

## Användning av kassaregisterdata från Systembolaget för egen uträkning av index

### *För beslut*

I samband med att beräkningar av prisindex för alkoholhaltiga drycker flyttas från Systembolagets datasystem till SCB föreslås följande principer i beräkningen av prisindex:

- Beräkna priskvot per produkt med genomsnittspris över 3 veckors försäljning;
- Tillämpa årliga produkturval;
- Tillämpa metodik med bytesvaror och kvalitetsjusteringar vid utgående artikelid;
- Exkludera pant från prisjämförelser.

### INNEHÅLL

#### ANVÄNDNING AV KASSAREGISTERDATA FRÅN SYSTEMBOLAGET FÖR EGEN UTRÄKNING AV INDEX1

<b>1</b>	<b>BAKGRUND</b> .....	<b>2</b>
<b>2</b>	<b>METOD</b> .....	<b>2</b>
2.1	NUVARANDE BERÄKNINGAR.....	2
2.2	HANTERING I EGET BERÄKNINGSSYSTEM .....	2
<b>3</b>	<b>GENOMFÖRANDE</b> .....	<b>3</b>
3.1	LEVERANS OCH LAGRING AV DATA.....	3
3.2	MARKNADSANALYS.....	3
3.3	VALIDERING, SELEKTERING, INLÄSNING OCH GRANSKNING .....	4
3.4	KVANTITETS- OCH KVALITETSVÄRDERING .....	4
<b>4</b>	<b>ANALYS</b> .....	<b>4</b>
4.1	EXEMPEL 1, CIDER.....	4
4.2	EXEMPEL 2, GLÖGG .....	5
<b>5</b>	<b>DISKUSSION</b> .....	<b>6</b>
<b>6</b>	<b>FÖRSLAG</b> .....	<b>6</b>



## 1 Bakgrund

Sedan en tid beräknar och levererar Systembolaget prisindex för ett tiotal produktgrupper i början av varje månad, enligt SCB:s beräkningsinstruktioner. I anslutning till Kassaregisterprojektet för dagligvaror (jämför nämndmöte nummer 244 "Kassaregister i KPI") har det diskuterats att flytta dessa beräkningar till SCB, och istället efterfråga leveranser av statistikuppgifter på produktnivå från Systembolagets centrala försäljningssystem (cirka 20 000 sålda produkter per månad över hela landet). Det primära motivet till denna förändring är att uppnå bättre kontroll över beräkningarna av index, bland annat ur ett kvalitetssäkringsperspektiv.

## 2 Metod

### 2.1 Nuvarande beräkningar

De prisnoteringar som används vid de månadsvisa jämförelserna avser varans försäljningspris den sista dagen i månaden, det vill säga pris inklusive moms inklusive pant. För varje produkt beräknas priskvoten relativt priset den 31:e december föregående år. Prisindex per produktgrupp beräknas med ett geometriskt medelvärde, viktat med respektive produkts försäljningsandel under föregående helår:

$$PI_{dec,y}^{m,y} = \exp \left( \sum_k \frac{Fsg_{k,y-1}}{\sum_l Fsg_{l,y-1}} \ln \left( \frac{Pr_{k,m,y}}{Pr_{k,dec,y-1}} \right) \right)$$

I ovanstående uttryck är underförstått att varor som ej sålts under jämförelsemånad alternativt referensmånad undantas summeringarna. En varas identitet bestäms av de artikel-ID som används i Systembolagets datasystem.

De olika produktgrupperna är fastställda enligt SCB:s önskemål. För att avgöra en varas produktgruppstillhörighet krävs en nyckeltabell, som kopplar Systembolagets varuklassificeringar till SCB:s klassificering. Denna nyckeltabell bör uppdateras kontinuerligt, åtminstone årligen, eftersom nya varuklassificeringar kan dyka upp i Systembolagets datasystem, medan andra försvinner. För beräkning av HIKP krävs en annorlunda produktgruppsindelning, som utgår från COICOP5 klassificeringen. Detta uppnås med en parallell nyckeltabell visavi Systembolagets varuklassificeringar.

En nyckel för COICOP5 är framtagen med hjälp av unika kombinationer av produktbeskrivningar, Systembolagets produktgruppsindelning och alkoholhalt från den senaste åren. Här kvarstår dialog med Systembolaget för att säkerställa att alla nya produkter tilldelas en COICOP5-grupp i SCB:s beräkningar.

### 2.2 Hantering i eget beräkningssystem

ES/PR eftersträvar enhetlighet i sina produktionssystem, med tanke på förvaltning och utveckling. Dagens centrala produktionssystem, Pi09, erbjuder tillfredsställande möjligheter till indexberäkning och granskning av enskilda produktnoteringar. Systemet erbjuder även möjligheten att arbeta med ersättningsvaror och kvalitetsvärderingar. Dock föreligger vissa prestandabegränsningar i det befintliga produktionssystemet, vilket man erfarit i det tidigare kassaregisterprojektet för dagligvaror.

Bedömningen är att ett urval av produkter ger bättre kontroll och gör materialet mer hanterbart. Fördelningen av försäljningsvolym per produkt är vanligen mycket skev, det vill säga med en mindre grupp av varor som står för en stor del av försäljningen, samt en större grupp av varor med var för sig försvinnande liten försäljningsandel (bland annat

beställningsvaror). Som lösning på problemet föreslås ett årligt urvalsförfarande av produkter, snarlikt det som tillämpas för dagligvaror. Bedömningen är att ett slumpmässigt urval om totalt cirka 1 000 produkter skulle resultera i en mycket god skattning av den totala prisutvecklingen.

Befintlig indexformel (se avsnitt 2.1) bedöms fortsatt adekvat, men med definitionsmässiga justeringar avseende de prisnoteringar som jämförs. Istället för pris per den sista i månaden föreslås ett medelpris över tre veckors försäljning. Med medelpris avses här totala försäljningsintäkter för produkten (exklusive pant) dividerat med antal sålda exemplar. Med andra ord,

$$Pr_{k,m,y} = \frac{\sum_{i=1}^3 Fsg_{k,v_i,m,y}}{\sum_{i=1}^3 Kv_{k,v_i,m,y}}$$

De tre ingående veckorna väljs enligt kriteriet: veckan som innefattar den 15:e dagen i månaden samt närmast föregående och närmast påföljande vecka.

### 3 Genomförande

Detta avsnitt avser att belysa det praktiska arbete som kommer att göras på ES/PR vid införandet av kassaregisterdata och egna beräkningar av Systembolagets data. Till stora delar liknar det arbetet som idag genomförs vid bearbetning av kassaregisterdata från övriga uppgiftslämnare. Arbetsmomenten som beskrivs är hämtade från arbetsbeskrivningar för dessa men är ännu inte testade på data för alkoholprodukter.

Arbetet med att förbereda produktionssystemet Pi09 återstår att genomföra. Det kommer att tidigast bli aktuellt att se över under hösten 2015.

#### 3.1 Leverans och lagring av data

Förhoppningen är att få på plats ett avtal mellan ES/PR och uppgiftslämnaren som reglerar att data skickas varje vecka via FTP. Enligt uppgifter från uppgiftslämnaren kommer leverans av veckovisa filer tidigast kunna ske under hösten 2015.

#### 3.2 Marknadsanalys

Vid bearbetning av övriga kassaregisterfiler som behandlas på enheten används EAN-koder som är kopplade till de produktkoder som används internt på ES/PR. EAN-kod är kopplade via en nyckel till ett produktnummer. Då en produkt ändrar utseende, kvantitet, smak eller andra kvaliteter ändras EAN-koden. Vissa ändringar i kvalitet får som effekt att produkten inte längre är representativ medan andra förändringar är av mindre betydelse. Om man strikt går på EAN-koder vid identifiering av produkter riskerar man att tappa prisobservationer. Vid marknadsanalys ser man över antalet sålda enheter för att identifiera produkter som är på väg att fasas ut och produkter där försäljningen gått ner men det finns en likvärdig ersättningsvara eller bytesvara. Idag sker detta arbetet manuellt på ES/PR.

I fallet med alkoholhaltiga drycker ingår inte EAN-koder i uppgiftslämnarens data. Istället används artikel-ID som är kopplat till ett artikelnamn. Därför kommer varje produktnummer att vara kopplat till ett artikel-ID istället för en EAN-kod. En fördel med detta är att artikel-ID inte uppdateras med samma känslighet som EAN-koder. Produkter med samma artikel-ID kan därför anses likvärdiga. Trots att detta minskar risken för att tappa prisobservationer måste en viss marknadsanalys genomföras. Dialog med Systembolaget rörande hur artikel-

ID och nyckeln mellan dessa och den av SCB specificerade produktgruppsindelningen hanteras är påbörjad.

### 3.3 Validering, selektering, inläsning och granskning

Dessa moment genomförs idag för kassaregisterdata. Nedan sammanfattas kort hur detta kommer att genomföras med de filer som kommer från Systembolaget. Alla steg kanske inte kommer att vara aktuella samt att ordningen kan komma att ändras med hänsyn till filernas utformning.

Efter marknadsanalysen kontrolleras hela indatafilen från uppgiftslämnaren, så kallad validering. Antalet prisobservationer jämförs med tidigare månad för att se att inga större förändringar ägt rum. I samband med selekteringen granskas de prisobservationer som ingår i urvalet för undersökningen.

En fördel med att beräkna index lokalt på SCB är att alla prisobservationer kan samlas i Pi09 vilket medför bättre översikt och möjlighet till smidigare mikrogranskning på ES/PR. Då selekteringen är gjord läses prisobservationer in till Pi09 i likhet med övrig kassaregisterdata.

Efter selektering och inläsning till Pi09 av de aktuella prisobservationerna genomförs en granskning i likhet med övrig kassaregisterdata. Priskvoter och prisnivåer granskas med hjälp av prishistorik.

### 3.4 Kvantitets- och kvalitetsvärdering

Vid införandet av urval för alkoholprodukter vill SCB göra det möjligt för byten. Som en följd till detta krävs möjlighet till kvantitet- och kvalitetsjusteringar.

Eftersom Systembolagets data innehåller uppgifter om volym (ml) bedöms ändringar i kvantitet vara lätta att upptäcka. Möjligen kan det bli aktuellt att införa en procentgräns för hur stora förändringar som tillåts kvantitetsjusteras. I exemplet med cider nedan föreslås hur en sådan kvantitetsjustering kan gå till.

Bedömningen av produktens kvalitet och värderingen av kvalitetsskillnaden bedöms kunna göras av handläggare på ES/PR och, i de fall där det krävs, i samråd med Systembolaget. Denna kvalitetsvärderingsmetod är alltså subjektiv och skiljer sig inte från den metod som i dagsläget görs på andra produkter i KPI.

## 4 Analys

Detta avsnitt avser att belysa behovet av att arbeta med ersättningsvaror, kvalitets- och kvantitetsvärderingar, så kallad marknadsanalys.

### 4.1 Exempel 1, Cider

Tabell 1 illustrerar kvarstående försäljningsandel 2014 vid tre tillfällen under 2015, då enbart försäljningen från artikel-ID som fortfarande finns till försäljning summeras. Noterbart är en tydlig nedgång avseende produktgruppen *cider* någon gång under januari månad.

Fenomenet får till sin förklaring i Tabell 2. Man kan notera att tidigare storsäljare (fiktiva artikel-ID 2, 4 och 6) upphör att finnas till försäljning. Parallellt förekommer motsvarande vara i mindre paketering (fiktiva artikel-ID 1, 3 och 5), men till betydligt lägre försäljningsandel under helåret 2014. Det förefaller rimligt att betrakta de tre senare varorna

som ersättningsvaror för tidigare storsäljare, med en kvantitetsjustering baserat på den relativa skillnaden ( $y/x = 1,14$ ).

**Tabell 1: Andelen kvarvarande produkter i sortimentet 1:a januari, 1:a februari respektive 1:a mars 2015, sett till deras försäljningsandel under helåret 2014.**

SCB Beskrivning	Procent		
	Januari	Februari	Mars
Övrigt vin	99	99	99
Övrig sprit	100	99	99
Whisky	99	99	99
Vita viner	99	99	99
Starköl	99	99	99
Röda viner	99	99	99
Mousserande viner	99	99	99
Mellanöl	99	98	98
Likör	100	100	100
Cider	99	73	73
Brännvin	100	100	100
Blanddrycker	100	99	99

**Tabell 2: Fiktiva artikel-ID och delvis undertryckta artikelnamn, deras försäljningsandel i produktgruppen Cider under 2014, samt varans pris 1:a januari, 1:a februari respektive 1:a mars 2015.**

Artikel-ID	ArtikelNamn	Andel 2014	Pris		
			Januari	Februari	Mars
1	A, burk x ml	1%	19,00	19,00	18,90
2	A, burk y ml	13%	19,40	.	.
3	B, burk x ml	1%	19,00	19,00	18,90
4	B, burk y ml	9%	19,40	.	.
5	C, burk x ml	0,2%	16,30	16,30	15,90
6	C, burk y ml	3%	16,90	.	.

#### 4.2 Exempel 2, Glögg

Säsongsmönster förekommer i försäljningen av alkoholhaltiga drycker. Tydliga exempel är roséviner sommartid, glögg kring jul, samt påsk- respektive julöl. Systembolaget anpassar sitt tillfälliga sortiment efter tidpunkt på året.

Tabell 3 avser att illustrera karaktäristiska fenomen avseende glöggförsäljningen i det centrala försäljningssystemet. Det är inte ovanligt att samma (eller i stort sett samma) artikelnamn dyker upp i systemet men med olika artikel-ID. Vid jämförelse mellan de två åren 2013 och 2014 förefaller det exempelvis rimligt att betrakta artikel-ID 3 som en ersättningsvara för artikel-ID 2, utan någon indikation om behov av kvalitetsjustering. Vi noterar i så fall en prisökning från 84 kronor till 88 kronor. Motsvarande fenomen dyker upp även upp för artikel-ID 5 och 6.

Artikel-ID 7-9 nedan avser tre besläktade produktbjudanden med något olika smak, och som delvis förkommer parallellt på marknaden. Storsäljaren 2013 förekommer således även under 2014, men ersätts i popularitet av ett nytt artikel-ID. Betraktar vi artikel-ID 9 som en ersättningsvara för artikel-ID 8, utan förändring i kvalitet, sker en prisökning från 97 kr till 100 kr mellan säsongerna.

Tabell 3: Fiktiva artikel-ID och undertryckta artikelnamn inom segmentet glögg, deras totala försäljning december 2013 respektive 2014, samt varans pris 1:a januari 2014 respektive 1:a januari 2015.

Artikel-ID	ArtikelNamn	Fsg 2013	Fsg 2014	Pris 140101	Pris 150101
1	Glögg A	79 kr	.	79 kr	.
2	Glögg A	13 milj	.	84 kr	.
3	Glögg A	.	15 milj	.	88 kr
4	Glögg B	79 kr	.	79 kr	.
5	Glögg B	11 milj	.	84 kr	.
6	Glögg B	.	13 milj	.	88 kr
7	Glögg C1	300 tusen	25 tusen	94 kr	94 kr
8	Glögg C2	20 milj	130 tusen	97 kr	97 kr
9	Glögg C3	.	15 milj	.	100 kr

## 5 Diskussion

De två exemplen i avsnitt 3 indikerar förhållandevis enkla val av ersättningsvaror. Kvalitets-/kvantitetsjusteringen i exemplen är endera obefintlig, eller baserad på relativ förändring i kvantitet. Behovet uppstår i vissa fall genom att samma vara tilldelas olika artikel-ID, i andra fall genom att producenterna genomför smärre justeringar i kvalitet, paketering och volym. Bibehållen kontinuitet i prisjämförelserna är av stor betydelse för att uppnå god kvalitet i prisindex.

Medelpriset över tre veckors försäljning i mitten av månaden bör anses tillräckligt representativt för konsumentens upplevda pris för produkten under aktuell månad, enligt det mönster som idag tillämpas vid beräkning av prisindex för dagligvaror. Denna metod medför bland annat bättre tidsmarginaler i SCB:s beräkning och kontroll av prisindex för alkoholhaltiga drycker.

Vanligen sker prisrevideringar av hela sortimentet per den förste i månaden. Med dagens metod ger en prisrevidering i månad  $t$  upphov till förändringar i prisindex mellan månaderna  $t-1$  och  $t$ . Den föreslagna alternativa metoden uppnår samma resultat i normalfallet, eftersom prisindex för månad  $t$  avser att spegla konsumentens upplevda pris under månaden.

Idag jämförs priser på alkoholhaltiga drycker enligt definitionen inklusive moms inklusive pant, medan icke-alkoholhaltiga drycker jämförs enligt definitionen inklusive moms exklusive pant. Det förefaller rimligt att tillämpa samma regel för såväl alkoholhaltiga som alkoholfria drycker.

## 6 Förslag

- Efterfråga veckovisa leveranser på produktnivå från Systembolaget, med statistik över aktuell försäljning.
- Möjliggör val av ersättningsvaror, kvalitets- och kvantitetsvärderingar i de fall betydande produkter i den årsvisa varukorgen upphör att finnas till försäljning, alternativt ändrar karaktär ur ett konsumentperspektiv.
- Jämför priser på produktnivå enligt definitionen inklusive moms exklusive pant.

- Tillämpa ett årligt urval av produkter, med urvalssannolikheter relativt föregående års försäljningsvärde.