

ES/PR
Rodica Bubioc
Brunilda Sandén
Miykal Tareke
John Johansson
Can Tongur

PM till Nämnden för KPI
Sammanträde nr 5
2018-10-16

Ny datakälla för Järnvägs-, flyg- och charterresor

För information

Undersökningarna för järnvägs-, flyg- och charterresor utgör tillsammans 2,5 procent av korgen i KPI. Dessa undersökningar har alla relativt små urval idag. Priserna samlas antingen in genom manuella provbokningar av biljetter på internet eller via blanketter till uppgiftslämnare. Prisenheten har nu tillgång till nya digitala datakällor som planeras ersätta den manuella prisinsamlingen i dessa undersökningar. Förutom en minskad uppgiftslämnarbörda och prisinsamlingsbörda, medför det också mer kvalitativa prismätningmetoder samt större urval.

Innehåll

1	SYFTE	1
2	FLYGCHARTER	2
2.1	BESKRIVNING AV DEN NYA DATAKÄLLAN	2
2.2	BERÄKNING AV PRISINDEX	3
3	JÄRNVÄGSRESOR	3
3.1	BEFINTLIGA PRISMÄTNINGAR	4
3.2	NYA PRISMÄTNINGAR	4
3.3	FÖRDELAR MED DEN NY DATAKÄLLA	4
4	FLYGRESOR	5
4.1	UNDERSÖKNINGEN I NULÄGET	5
4.2	NY DATAINSAMLINGSMETOD	5
5	REFERENSER	6

1 Syfte

Syftet med denna PM är att informera om nya datakällor för delundersökningarna järnvägs-, flyg- och charterresor, samt informera om hur undersökningarna kan komma att utformas i



och med dessa nya möjligheter. Prisenheten ämnar designa om dessa undersökningar från och med indexåret 2019. Nämnden välkomnas komma med synpunkter på de planerade ändringarna.

I följande tre avsnitt följer mer ingående information om de tre olika delundersökningarna.

2 Flygcharter

KPIs delundersökning Flygcharter avser mätning av priser för paketresor. Priser samlas in för ett urval av typiska charterdestinationer genom att undersökningen täcker dimensionerna charteroperatör och resepaket i form av olika specifikationer av boende och resesällskap.

Undersökningen genomförs manuellt – genom enkäter som månatligen skickas ut till operatörerna innehållande ett antal resespecifikationer, s.k. produkterbjudanden enligt KPI. Eftersom det är en undersökning som resulterar i tämligen tung uppgiftslämnarbörda måste urvalen hållas begränsade. Urvalen är samtidigt inte så stora, i relation till mångfalden i utbudet.

Undersökningen är behäftad med variansproblem vilket har konstaterats vid genomgång av KPIs delundersökningar, se nämndpappret ”En modell för resursallokering inom KPI” (Norberg, Strandberg & Ståhl, 2018). För att motverka urvalsosäkerheten skulle urvalet behöva ökas med en faktor 2,8 vilket i praktiken är svårt med tanke på dagens uppgiftslämnarbörda.

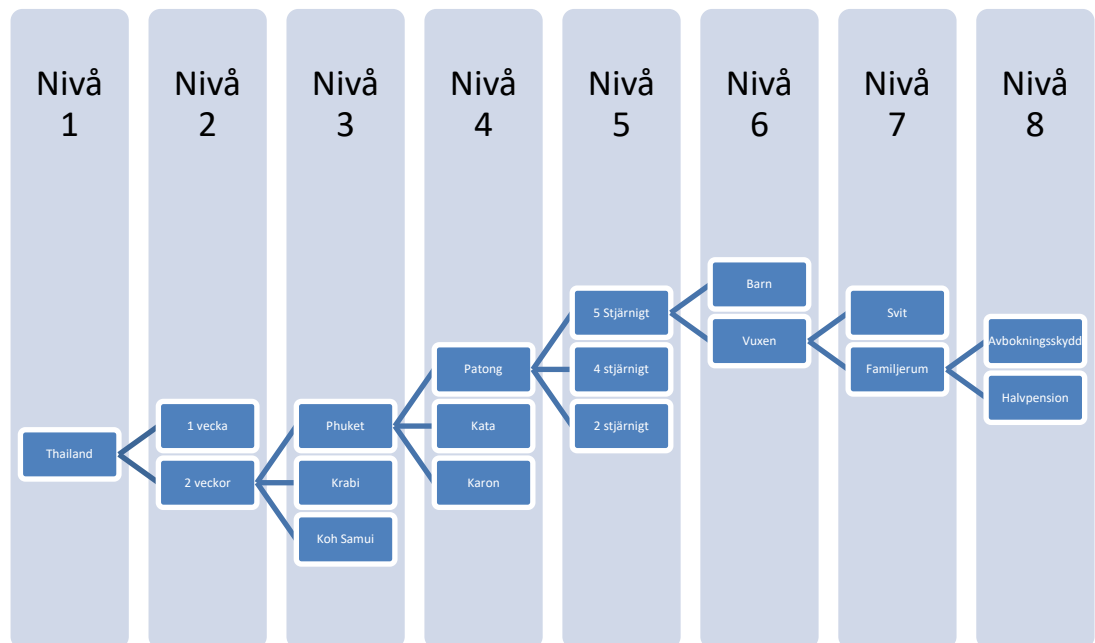
Som ett steg i att öka användningen av digitala datakällor har Prisenheten det senaste året haft kontakter med flera charteroperatörer. Prisuppgifter likt kassaregisterdata har undersökts från respektive operatör, vilka planeras användas i produktion från 2019.

För att behandla två starka säsonger tillämpas sedan 2017 en metod som sammanväger ett vinterindex med ett sommarindex, benämnd som ”Metod 2” i tidigare nämndunderlag (Johansson & Ståhl, 2017). Nämnden gav sitt gillande till denna metod av fyra möjliga för att hantera säsongproblematiken. Prisenheten ämnar inte byta säsongsimputeringsmetod i och med den nya datakällan, utan använder den redan implementerade metoden från 2017.

2.1 Beskrivning av den nya datakällan

Det nya datamaterial som SCB har undersökt innefattar paketresor d.v.s. både flyg och logi. Data levereras som aggregat med snittpriser uppdelat på avreseort, destination, hotell, rumstyp och reslängd. Uppgifter om antal försäljningar finns också. Data täcker företagets hela försäljning av paketresor.

I figur 1 illustreras olika nivåer vid specifikation av en resa. Modellen illustrerar åtta nivåer av detaljering, subjektivt formulerade här för att visa på ingående faktorer vid val av resa.



Figur 1: Exempel på olika nivåer vid specifikation av en resa

Det kan ses från figur 1 att detaljeringen kan bero på förutsättningarna – exempelvis kan undervariationerna vara många fler för en destination (nivå 4) genom att det finns olika typer av hotell (nivå 5) medan andra destinationer kanske är begränsade avseende underliggande utbud.

Nuvarande urvalsundersökning görs genom specificerade produkt erbjudanden ända till nivå 8, d.v.s. ytterst detaljerade specifikationer som dessutom besvaras genom enkät till uppgiftslämnarna. Det nya tillgängliga datamaterialet motsvarar en detaljering på nivå 7, det vill säga att det inte går att särskilja bokningsspecifika detaljer såsom avbokningsskydd och matpaket.

2.2 Beräkning av prisindex

Prisenheten avser att beräkna index enligt standardmetod, d.v.s. ett viktat geometriskt index där vikterna är proportionella mot total omsättning under föregående helår. De enskilda produkterna (produkt erbjudanden) kommer att definieras enligt nivå 7. Företagsurvalet är oförändrat, och täcker mer än 80 procent av marknaden för försäljning av charterresor. Täckningen av företagets försäljning kommer med nya datamaterialet vara näst intill 100 procent, istället för den bråkdel som nuvarande urvalsundersökning täcker.

I och med en fast korg kommer inte nya resmål komma in i indexberäkningen under året. För att kunna inkludera nya resmål i beräkningen, skulle det krävas beräkning av enhetsvärden på en mindre detaljerad nivå än nivå 7 (eller en ny indexkonstruktion). Preliminära indexberäkningar baserade på enhetsvärden på en mindre detaljerad nivå visar snarare på mixproblem, än en ökning av täckning i korgen.

3 Järnvägsresor

För KPI samlas priser på järnvägsresor in för ett urval av flera operatörer och olika destinationer. Insamlingen är manuell, via hemsidor på internet och genom ”provbokningar” som sparas undan för att kunna verifiera mätningarna.

I samband med ny översyn av urvalet av tågresor har SCB kunnat få tillgång till veckovisa transaktionsdata från en operatör som kan ersätta manuell prisinsamling via internet. Vägningstalet för tågresor är runt 4 promille av korgen och priserna är inte särskilt volatila.

3.1 Befintliga prismätningar

För produktgruppen Järnvägsresor ingår flera tågoperatörer i urvalet och det nuvarande upplägget för undersökningen tillämpas sedan 2011 (Carlsson och Sandén, 2010). Prisinsamling görs manuellt via tågoperatörernas hemsidor. De priser som samlas in formar produkterbjudanden som har specificerats för att täcka olika varianter av avresedagar, varierande framförhållning i bokningstid, typ av resenär (vuxen/ungdom) för olika destinationer och bokningstyp (ombokningsbar/ej ombokningsbar). Genom ett antal sådana komponenter antas prisutvecklingen under hela månaden fångas. Tabell 1 exemplifierar hur ett bokningsschema kan se ut.

Tabell 1. Exempel på bokningsschema för järnvägsresor

Kategori	Avgångstid	Destination	Framförhållning vid bokning	Tågtyp	Bokningstyp
Vuxen	Måndag, morgon	Sthlm - Gbg	1 dag	Snabbtåg	Ombokningsbar
Vuxen	Tisdag, f.m.	Sthlm - Åre	21 dagar	Intercity	Ej ombokningsbar
Vuxen	Onsdag, e.m.	Uppsala - Gävle	Samma dag	Nattåg	Ej ombokningsbar
Ungdom	Fredag, e.m	Gbg - Skövde	3 dagar	Regional	Återbetalningsbar

Notering: Med "Framförhållning vid bokning" avses hur lång tid innan resan ska provbokas, alltså tidpunkten för datafångst. Denna konstanthålls i teorin men rubbas ibland av SCBs faktiska arbetstider/röda dagar.

3.2 Nya prismätningar

Det nya dataunderlaget innehåller daglig omsättning över alla sålda resor (från första bokningen till sista bokning innan avgång), per destination, med avgång den specifika dagen. Nedbrytningen är efter typ av resenär (vuxen/ungdom) och avgränsas till andra klass och endast privatköp, alltså ej första klass eller företagsköp. Med aggregering menas här att det är summan av all försäljning över alla avgångar under ett dygn, mellan varje par av destinationer och över alla tågtyper, i de få fall där dessa varierar. Materialet är inte nedbrutet på enskilda avgångar (alltså klockslag), bokningstyp (ombokningsbara/ej ombokningsbara biljetter) eller tågtyp (snabbtåg, intercity, nattåg, regional).

Genom underlaget fångas ett genomsnittspris per dag/ vecka eftersom det avser all omsättning över alla avgående tåg en viss dag, vilket motsvarar ett enhetsvärdesindex där tjänsten, tågresan, i princip hålls konstant.

3.3 Fördelar med den nya datakälla

SCB har inte tillgång till tidigare datauppgifter för den nya datakällan utan det underlag som hittills tagits emot under hösten 2018 visar på prisnivåer enligt förväntan och SCB bedömer att byte av insamlingsmetod blir positivt för undersökningen. En direkt nytta av den nya

datakällan jämte övrigt underlag är att urvalet av destinationer hos denna aktör kan ökas till en cut-off täckande minst 60 procent av omsättningen för sålda resor

4 Flygresor

Sedan 2012, efter beslut i nämndens sammanträde (möte nr. 244) har prisinsamling av utrikes flygresor gjorts manuellt genom provbokning av resor via internet (se Andersson, Hulth, och Nilsson, 2011).

Generellt fluktuerar priserna för flygresor som ett resultat av olika faktorer, bland annat avresetid, säsong/kalendereffekter, reslängd och geografiskt avstånd. Med anledning av detta samlas priserna in manuellt enligt ett schema täckande olika tillfällen för att avspegla faktorerna ovan. Detta innebär en ganska stor manuell insamling och omfattande granskning/verifiering som kontroll av insamlingen.

Vägningstalen för flygresor i KPI är runt 9 promille av korgen år 2018, där priserna kännetecknas av stor variation, vilket innebär att prisförändringarna ofta får en stor påverkan på KPI totalt.

Under året har SCB fått tillgång till ett verktyg från en stor internetsida för resor vilket innebär en automatisk datafångstprocess, ett s.k. API (application programming interface - applikationsprogrammeringsgränssnitt). Detta innebär en signifikant ändring av insamlingsmetod, från manuell till digital. Urvalet kan ökas och insamlingsarbetet effektiviseras.

4.1 Undersökningen i nuläget

Idag samlas priser in för 13 destinationer. Varje månad provbokas 78 resor (6 per destination). Resorna är specificerade enligt destination, bortavaro (BV), tid för utresa, byten samt maxtid för resan, se tabell 2. För varje resetillfälle registreras de 5 lägsta priserna. Varje destination bokas två månader i förväg (m-2), en månad i förväg (m-1) och i samma månad (m-0) som avresan. Mättillfällena inträffar två gånger i månaden (T1 och T2).

Tabell 2. Exempel på bokningsschema för flygresor

Resdag	Resa	Framförhållning vid bokning	MT	BV	Starttid	Byte	Maxtid
1:a onsdag	STO ARN – CHI KIV	m-2	T1	4	06 - 12	2	20
2:a söndag	STO ARN – CHI KIV	m-2	T1	4	12 - 18	2	20
2:a fredag	STO ARN – CHI KIV	m-1	T2	7	06 - 12	1	12
3:e torsdag	STO ARN – CHI KIV	m-1	T2	4	12 - 18	1	12
4:e måndag	STO ARN – CHI KIV	m-0	T1	4	12 - 18	0	7
1:a söndag	STO ARN – CHI KIV	m-0	T1	4	18 - 24	0	7

4.2 Ny datainsamlingsmetod

Genom den nya datafångsmetoden (API) kommer det befintliga undersökningsupplägget kunna behållas. Resorna kommer att vara specificerade på samma sätt som nuvarande undersökning. Ändringarna som kommer införas från indexåret 2019 är byte av prisinsamling från manuell till digitalt, utökning av urvalsdestinationer (från 13 till cirka 50)

samt utökning av mättillfällena (från 2 till ca 10). En begränsning med API kan vara att en specifik resa inte är sökt vilket hanteras genom en sökrobot som istället gör de sökningar som är av intresse enligt bokningsschemat.

5 Referenser

Andersson, C., Hulth, A. och Nilsson, P. (2011). *Översyn av index för utrikes flygresor*. PM till nämnden för Konsumentprisindex, möte nr. 244.

Carlsson, E., Sandén, B. (2010). *Översyn av prisindex för transporttjänster i KPI*. PM till nämnden för Konsumentprisindex, möte nr. 242.

Johansson, J. och Ståhl, O. (2017). *Hantering av säsongprodukter i KPI*. PM till nämnden för Konsumentprisindex, möte nr. 3.

Norberg, A., Strandberg, K. och Ståhl, O. (2018). *En modell för resursallokering inom KPI*. PM till nämnden för Konsumentprisindex, möte nr. 4.