020*, version 2020-02-19.

https://www.scb.se/contentassets/a1e257bb3a574420b9d3f2ff59851c0a/pr0101_staf_2020_kk_200219.pdf

SCB (2020b), *Kvalitetsdeklaration, Nationalräkenskaper, kvartals- och årsberäkningar*, version 2020-05-29.

https://www.scb.se/contentassets/66e9dae3a5d94bf8b4c299ce25294348/nr0103_kd_2020_hk_020527.pdf

Bilaga: Om underlagen för vägningstalsberäkningarna

Hushållskonsumtion i NR

Nationalräkenskaperna publicerar både års- och kvartalsvisa beräkningar över hushållskonsumtionen. Årsberäkningarna, vilka är betydligt mer detaljerade, publiceras 17 månader efter referensåret medan kvartalsberäkningarna publiceras ca två månader efter referenskvartalet.

Skillnaderna är relativt stora mellan förfarandet bakom års- och kvartalsberäkningarna. Årsberäkningarna används för att bestämma själva nivån på hushållskonsumtionen, medan kvartalsberäkningarna främst bygger på framskrivningar av senast beräknad årsnivå utifrån olika typer av korttidsindikatorer.

Kvartalsberäkningarna revideras i samband med varje ny publicering. När en årsberäkning sedan blir färdig så används den även som benchmark för motsvarande kvartal; dessa revideras då alltså en sista gång, på sådant sätt att helårstotalerna enligt bägge serier stämmer överens.

Nationalräkenskapernas redovisning av hushållskonsumtionen sker fördelat på ändamål enligt COICOP, d.v.s. enligt samma klassifikation som används inom KPI och HIKP. Kvartalsuppgifterna redovisas i SSD på huvudgruppsnivå, men prisenheten har även tillgång till mer detaljerade uppgifter omfattande de ca 100 aggregat som nämns i denna pm.

Underlag för KPI och HIKP:s vägningstalsberäkningar

I samband med att hushållskonsumtionsberäkningen för det fjärde kvartalet publiceras så beräknar NR även en preliminär årstotal bestående av summan av de fyra ingående kvartalsskattningarna. Det är (en reviderad version av) dessa ”preliminära årsberäkningar” som används av prisenheten som huvudkälla för såväl KPI:s som HIKP:s vägningstalsberäkningar.¹¹

För fördelningar på finare nivå används uppgifter från NR:s årsberäkningar för ett tidigare år ($y-3$). Inom delar av COICOP 03, 05, 09 och 12 används även uppgifter från undersökningen Hushållens utgifter (HUT), där de senast publicerade resultaten avser år 2012. Därutöver utnyttjas ytterligare ett antal interna och externa källor vilka i vissa fall finns tillgängliga för ”rätt” tidsperiod, d.v.s. avseende helåret ($y-2$), men i andra fall enbart för en tidigare eller senare period.

Mer information om NR

För ytterligare information om nationalräkenskapernas kvartals- och årsberäkningar hänvisas till SCB (2020b).

¹¹ För de belopp avseende ($y-3$) som går in i årslänkens vägningstalsberäkning utnyttjas de definitiva årsberäkningarna som huvudkälla.