

# Redesign av butiksurval för KPI:s lokala prisundersökningar

*För beslut*

---

Prisenheten föreslår att urvaldesignen för butiksurvalet för KPI:s lokala prisundersökningar på ett kostnadseffektivt sätt tar hänsyn till geografiska aspekter.

Förslaget föreslås verkställas inför 2016, dels genom att införa cut-off på kommunnivå motsvarande 5 alternativt 10 procent av total handelsomsättning i riket, dels genom dragning av butiksurval i två steg med urval av regioner i ett första steg.

---

## INNEHÅLL

<b>REDESIGN AV BUTIKSURVAL FÖR KPI:S LOKALA PRISUNDERSÖKNINGAR</b> .....	<b>1</b>
<b>1 BAKGRUND</b> .....	<b>2</b>
1.1 PRISINSAMLING .....	2
1.2 NUVARANDE BUTIKSURVAL FÖR KPI:S LOKALPRISMÄTNING .....	3
<b>2 SYFTE</b> .....	<b>4</b>
2.1 EFFEKTMÅL .....	4
2.2 DELMÅL .....	4
<b>3 RIKTMÄRKEN</b> .....	<b>5</b>
3.1 COMPENDIUM OF HICP REFERENCE DOCUMENTS .....	5
3.1.1 <i>Förordningar av särskilt intresse i kompendiet</i> .....	5
3.1.2 <i>Rekommendationer</i> .....	5
3.2 CONSUMER PRICE INDEX MANUAL: THEORY AND PRACTICE .....	6
3.3 ANDRA LÄNDER .....	7
<b>4 UTREDNINGSMETOD</b> .....	<b>8</b>
<b>5 RESULTAT</b> .....	<b>9</b>
5.1 ANALYS AV RESKOSTNADER .....	9
5.1.1 <i>En iakttagelse</i> .....	9
5.1.2 <i>Regionala skillnader – kanske inte som väntat</i> .....	9
5.1.3 <i>Skillnader beroende på antal butiker per intervjuare</i> .....	10
5.2 VARIANSSTRUKTUR FÖR PRISFÖRÄNDRINGAR .....	11
5.3 STATISTIK OM URVALSRAMEN .....	13
<b>6 FÖRSLAG</b> .....	<b>15</b>



6.1	CUT-OFF AV KOMMUNER .....	15
6.2	REGIONURVAL BASERAT PÅ KOMMUNER OCH POSTNUMMER.....	15
6.3	SAMU-ROTATION.....	16
6.4	TRÖGHETSREGEL OCH BALANSERING.....	16
<b>7</b>	<b>DISKUSSION .....</b>	<b>16</b>
7.1	TÄCKNING.....	16
7.2	REGIONURVAL .....	17
7.3	EFFEKT .....	17
7.4	FAKTISKA URVALSSANNOLIKHETER .....	17
7.5	"TRÖGHETSREGELN" .....	20
7.6	BALANSERING.....	20
7.7	AVGRÄNSNING.....	21
<b>8</b>	<b>REFERENSER .....</b>	<b>21</b>
<b>9</b>	<b>APPENDIX .....</b>	<b>23</b>
9.1	A1: REKOMMENDERAT INSAMLINGSSÄTT PER BRANSCH .....	23
9.2	A2: VARIANSANALYS.....	25
9.3	A3: BUTIKSVIKTEN I URVALSSTRATUM .....	26

## 1 Bakgrund

### 1.1 Prisinsamling

Prisinsamling i butiker, restauranger, frisörer och andra tjänsteställen utförs av SCB:s intervjuare. Insamlingen sker endera vid besök varje månad eller via telefonsamtal. Det finns framtaget av enheten för prisstatistik, härefter prisenheten, en lista som för varje bransch anger rekommenderad insamlingsmetod, se appendix 1. När intervjuaren anser det nödvändigt kan besök genomföras även om telefonmätning rekommenderats i första hand.

Manuell prisinsamling vid besök innebär en betydande kostnad. Resekostnader är en väsentlig kostnadspost. Det finns ett intresse av att prisinsamlingen skall vara effektiv i den meningen att större delen av intervjuarnas tid skall åtgå för faktisk observation, inklusive urval av observationsobjekt och i förekommande fall kvalitetsvärdering. Som ett led i denna effektivisering och förbättring av prisinsamlingen tar SCB nu fram ett nytt verktyg för prisinsamling som avser ersätta den nuvarande handdatoren. Det nya verktyget för prisinsamling är en helhetslösning i form av plattformsoberoende webbaserad mjukvara med gränssnitt för prisinsamling i kombination med en surfplatta med egen internetaccess. Det nya insamlingsverktyget kommer att rullas ut gradvis inför 2016 för att säkerställa övergången från det nuvarande insamlingsverktyget.

Manuell prisinsamling har redan reducerats för mätningar i dagligvaruhandeln genom användandet av kassaregisterdata.

Totalt utfördes prismätning i 681 butiker av 87 intervjuare i februari 2015, av vilka omkring 80 intervjuare fysiskt besökte butik och resterande intervjuare hade enbart insamling via telefonintervjuer. Antalet butiker per intervjuare som skall besökas varierar mellan 1 och 10. Därutöver besöks även ett antal butiker för att det bedöms mer lämpligt med besöksintervju än att be om uppgifterna per telefon. Priser insamlas för bensinstationer av telefongruppen i Örebro och för internetbutiker via webb av prisenheten i Stockholm.

Tabell 1 Fördelning av antalet butiker som skall besökas per intervjuare i februari 2015

Antal besöksbutiker per intervjuare	Antal intervjuare
1	6
2	13
3	14
4	13
5	9
6	13
7	8
8	5
9	4
10	2
<b>Totalt antal intervjuare</b>	<b>87</b>

## 1.2 Nuvarande butiksurval för KPIs lokalprismätning

Urvalet av butiker för prisinsamling lokalt av intervjuare i butik dras enligt SAMU som är SCB:s centrala metod, system, verktyg och register för SAMordnade Urval inom företagsstatistiken (Ohlsson 1995). SAMU-register framställs fyra gånger om året. Registren är ”frysningar” av SCB:s företagsregister FDB, eftersom FDB ändras fortlöpande. Den ram som används för butiksurval etableras i maj, som är en av fyra SAMU-ramtillfällen under året. Med SAMU erhålls exempelvis negativt samordnade urval för olika grupper av undersökningar, vilket är bra ur uppgiftslämnarsynpunkt då det sprider bördan. Urvalen är positivt samordnade över tiden med långsam rotation över tid vilket balanserar krav för effektiva förändringsskattningar mot spridning av uppgiftslämnararbetet.

Ramen rensas från butiker som ligger mycket långt från närmaste intervjuare. Butiker i kommunerna 2421, 2425, 2462, 2463 och 2506, 2510, 2521, 2523, 2422, 2505, samtliga i Norrlands inland, exkluderas. Dessa utgör 0,76% av antalet butiker i landet och har 0,57% av handelns omsättning, enligt information i urvalsramen. I stället vägs storleken per butik upp i övriga kommuner i länen 24 och 25, Västerbotten och Norrbotten, exklusive residenskommunerna 2480 och 2580, så att det sammanlagda storleksmättet för dessa län blir oförändrat.

Ramen över försäljningsställen stratifieras efter SNI-bransch, men ej efter storlek, region, eller kedjetillhörighet. Vissa SNI-branscher sammanslås, t.ex. alla typer av klädbutiker, SNI-branscher 47711 – 47714. Specialiserad butikshandel med virke och byggvaror (SNI = 47521) slås ihop med Specialiserad butikshandel med järn- och VVS-varor (SNI = 47522), även Restauranger (SNI = 56100) och barer (SNI = 56300) slås ihop.

Storleksmättet vid urvalsdragning med varierande sannolikheter (presenterat på möte 206 i nämnden för KPI) är en sammanvägning av till lika delar antal anställda plus ett och en omsättningsuppgift baserad på momsuppgifter till Skatteverket. För flerbutiksföretag har omsättningsuppgiften för företaget fördelats på arbetsställen i proportion till antalet anställda per arbetsställe.

Urvalet allokeras per stratum med hänsyn till mellanbutiksvarians, vägningstal och kostnader för insamling, allt beräknat årligen baserat på innevarande undersökningsomgång, så som tidigare har dragits i nämnden för KPI. Kostnaderna varierar väsentligt mellan butiker som fordrar månatliga besök och butiker i branscher där telefoninsamling rekommenderas, vilket beaktas i allokeringen av urvalet.

Metoden för urval av försäljningsställen är s.k. roterat, stratifierat, sekventiellt Poisson-urval med urvalssannolikheter proportionella mot storleken av respektive försäljningsställe (Ohlsson 1990). Butikerna per stratum (bransch) rangordnas efter  $X_i/S_i$  i fallande ordning,

där  $X_i$  är det permanenta slumptalet (PRN), ett likformigt slumpstal, och  $S_i$  är butikens storlek. Med enbart detta förfarande finns det inga objekt som med säkerhet får så små kvoter att de hamnar i toppen av listan och väljs med säkerhet. Det är bara för att vi i förväg redan har identifierat objekten vars  $n \cdot S_i / \sum S_j > 1$  som vi har objekt vi påstår är valda med säkerhet.

Två avvikelser mot principen om proportionalitet mot storlek görs.

- 1) **Tröghetsregeln.** När butiker ändrar storlek mellan åren så hoppar de ut och in i urvalen eftersom urvalssannolikheten är proportionell mot storleken. Från och med 2009 års urval "favoriseras" butiker som redan är med i KPI:s prismätningar genom att storleksmättet för dem multipliceras med faktorn 2 och urvalssannolikheten, betingat föregående års urval, ökas. Ingen åtgärd görs vid estimationen, vilket principiellt innebär att en bias införs. Förfarandet innebär att objekt som minskat i storlek överrepresenteras och objekt som ökat i storlek underrepresenteras.
- 2) **Balansering.** Två butiker som är lika varandra i något visst avseende skulle kunna finnas samtidigt i urvalet ur ett branschstratum eftersom ingen stratifiering efter kedja, region eller dylikt görs. Detta förhindras aktivt på följande sätt i syfte att erhålla effektivare urval:
  - (a) I ett första steg sorteras hela urvalsramen enligt det generella förfarandet vid Poission-urval.
  - (b) I ett andra steg tilldelas butiker ett gruppnummer enligt en relevant gruppering (kommuner, dagligvarukedjor etc.). Sorteringen från första steget justeras därefter så att två butiker i samma grupp inte kan ligga för nära varandra i ordningen, men ändå ganska nära den plats i ordningen som de fick i första steget.

## 2 Syfte

### 2.1 Effektmål

Förslaget till re-design syftar till att effektivare utnyttja KPI:s resurser för att uppnå bra total statistisk kvalitet.

### 2.2 Delmål

Det är åtminstone följande aspekter att betrakta vid design av lokalprismätningarna.

1. Reskostnader kan minskas om urvalet av butiker kan begränsas till färre orter eftersom flera butiksbesök då skulle utföras i en följd under samma dag. Minskade resor kan åstadkommas genom:
  - a. att butiker på långa reseavstånd från närmaste intervjuare exkluderas från rampopulationen.
  - b. ett initialt urval av regioner, så pass litet att vanligtvis fler än en besöksbutik skall väljas per ort. Tvåstegsurval medför i sig alltid en viss ökning av samplingvariansen för samma storlek av det totala butiksuralet, men denna ökning bedöms bli mycket liten för KPI, på grundval av observerade data. För given kostnad bedöms variansen bli mindre om vunna resurser används för större urval av butiker.
  - c. att medvetet införa en begränsning av rampopulationen för urvalet av butiker, t.ex. att bara välja butiker i säg 20 större städer eller i säg 40 orter med viss storleksvariation där det bor en intervjuare. Denna typ av medveten cut-off medför en bias i skattningen av prisutvecklingen, medan samplingvariansen kanske blir mindre.

- d. att exempelvis urvalet av butiker fördelat per intervjuare sker till hälften bestämt i samband med urvalsdragningen, till hälften bestämt av intervjuarens egen bedömning. Detta lär kräva en utökad administrativ process kring urvalsarbetet.
2. Om KPI år efter år kan anlita samma intervjuare och därmed ge dessa en större arbetsuppgift vardera så kan kompetensutveckling av intervjuarna på sikt förbättra mätningarna. Därmed kunde eventuellt kvaliteten i intervjuarbetet bli bättre säkrad. Antalet intervjuare kan vara av storleksordningen 30-50 personer.

## 3 Riktmärken

### 3.1 Compendium of HICP reference documents

Eurostat har samlat förordningar (Legislation) och rekommendationer i kompendiet <http://ec.europa.eu/eurostat/documents/3859598/5926625/KS-RA-13-017-EN.PDF>

#### 3.1.1 Förordningar av särskilt intresse i kompendiet

**Avsnitt 1.1, *Framework regulation***, council regulation 2495/95 skrevs 1995 och uppdaterades 2003 och 2009. Article 3 nämner den geografiska täckningen:

The HICP shall be based on the prices of goods and services available for purchase in the economic territory of the Member State for the purposes of directly satisfying consumer needs. Questions concerning weighting shall be decided on by the Commission under the procedure laid down

**Avsnitt 1.2, *Initial implementing measures***, council regulation No 1749/96 skrevs 1996 och uppdaterades 1998 och 2007.

Article 8 under avsnittet "II. MEASURES FOR ENSURING COMPARABILITY, RELIABILITY AND RELEVANCE OF THE HICP" avser urvalsmetoder:

#### **"Minimum standards for sampling**

HICPs constructed from target samples which, for each category of COICOP/HICP and taking into account the weight of the category, have sufficient elementary aggregates to represent the diversity of items within the category and sufficient prices within each elementary aggregate to take account of the variation of price movements in the population shall be deemed reliable and comparable. "

**Avsnitt 1.17, *Sampling***, council regulation No 1334/2007, är från 14 November 2007. I följande stycken ställs allmänna krav på sampling:

"(9) It is therefore necessary to further specify the purposes and definition of the HICP, to clarify where those determine the actual practices of sampling, replacement, and quality adjustment, and to establish the required representation of the HICP and its form, and to establish further minimum standards with respect to sampling, replacement, quality adjustment and aggregation procedures.

(10) In particular, it is necessary to set a clear statistical target for the purposes of sampling, replacement and quality adjustment and ensure that the HICP measures close to the target, with a reasonably small uncertainty or error in terms of bias and variance. A trade-off between unbiasedness and precision must be considered."

#### 3.1.2 Rekommendationer

Det finns ingen förordningstext om samplingmetoder. Det finns i avsnitt **2.4 *Treatment of internet purchases***, ett underkapitel (3.6) om internethandel där det går att läsa att den

normala metoden, förutom för just internethandel, är att konstruera en urvalsram och att använda en ”proper sampling method”:

” A related challenge to be solved with the proposed Recommendations is the sampling of outlets (sellers). Normal procedure would be to construct a target universe of outlets and then use a proper sampling method. With internet purchases the target universe of sellers is in many cases unknown to the Statistical Office. Hence, more subjective sampling methods have to be used, based on information from e-commerce surveys or branch organisations with knowledge in the field of e-commerce.”

### 3.2 Consumer price index manual: Theory and practice

International Labour Office, ILO, har tillsammans med andra internationella organ givit ut en manual för konsumentprisindex, ursprungligen från 2004, som fungerar som vägledning i frågor rörande konsumetprisindex. Här är en länk:

[https://books.google.se/books?id=uYQeQH\\_hsFgC&pg=PA456&lpg=PA456&dq=hicp+manual&source=bl&ots=K1zGQ0Wj2i&sig=RIm8G-m8WZQw8ES87CSpljUM8B4&hl=sv&sa=X&ei=3UwZVaf-I4TYU-GOgvAN&ved=0CFQQ6AEwBg#v=onepage&q=hicp%20manual&f=true](https://books.google.se/books?id=uYQeQH_hsFgC&pg=PA456&lpg=PA456&dq=hicp+manual&source=bl&ots=K1zGQ0Wj2i&sig=RIm8G-m8WZQw8ES87CSpljUM8B4&hl=sv&sa=X&ei=3UwZVaf-I4TYU-GOgvAN&ved=0CFQQ6AEwBg#v=onepage&q=hicp%20manual&f=true)

Kapitel 5 i manualen avser sampling.

Här är avskrift av tre avsnitt som utgör rekommendation för urvalsmetoder:

“5.3 The universe usually has three dimensions. There is a product dimension, consisting of all purchased products and varieties of products. There is a geographical and outlet dimension consisting of all places and channels where a products is sold. Finally, there is a time dimension consisting of all sub-periods within an index period. The time dimension will be given less attention since price variation is usually smaller over a short time span and since temporal aspects may be dealt with in product and outlet specifications.”

“5.27 Modern statistical sampling theory focuses on probability sampling. Use of probability sampling is also strongly recommended and standard practice for all kinds of statistical surveys, including economic surveys. But price index in practice in most countries is still dominated by non-probability techniques. It may then be fruitful to speculate somewhat about the rational and irrational reasons for this situation. In the following section we discuss a number of such possible reasons, one by one. We then go on to consider some non-probability techniques.”

Motiv för icke-sannolikhetsmetoder som nämns är (förkortat och tolkat):

5.28 Det finns ingen urvalram (vanligt för produkter, för butiker finns vanligtvis ett register).

5.29 Bias kan bedömas vara försumbar (exempel från produktsidan).

5.30 Urvalet skall kunna vara observerbart framgent (man bör inte välja kortlivade produkter).

5.31 Ett urval för basperioden är inte ett urval för jämförelseperioden (produkter).

“5.32 *Price collection must take place where there are price collectors.* This argument applies to geographical sampling only. It is, of course, cheaper to collect prices near the homes of the price collectors, and it would be difficult and expensive to recruit and dismiss price collectors each time a new sample is drawn. This problem can be reduced by having good coverage of the country in terms of price collectors. One way to achieve this is to have a professional and geographically distributed interviewer organization within the national statistical agency, which works on many surveys at the same time.

Another way of reducing the problem is to have a first-stage sample of regions or cities or locations which changes only very slowly.”

5.33 Mycket små urval per stratum (kan t.ex. gälla urval av geografiska områden och med ”otur” i urvalet kan detta bli väldigt ”snett”).

5.34 Urvalet skall göras exempelvis när prisinsamlarna väljer produktbjudanden i butik.

### 3.3 Andra länder

Eurostat har sammanställt information medlemsländernas metoder i olika avseenden, bl.a. urval av butiker och andra försäljningsställen. Informationen finns här:

<http://ec.europa.eu/eurostat/web/hicp/methodology/metadata-and-national-practices>

Informationen är många gånger kortfattad. Det går att utläsa att vissa länder har en design för urval som innebär att man i ett första steg gör geografiska urval. Urvalsmetoderna i detta steg är sparsamt beskrivna. Nedan presenteras alla intressanta texter.

Som kan noteras i följande förteckning, verkar flera KPI i flera länder mätas med geografisk hänsyn.

**Belgien:** “Outlets are selected in 65 localities, which have been identified as being important also for the surrounding area. Outlets are in general selected by price collectors based on their experience and knowledge of local shopping habits.”

**Estland:** “There are around 1800 selected outlets including services. Outlet selection is purposive and sample is updated annually. The selected outlet must have significant turnover and availability of selected products/services. The general strategy is to select typical outlets.”

**Finland:** “For consumer durables and services, the price collectors select the outlets based on their local knowledge and potentially some turnover information (purposive selection). The sample includes around 2 430 outlets.”

På Finlands hemsida finns mer information avseende dessa produkter: ”It has not been possible to use a statistical sampling method for selecting specialised stores for the price collection, so Statistics Finland’s interviewers have chosen suitable outlets from their area according to specified criteria and by drawing on their knowledge of the area. The sample comprises good 2,400 outlets from 75 municipalities.” Finland har 320 ”municipalities”. Enligt mailväxling med statistikbyrån är det 20 stora städer samt ungefär 50 övriga kommuncentra.

**Frankrike:** “Outlets outside conurbations are not covered by the sampling procedure.”

I följande dokumentet på franska statistiskbyråns hemsida

[http://www.insee.fr/en/indicateurs/ind29/IPC\\_m-EN.pdf](http://www.insee.fr/en/indicateurs/ind29/IPC_m-EN.pdf)

läser man ”Sampling frame and number of units surveyed:

The sampling plan is stratified according to three criteria:

- geographic criterion: prices are collected in 106 urban units of more than 2,000 inhabitants throughout France and of all sizes and in four overseas departments
- product types: a sample of slightly over 1,000 goods and services families, called "items" (variétés) is defined to allow for product heterogeneity within 161 groups of products. ...
- "type of sales outlet": a sample of 27,000 outlets, stratified by sales channel, was constructed to represent goods and services diversity by brand, distributor, and consumer-

purchasing modes, and to take into account the differences in price movements for each sales channel.”

**Italien:** “. . .the survey at territorial level is carried in 84 municipalities (out of a total of 110) by Municipal Offices of Statistics (MOS).”

**Norge:** “The outlets/firms are selected from Statistics Norway’s Business Register. The selection is made after stratifying the population by industry and region. The probability to be selected is proportional to the size of the turnover.”

**Storbritannien:** “Shopping locations are statistically selected from across the UK using a probability proportional to size (PPS) sampling frame, stratified by region. Outlets within locations are first classified by shop type and then sampled using a PPS methodology.”

På ONS hemsida hittar man ytterligare information: “Prices are collected in around 150 locations across the 12 government regions within the UK. Location selection takes place separately within each region, with the probability of a particular location being selected proportional to the number of employees in the retail sector in that location.”

Within a location, outlets are selected randomly.”

**Tyskland:** “Approximately 300 000 price observations for approximately 700 commodities are collected from approximately 30 000 retail outlets in 188 municipalities throughout Germany on a monthly basis.” Antalet municipalities, eller s.k. “Gemeinden”, uppgår till cirka 12 000 för år 2015 (se t.ex. Wikipedia eller Destatis.de).

**Österrike:** “About 4 200 outlets in the country are chosen for the survey. The outlets are chosen by the regional offices as being representative of the particular product groups and of the places where households make their purchases. Outlets may be chosen outside the administrative border of the cities, e.g. shopping centres, if they are popular.”

**USA:** (hemsida) “The smallest geographic areas in which pricing is done for the CPI are called primary sampling units (PSUs). Within these areas, sales outlets are chosen where people shop and live. . . . For the purpose of selecting the 1998 CPI PSU sample, the entire United States was divided into PSUs. First, BLS used the U.S. Office of Management and Budget’s (OMB’s) definition of Metropolitan Areas (MAs) to divide the country into metropolitan and nonmetropolitan areas. The PSUs within the metropolitan area are, with five exceptions, OMB-defined MAs. In the nonmetropolitan areas, BLS defined the PSU boundaries. In general, a PSU is delineated by county borders (with some exceptions in New England) and can comprise several counties. . . . select one PSU per stratum so that the selected PSUs are well-distributed over the States”

Sammanfattningsvis är det vanligt att statistikbyråer som skapar butiksurvalet med ett första steg av urval av regioner. Storbritannien gör detta med sannolikhetsurval, för övriga är det oklart hur det görs eller så är det uppenbart att det är medvetna urval alternativt cut-off-urval.

## 4 Utredningsmetod

KPIs intervjuarinsamlade data för fyra år, 2011-2014 kommer att analyseras dels genom deskriptiv analys av kostnader för lokala prisinsamlingen, dels genom variansanalys av prisförändringar. Fokus kommer ligga på antal butiker per intervjuare, butikernas geografiska spridning och restiden. Som alltid får man vara försiktig med generaliserbarheten av ett empiriskt resultat baserat på ett begränsat antal år.



## 5 Resultat

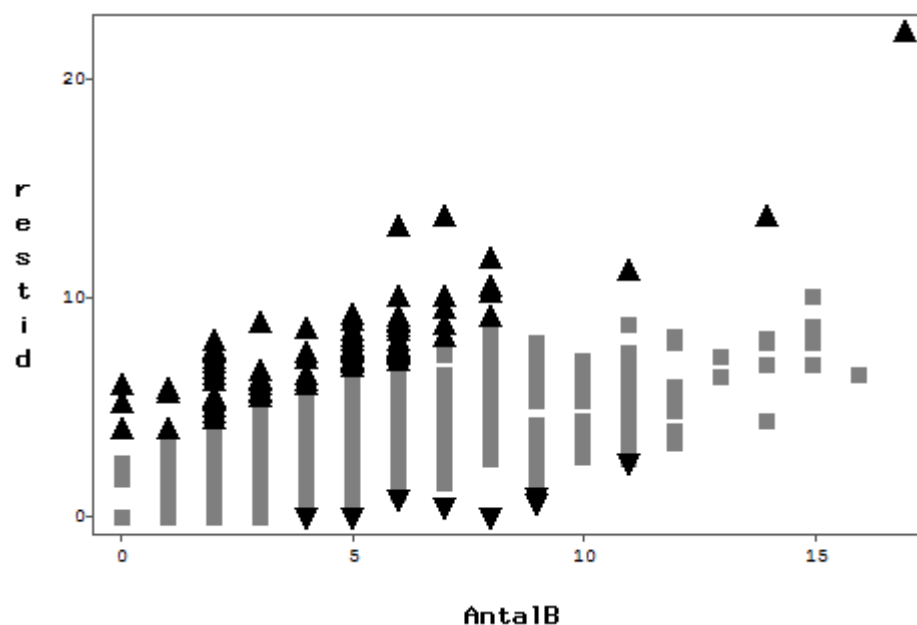
### 5.1 Analys av reskostnader

Bokföringen av intervjuarnas restid avser hela månaden utan fördelning över deras butiker, vilket försvårar entydiga analyser.

#### 5.1.1 En iakttagelse

I diagram 1 kan ses att några intervjuare bokför restid trots att prisinsamling borde kunna göras med telefonsamtal. Det är ett antal fall där intervjuare har besökt butiker mer än en gång per år fast telefoninsamling rekommenderas för alla tilldelade butiker (AntalB=0), men detta har oftast praktiska förklaringar.

Diagram 1 Total restid mot antal butiker som skall besökas. Per intervjuare 2014 och 2015. Den horisontella axeln är antal butiker och den vertikala axeln är sammanlagd restid per intervjuare.



#### 5.1.2 Regionala skillnader – kanske inte som väntat

Diagrammen 2 och 3 visar sambandet mellan total restid och antal butiker som intervjuaren rekommenderas besöka. Om man hade en hypotes om att restiderna är långa i Norrland så kan den förkastas – kostnadsdata visar inte på sådana samband mellan regioner och reskostnader. Debitering av restider ser ut att vara ungefär lika hög i regionen Stockholms, Uppsala och Gotlands län (diagram 2).

Diagram 2 Total restid mot antal butiker som skall besökas.  
 Per intervjuare 2014 (alla månader) och 2015 (januari och februari).  
 Region 1 = Stockholms, Uppsala och Gotlands län

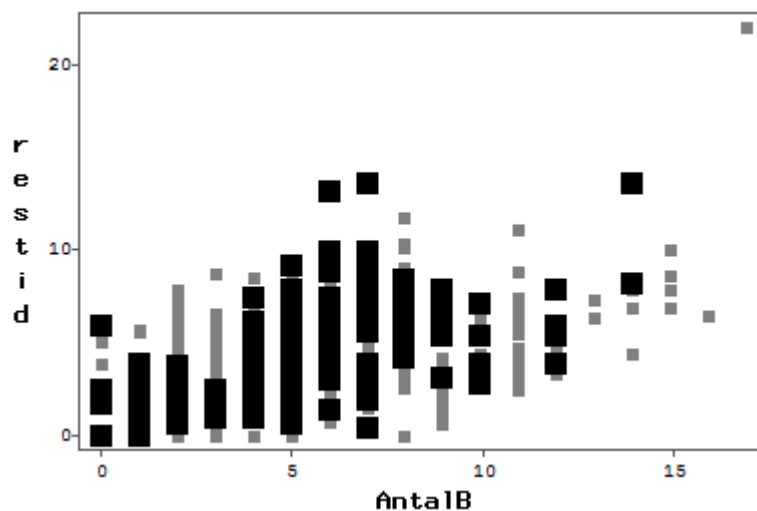
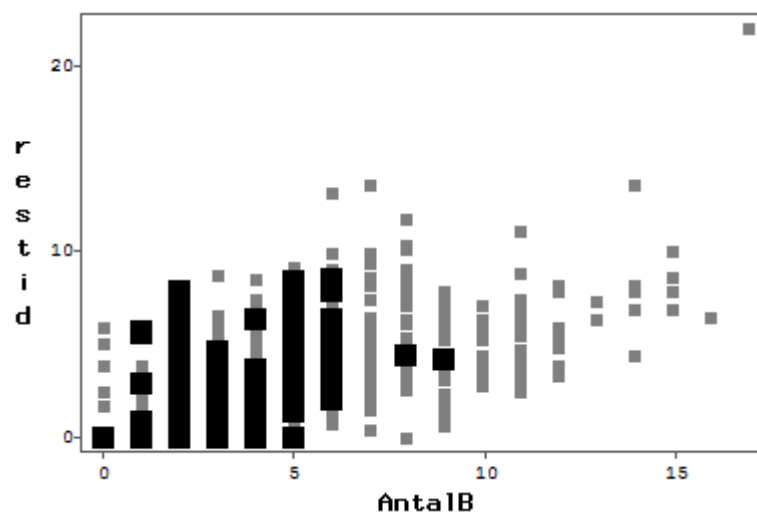


Diagram 3 Total restid mot antal butiker som skall besökas. Per intervjuare 2014 och 2015.  
 Region 14 = Dalarnas, Gävleborgs, Västernorrlands, Jämtlands, Västerbottens och Norrbottens län.

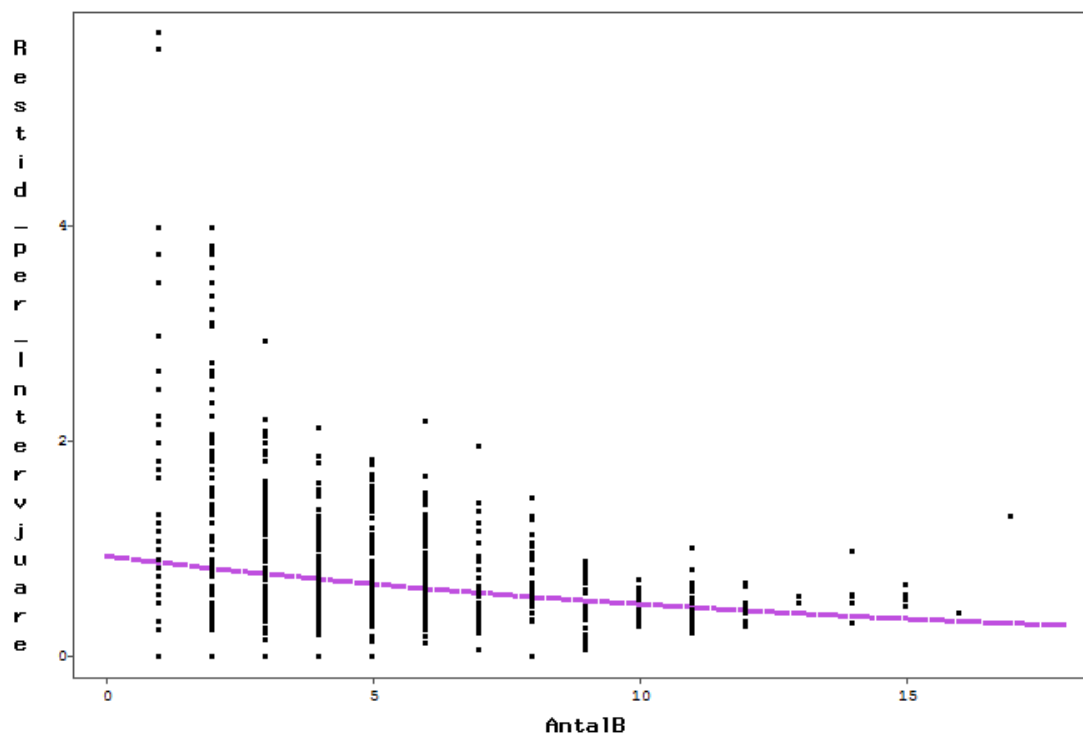


### 5.1.3 Skillnader beroende på antal butiker per intervjuare

Diagram 4 visar att restid per butik för vilken besök rekommenderats (enligt de praktiska riktlinjer som används i KPI) tenderar att vara högre för intervjuare med få tilldelade butiker än för intervjuare med stor tilldelning. Det finns även ett samband med antal butiker som skall kontaktas med telefon. Förklaringen är helt hypotetisk. Det förefaller tänkbart att en intervjuare med 10 tilldelade butiker kan planera resorna så att flera butiker besöks på samma resa än om intervjuaren bara har två butiker.

Med anpassning av en exponentiell modell får man skattningen att intervjuare som har två besöksbutiker använder i genomsnitt 50 minuter restid per butik medan intervjuare som har 8 butiker använder i genomsnitt 34 minuter restid och intervjuare som har 14 butiker använder i genomsnitt 23 minuter restid.

Diagram 4 Total restid dividerat med antal butiker för vilka besök anbefalles mot antal butiker. Per intervjuare 2014 och 2015.



## 5.2 Variansstruktur för prisförändringar

Finns det skillnader i prisutveckling mellan orter av olika storlek? För att belysa detta har vi utfört variansanalyser.

I produktdimensionen finns cirka 360 s.k. produktgrupper (consumption segments) totalt. Intervjuarna samlar in priser för ett urval av produkter inom respektive produktgrupp. De flesta produktgrupperna har endast en vald produkt, även kallad representantvara. Motordrivna trädgårdsverktyg representeras med två produkter med de generiska beskrivningarna,

- ”Gräsklippare, bensin. Motor 3,5 – 5 hk el motsvara 2573 – 3676 w, klippbredd 42-48 cm. Utan uppsamlare. Får ha utkastare. Får vara senaste pris vintertid”.
- ”Gräsklippare, bensin. Motor 3,5 – 5 hk el motsvara 2573 – 3676 w, klippbredd 42-53 cm. Med uppsamlare. Får ha utkastare. Får vara senaste pris vintertid”.

Urvalet i produktdimensionen skulle i detta fall med god vilja kunna sägas vara stratifierat med urvalsstorleken  $n = 2$  där motordrivna trädgårdsverktyg är stratumet. Fast mer rätt är att beskriva de valda motorgräsklipparna som ett helt enkelt medvetet urval bland alla möjliga gräsklippare, motordrivna häcksaxar, jordberedningsmaskiner, motorsågar m.m. som finns och som motsvarar vägningstalet för produktgruppen motordrivna trädgårdsverktyg. Det är möjligen ett storleksurval.

I butiksdimensionen har KPI en urvalsmetod som levererar nästan läroboksmässigt korrekta urval (sånär som på cut-off av långväga butiker, intervention för att öka den positiva samordningen mellan år och justering för att uppnå balans i urvalen).

I det slutliga urvalssteget väljer intervjuaren ett produkterbudande. Detta görs dels när en butik är ny i urvalet och dels vid varje nödvändig ersättning av ej längre sålda produkter. Det finns instruktioner för denna process, med representativitet (mest sålda) och jämförbarhet

(lika som föregående vid ersättning) som önskade egenskaper för valda produkterbjudanden. Det är dock inte så långt ifrån en korrekt beskrivning att påstå att urvalet är godtyckligt, om än inte riktigt slumpmässigt.

För ett antal s.k. produktområden kan vi formulera en modell för logaritmerad priskvot mellan december år  $y-1$  och december år  $y$ . Det är inte lika intressant att få en statistisk belysning av förekomst av butikseffekt som det är att se om det finns någon regional effekt, enligt någon regional indelning som är av intresse vid design av butiksurvalen. Därför tillämpas modellen:

$$\ln(\text{priskvot}_{p,b,e}) = \text{Produktgrupp}_p + \text{Region}_r + \varepsilon_e$$

där  $\ln(\text{priskvot}_{p,b,e})$  är logaritmerad priskvot

Produktgrupp<sub>p</sub> = genomsnittlig priskvot för produktgrupp p

Region<sub>r</sub> = genomsnittlig priskvot för regionen r

och  $\varepsilon_e$  = en ”slumpeffekt”, vilken uppstår av intervjuarens val av produkterbjudande.

Vi har testat förekomsten av skillnad mellan regioner med definitionen av region som 1 = kommuner med slutsiffra 80 eller 81 som innebär stor kommun och 2 = övriga kommuner. Dessa två regioner är ungefär lika stora i termer av population och urval av butiker. För försäljningsställen som skall insamlas vid personliga besök av intervjuaren ser fördelningen praktiskt taget jämn ut.

Tabell 2. Fördelning i ramen av butiker över de två definierade regionerna, i procent

Region	Andel av antal butiker i hela rampopulationen	Andel av antal butiker i rampopulationen som skall besökas	Andel av storleken av butikerna i hela rampopulationen	Andel av storleken av butikerna i rampopulationen som skall besökas
1 = kommuner med slut-siffra 80 eller 81	46	46	55	51
2 = övriga kommuner	54	54	45	49
Totalt	100	100	100	100

Resultat per år för nio produktområden under åren 2011-2014 redovisas i appendix 2. Av 36 tester finns tre ”signifikanser” på 5-procentsnivån med tolkningen att det kan finnas en regional skillnad i prisutveckling. Med tanke på masssignifikansproblemet blir slutsatsen ändå att det för dessa fyra år är någon skillnad i prisutveckling mellan regioner för ingående produktområden knappt urskiljbar. I produktdimensionen förekommer många fler ”signifikanser”, vilket kan tolkas som att det faktiskt skulle finnas en skillnad i prisutveckling för olika produktgrupper inom t.ex. området Hemelektronik. Här finns produktgrupperna liten TV, stor TV, CD-kassettradio, stereopaketer, videokamera, DVD-spelare, hemmabiopaket, MP3-spelare, digitalkamera, spelkonsol och TV- & dataspel.

Liknande slutsats, d.v.s. att variationer i prisförändringskvoter inte låter sig förklaras av regionala eller butikseffekter och inte av produktgrupper utan framstår som vitt brus finns i tabell 9 på sidan 23 i Norberg (2004).

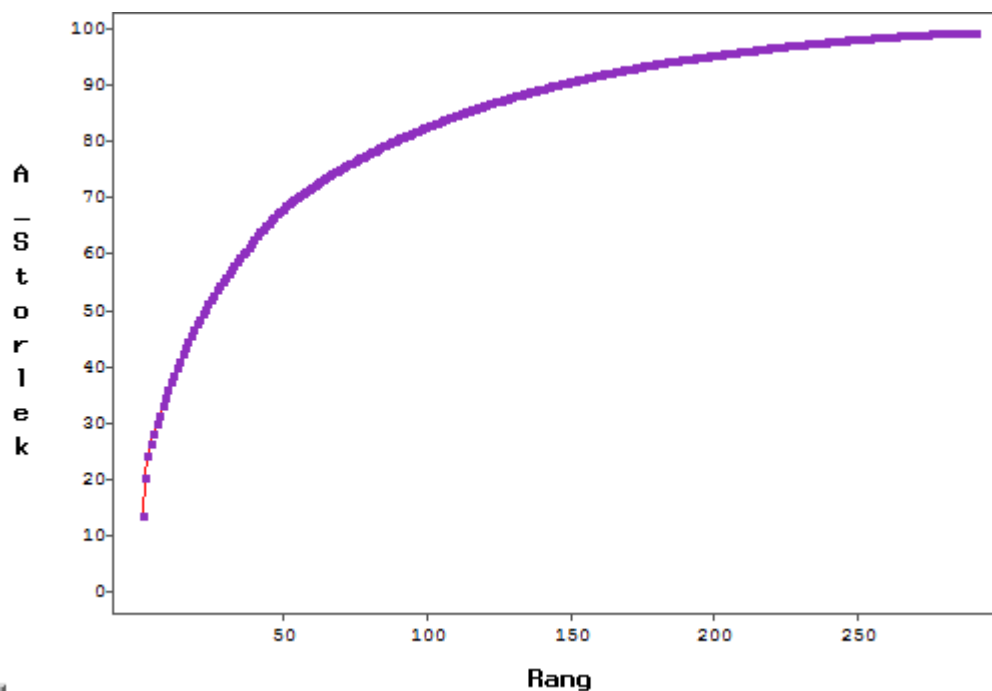
Slutsatsen av denna analys är att KPI skulle vinna på stora urval av produkterbjudanden där det är relativt sett billigt att registrera data, d.v.s. i små urval av försäljningsställen på korta reseavstånd för intervjuarna.

### 5.3 Statistik om urvalsramen

I detta avsnitt presenteras deskriptiv statistik över butikernas geografiska fördelning i landet. Syftet är främst att föreslå att den geografiska täckningen av landet kan minskas från 99,43 procent till en betydligt mindre andel, förslagsvis 90-95 procent.

Diagram 5 visar andelar av den totala storleken av rampopulation per kommun i fallande ordning. Stockholms kommun inleder grafen med drygt 13 procent av hela rampopulationen och 10 procent räknat på endast besöksbutiker. 23/26 kommuner når upp till 50 procent av storleken, 87/90 kommuner når 80 procent, 142/142 kommuner räcker för 90 procent och med 189/188 kommuner täcks 95 procent av storleken av hela rampopulationen respektive de försäljningsställen som skall besökas. Notera att med dagens design av KPI-urvalet så exkluderas endast 0,57 procent.

Diagram 5 Kommunernas kumulativa andel av total storlek av rampopulationen

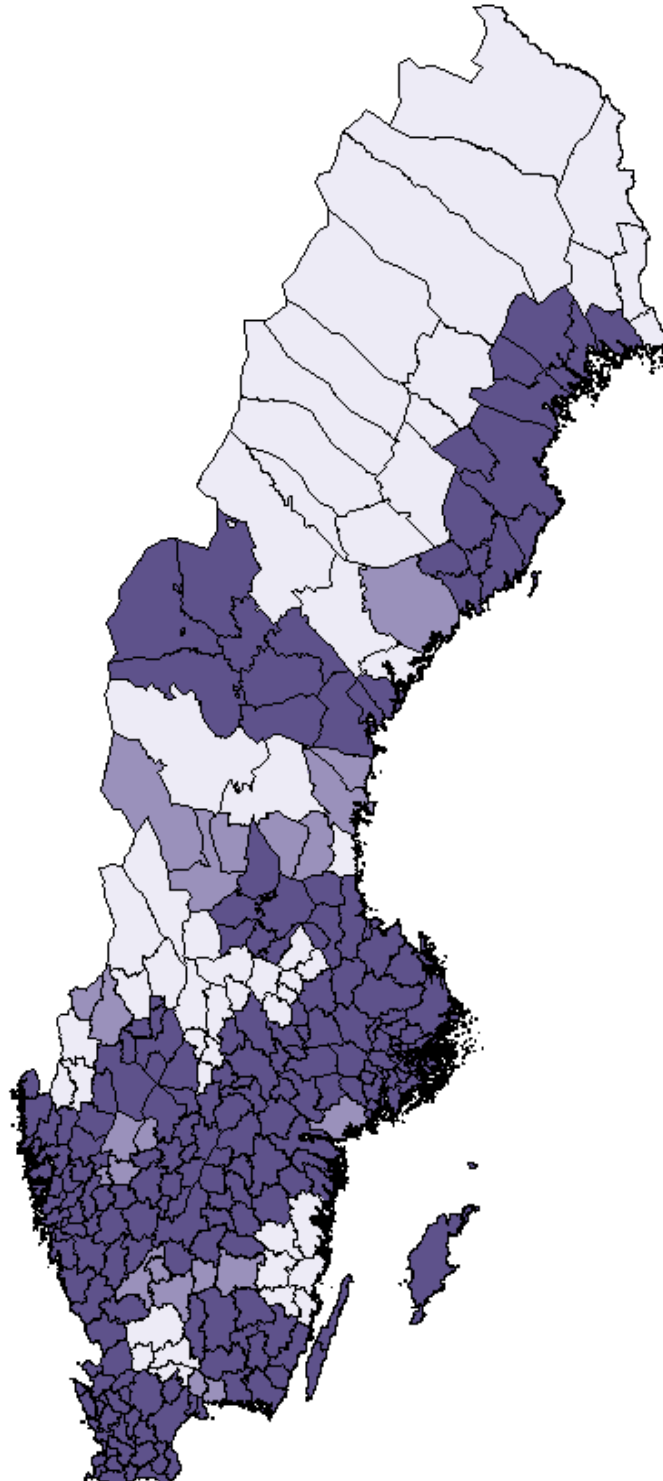


Vi ser emellertid att vissa ”små kommuner” inte är en bra avgränsning för att exkludera i urvalsramen då en sådan kommun faktiskt kan ligga nära en stor centralort på bekvämt avstånd. Den relevanta avgränsningen bör göras i samråd med SCB:s intervjuorganisation och med hänsyn till intervjuarnas adresser. Detta är inte utfört till föreliggande promemoria.

För att illustrera hur utfallet kan bli väljer vi i stället begreppet *lokala arbetsmarknader* (se källhänvisning till SCBs hemsida för indelningen) som vardera består av en eller flera kommuner. Det finns 73 sådana regioner för år 2013. Av dessa regioner består 25 av endast en kommun, ganska ofta är det så för ytmässigt stora inlandskommuner i Norrland med liten befolkning. I diagram 6 visas vilka kommuner som ingår i regioner som har 5 respektive 10 procent av all handel. Syftet med lokala arbetsmarknader är att kunna beskriva arbetsmarknadens funktionssätt för geografiska områden som är relativt oberoende av omvärlden med avseende på utbud och efterfrågan av arbetskraft. De lokala arbetsmarknaderna är således funktionella och kan förändras över tiden beroende på förändringar i pendlingsströmmarnas riktning och styrka. Definitionen är naturligtvis inte perfekt för föreliggande tillämpning men tillräckligt bra.

Av Sveriges kommuner ingår 50 i de minsta lokala arbetsmarknader som tillsammans har fem procent av handeln, enligt vårt storleksmått. Ytterligare 23 kommuner svarar för fem procent av handeln.

Diagram 6 Kommuner som tillsammans har 5 respektive 10 procent av all handel i Sverige. Ljus färg är de minsta fem procenten, mellanljus färg motsvarar också fem procent men är större. Mörk färg motsvarar resterande 90 procent av handeln.



## 6 Förslag

Följande metoder och omfattning föreslås införas inför 2016 i urvalsdragningen av butiker till KPI. För detaljutformning av regioner och bestämning av urvalsstorlekar skall SCB:s metodenheter utföra ytterligare metodstudier i vanlig ordning.

### 6.1 Cut-off av kommuner

Kommuner med liten handel som ligger på sådant avstånd från intervjuare att resekostnader för eventuella valda butiker blir stora skall exkluderas från urvalsramen. Kompensation för detta skall göras länsvis så att butiker i andra kommuner i länet, exklusive kommuner med slutsiffror 80 och 81, får större urvalssannolikheter. Totalt skall detta s.k. cut-off-stratum utgöra cirka 5 alternativt 10 procent av alla handel i landet.

### 6.2 Regionurval baserat på kommuner och postnummer

#### Steg 1

Ett första urvalssteg av regioner införs i urvalsdesignen. Riket, exklusive cut-off-kommunerna, delas i postnummerområden, definierade av postnummer. Det finns drygt 10 000 postnummerområden i Sverige. De enstaka butiker som inte har postnummer eller har postnummer i annan kommun än vad kommunangivelsen för företagsregistret anger får ett imputerat postnummer som är medianen av postnummer i respektive kommun. Vidare delas ett eventuellt postnummerområde som sträcker sig in i två kommuner även efter kommun.

Butikernas storleksmått (se avsnitt 1.2 för en förklaring) summeras per postnummerområde, oavsett bransch. Som storleksmåtten för postnummerområden används roten ur den framräknade storleken. Ett systematiskt  $\pi$ ps-urval av 800 postnummerområden väljs efter att ramen sorterats efter kommun och postnummer. Motivet att använda roten ur storleken är att ge mindre områden större urvalssannolikhet än om storleken används rakt av. Detta har visat sig vara fördelaktigt då stora butiker i små områden annars inte får en så stor urvalssannolikhet som vi önskar.

För att bibehålla rotationsegenskapen i SAMU föreslås att riket delas in i fem lika stora rotationsgrupper efter län och kommun, var rotation i varje grupp tillåts vart femte år. För varje år byts det systematiska  $\pi$ ps-urvalet av postnummerområden ut helt, för en av de fem grupperna. Sannolikheten är stor att stora postnummerområden kommer att ingå i flera femårsperioder i följd.

#### Steg 2

I ett andra urvalssteg väljs butiker med sekventiellt Poisson-urval med urvalssannolikheter proportionella mot storlekmultiplicerat med inverterad urvalsannolikhet från steg 1. Till skillnad från hittills sorteras inte butiker som i princip skulle ha urvalssannolikhet 1,0 ut i förväg. Poisson-urval ger därmed inte dessa stora butiker en faktisk urvalssannolikhet = 1,0. I praktiken gäller detta endast ett litet antal branscher, bl.a. musikaffärer.

En simulering med 4 000 urvalsdragningar visar att denna metod fungerar utmärkt för branscher där många butiker ligger i "handelsområden", som därför är "stora".

I mycket små postnummerområden blir urvalssannolikheten för enskilda butiker mindre än den önskvärda, men dessa svarar för en mindre andel av total handel. I diagrammen 8.1-8.5 används färger från violett till rött enligt regnbågens ordning för att demonstrera postnummerområdenas storlek från minsta till största.

Ytterligare metodstudier kan förfina metoden. Verkligt stora butiker som bör vara med i urvalet med säkerhet, år efter år, kan ”handplockas”. De minsta postnummerområdena kan uteslutas, likt cut-off ovan och av samma skäl, nämligen att där finns kanske bara en butik.

### 6.3 SAMU-rotation

Det skulle innebära önskat hög grad av urvalsbyten varje år om dels urvalet av en femtedel av postnummer byts enligt steg 1 och dels cirka 20 procent av övriga fyra femtedelar av butiksurlvalet roteras enligt SAMU. Vi föreslår därför följande modifierade rotation av butiker.

- I de fyra femtedelar av urvalet av postnummerområden som bibehålls från förra året ändras inte de permanenta slumpalen (PRN) så som SAMU gör utan de behålls. Nya butiker i urvalsramen tilldelas PRN från SAMU.
- I den femtedel av postnummerområden som är nyvald används butikernas aktuella PRN från SAMU-registret.

Detta betyder för butiker som lever i många år att KPI skapar PRN-talen i samband med att nya postnummerområden väljs, d.v.s. vart femte år. Den ”totala” avsiktliga rotationen bör bli cirka 20 procent, som hittills.

### 6.4 Tröghetsregel och balansering

Tröghetsregeln har liten effekt men skadar inte, den behålls därför.

Balansering inverkar inte menligt på urvalssannolikheterna och behålls. Villkoren skall ändras så att det helst inte väljs mer än butik per bransch och postnummer när det finns 800 valda regioner.

## 7 Diskussion

### 7.1 Täckning

Riktmärken i form av regler för HICP och rekommendationer från ILO samt tillämpade metoder inom EU visar att Sveriges KPI och HICP har haft höga krav på hög täckningsgrad för rampopulationen och att metoder med sannolikhetsurval skall tillämpas. Därmed inte sagt att enstegsurval av butiker från hela riket, utan geografisk eller annan stratifiering än bransch, är helt optimalt.

Då kostnader för intervjuares resor för prismätning i butik är relativt stor och det inte går att påvisa någon regional effekt på prisindex så finns ett utrymme för att i viss mån minska täckningsgraden genom medveten undertäckning av kommuner. Variansanalyserna visar att regionala effekter är mindre än effekter av produktgrupper inom olika produktområden. Beaktat det sätt som används för att välja representativa produkter inom produktgrupper, exemplifierat med gräsklippare i avsnitt 5.2 så framstår det som rimligt att urval av butiker i regioner som medför långa resor inte nödvändigtvis är motiverade ur en kvalitetsaspekt. Resurserna bör hellre användas till att samla fler priser där det är kostnadseffektivt att samla in.



## 7.2 Regionurval

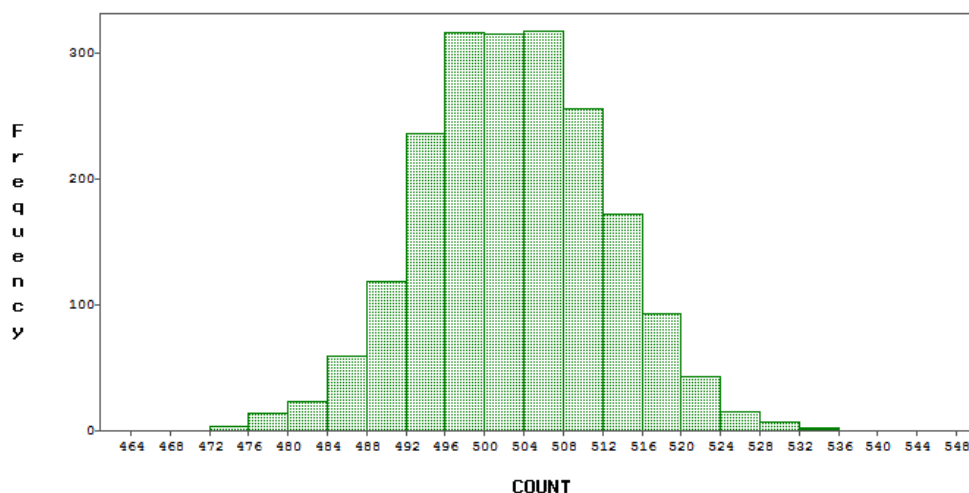
Många nationella statistikbyråer väljer försäljningsställen i två steg, först regioner och sedan enskilda försäljningsställen. Vi kan inte med säkerhet tolka flera av de beskrivningar som lämnats av statistikbyråerna till Eurostat, men det är uppenbart inte ovanligt att regioner valts utan sannolikhetsurval, alltså istället medvetet. Storbritannien och USA är dock exempel på att sannolikhetsurval av regioner också tillämpas.

För Sveriges KPI innebär SAMU-systemet för urval inom förestagssektorn stora fördelar som bör behållas vid en re-design. För små butiker, d.v.s. butiker med en urvalsannolikhet mindre än tio procent, gäller att de förväntas vara med i urvalet i fem år om de väljs. Om vi inför ett första steg med regionurval kan SAMU-metoden ändå behållas i andra steget som nämnt i avsnitt 6.3. Om det inte sker större omläggningar i postnummersystemet och vid bibehållna strukturer på butikerna avseende storleksmått torde urvalen bli tämligen beständiga över tiden, utöver den förutbestämda rotationen.

Samtidigt kan vi behålla och effektivisera algoritmer som ger spridning av butikerna i urvalen per bransch regionalt samt förbättra tröghetsregler som på ett ”gynnsamt” sätt skapar positiv samordning mellan åren.

## 7.3 Effekt

Diagram 7 Fördelning av antal postnummerområden per urval i simuleringarna



En önskad effekt av en redesign var att få en geografisk koncentration av butiksurlvalet. Enligt diagram 7 kan ses att det ingår oftast runt 500 postnummerområden i urvalen, vilket ska ställas mot exempelvis 2015 års urval där det ingick drygt 1 300 postnummerområden.

## 7.4 Faktiska urvalssannolikheter

För beräkning av *faktiska* urvalssannolikheter har 4 000 butiksurlval dragits i en simuleringsstudie. Med *faktisk* urvalssannolikhet åsyftas i detta sammanhang, när  $\pi$ ps-urval används som urvalsdesign, den faktiska förekomsten av objekt i urvalet när objekten har varierande storleksmått. Detta måste observeras genom upprepade slumpmässiga dragningar av (olika) urval från ramen. Begreppet kan ställas mot *önskad* urvalssannolikhet, vilket är en funktion av objektets storleksmått ställd i relation till det sammanlagda storleksmättet över alla objekt

som kan sägas vara det förväntade värdet för objektets urvalssannolikhet. Avvikelser mellan den *faktiska* och den *önskade* urvalssannolikheten kan säga göra skattningen icke väntevärdesriktig. I praktiken skulle det innebära att prisnoteringarnas designvikter inte motsvaras av deras vägningstal och urvalet blir inte längre självvägande.

I diagrammen 8.1-8.5 visas fem branscher av sammanlagt 47 olika branscher som prismätningar görs i (samtliga branscher ej redovisade i diagramform här). Det har noterats att det i de flesta branscher fås en hög samstämmighet mellan faktisk urvalssannolikhet och önskad urvalssannolikhet baserat på simuleringsstudien. Några fall sticker ut och redovisas i diagrammen 8.4 och 8.5.

Diagram 8.1 Faktisk urvalssannolikhet mot önskad urvalssannolikhet. Dagligvarubutiker SNI 47112

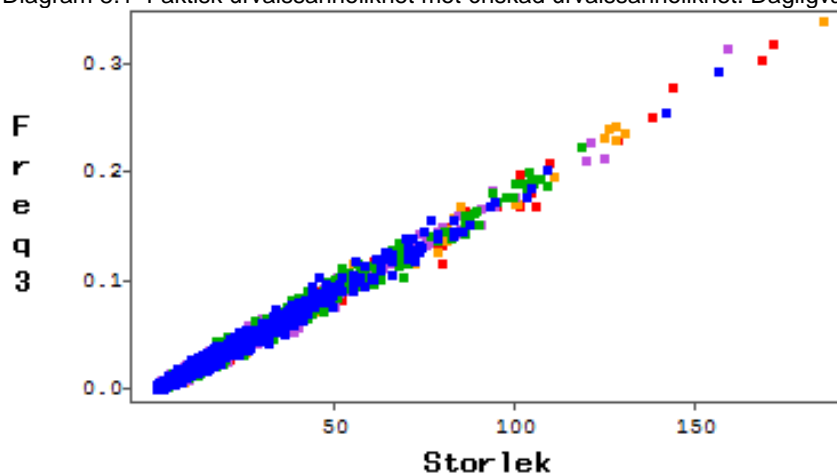


Diagram 8.2 Faktisk urvalssannolikhet mot önskad urvalssannolikhet. Stormarknader SNI 47111

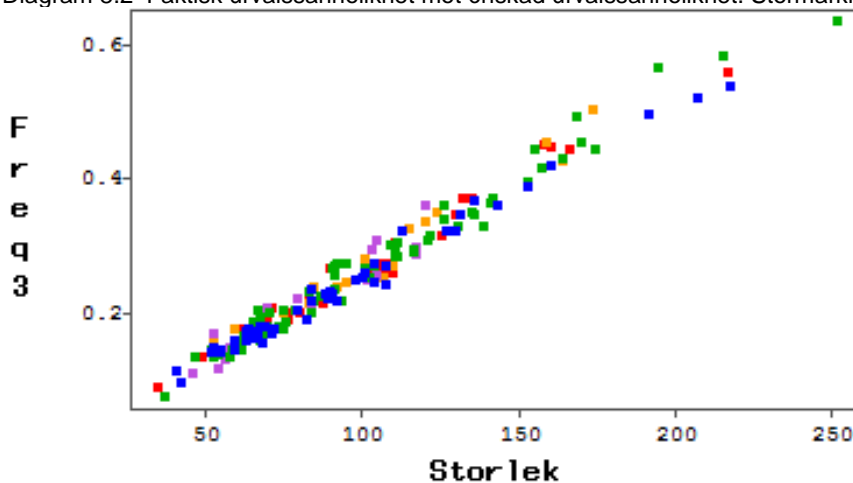


Diagram 8.3 Faktisk urvalssannolikhet mot önskad urvalssannolikhet. Radio- och TV-affär SNI 47430

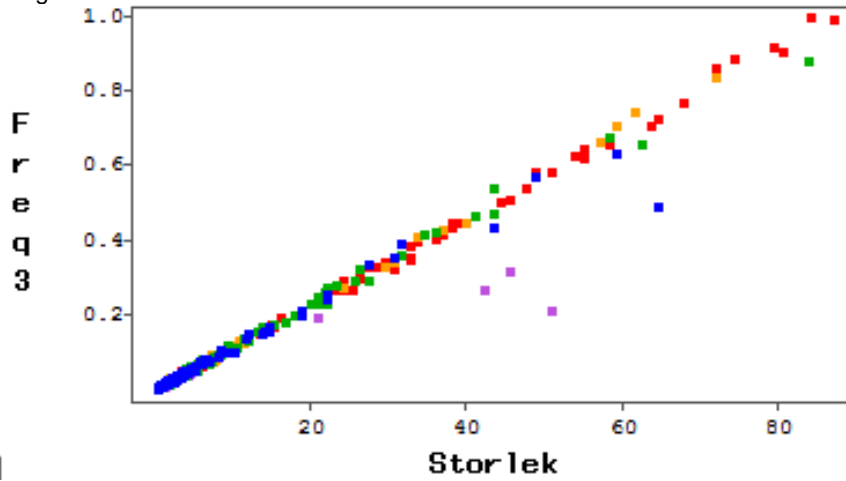


Diagram 8.4 Faktisk urvalssannolikhet mot önskad urvalssannolikhet. Hemtextilbutik SNI 47532

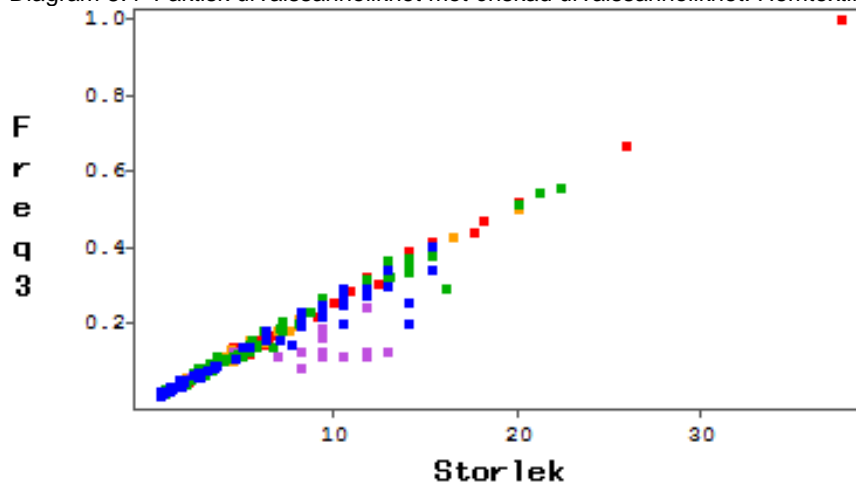


Diagram 8.5 Faktisk urvalssannolikhet mot önskad urvalssannolikhet. Specialiserad butikshandel med kontorsförbrukningsvaror | SNI 47622

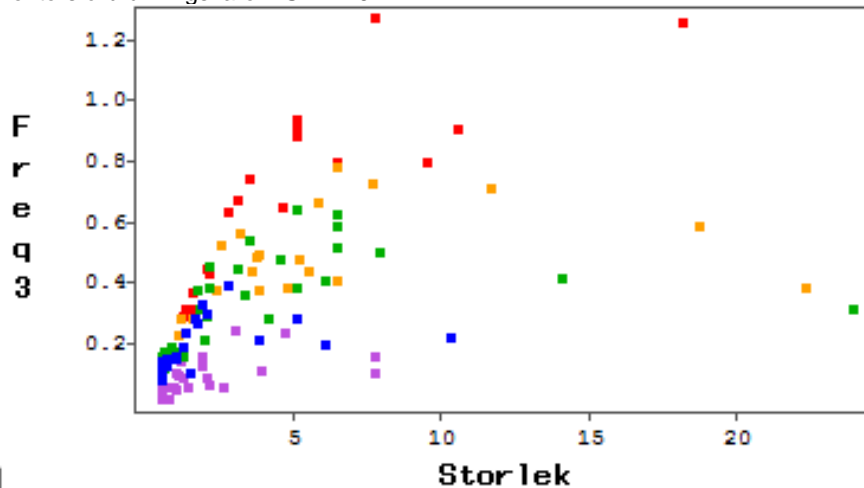


Diagram 8.5, Specialiserad butikshandel med kontorsförbrukningsvaror är en bransch där insamling rekommenderas genom telefonintervju. Butikerna ligger förmodligen mer sällan i köpcentra utan oftare i industri- och kontorsområden. Enligt simuleringarna finns det stora variationer i utfallet av faktiska urvalssannolikheter visavi önskade urvalssannolikheter för denna bransch. Urvalsstorleken av butiker är så stor att det inte säkert går att hitta önskat bruttourval av butiker i urvalet av postnummerområden och därför blir den faktiska urvalssannolikheten i landet för låg.

## 7.5 "Tröghetsregeln"

Urval som från år till år är samordnade är bra för förändringsskattningar. Med SAMU-systemet förväntas 20% av små objekt ut varje år bytas ut, vilket är bra för att uppgiftslämnarbördan skall bli jämnare, på sikt. När objekten ändrar storlek mellan åren så hoppar butiker ut och in i urvalet eftersom urvalssannolikheten är proportionell mot storleken. Att butiker hoppar ut och in innebär tre nackdelar:

- Det är en praktisk och ekonomisk olägenhet för SCB eftersom alla butiker i urvalet för det kommande året som inte redan är med i urvalet skall sortimentundersökas vid ett besök i oktober av en prisinsamlare.
- Förändringsskattningar, även fast KPI är en kedjeindex, får bättre kvalitet ju större andel av urvalen som är gemensamma två år i rad.
- Det kan av butiksägarna gissningsvis uppfattas ett obegripligt förfarande att vara med i urvalet vartannat år eller så.

Från och med 2009 års urval "favoriseras" butiker som redan är med i KPI:s prismätningar (inte hela bruttourvalet) genom att storleksmättet multipliceras med faktorn 2 och urvalssannolikheten, betingat föregående års urval, därmed ökas. Förfarandet innebär att objekt som minskat i storlek överrepresenteras och objekt som ökat i storlek underrepresenteras.

Analysen visar att effekten av den vidtagna åtgärden är liten, det är knappt tio butiker färre som återkommer i urvalet efter ett års frånvaro med faktorn 2 än med faktorn 1. Vi ser exempel på att butiker är borta ur urvalet ett år och återkommer igen beroende på att branschkodningen i företagsregistret tillfälligt har varit en annan. Några exempel, utan anspråk på att det är de vanligaste, är bensinstationer som omklassas som biluthyrning och sedan tillbaka till bensinstation, bensinstationer som omklassas till restaurang och sedan tillbaka, bensinstation till livsmedelsbutik och tillbaka, hemtextilbutik till möbelbutik och tillbaka.

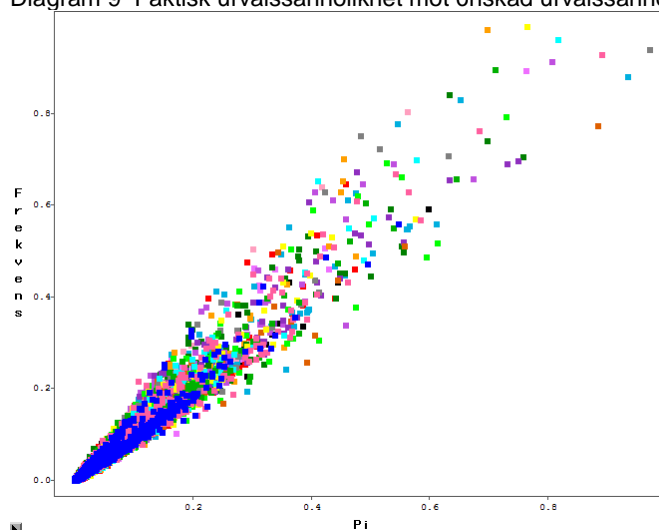
Om vi vill ha en kraftigare effekt kan vi intervensera genom att välja faktorn 4 eller kanske ännu högre, men inte så högt att den 20-procentiga SAMU-rotationen sätts ur spel. En alternativ ansats vore att införa en hård regel av typen att butiker som varit med i urvalet två år och sedan det tredje året inte är med i urvalet faktiskt inte skall kunna vara med i urvalet på förslagsvis fem år.

## 7.6 Balansering

Det är inte optimalt för precisionen i statistiken och det är inte bra för fördelningen av den lokala prisinsamlingen som utförs av SCB:s intervjuare om designen är sådan att ett urval kan råka bli så att två i väsentliga avseende likadana butiker hamnar i urvalet samtidigt. För närvarande förhindras detta genom en intervention som beskrevs i avsnitt 1.2.

Ett optimalt förfarande för denna justering skulle inte behöva ändra den första ordningens urvalssannolikheter medan andra ordningens skulle påverkas gynnsamt och därmed generera lägre varians. Med andra ordningens urvalssannolikheter menas sannolikheten för två butiker att vara med samtidigt i urvalet. Vi har inte funnit denna optimala metod. Tyvärr innebär tillämpad metod att även första ordningens urvalssannolikheter inte är exakt de önskade, men avvikelserna är inte stora, vilket visas i diagram 9.

Diagram 9 Faktisk urvalssannolikhets mot önskad urvalssannolikhets



Metoden bedöms fungera tillfredsställande för små urval, storleksordning  $n = 10 - 30$ , då det är opraktiskt att stratifiera på t.ex. län, kommun, butikskedja och liknande.

## 7.7 Avgränsning

Den ovan utförda analysen med urval av postnummerområden avser samtliga branscher, alltså både där insamlingssättet är besöksintervju och telefonintervju. En kostnadsbesparing bör kunna uppnås genom att tillämpa metoden enbart för branscher där rekommenderat insamlingssätt är besöksintervju och inte telefonintervju.

## 8 Referenser

- Ohlsson, E. (1990). Sequential Poisson Sampling From a Business Register and its Application to the Swedish Consumer Price Index. R&D Report 1990:6, Statistics Sweden
- Ohlsson, E. (1995). Coordination of Samples Using Permanent Random Numbers. In Cox et al: Business Survey Methods, Wiley & Sons, pp. 153–169.
- Norberg, A. (2004). Comparison of Variance Estimators for the Consumer Price Index. 8th Ottawa Group Meeting - Helsinki - 23-25 August 2004 (läst 2015-05-12), [http://www.ottawagroup.org/Ottawa/ottawagroup.nsf/home/Meeting+8/\\$file/2004%208th%20Meeting%20-%20Norberg%20Anders%20-%20Comparison%20of%20Variance%20Estimators%20for%20the%20Consumer%20Price%20Index.pdf](http://www.ottawagroup.org/Ottawa/ottawagroup.nsf/home/Meeting+8/$file/2004%208th%20Meeting%20-%20Norberg%20Anders%20-%20Comparison%20of%20Variance%20Estimators%20for%20the%20Consumer%20Price%20Index.pdf)

Statistiska centralbyråns hemsida SCB.se, sök efter “Lokala arbetsmarknader” (läst 2015-05-12),



[http://www.scb.se/sv\\_/Hitta-statistik/Statistik-efter-amne/Arbetsmarknad/Sysselsattning-forvarvsarbete-och-arbetstider/Registerbaserad-arbetsmarknadsstatistik-RAMS/7899/Lokala-arbetsmarknader-LA/](http://www.scb.se/sv_/Hitta-statistik/Statistik-efter-amne/Arbetsmarknad/Sysselsattning-forvarvsarbete-och-arbetstider/Registerbaserad-arbetsmarknadsstatistik-RAMS/7899/Lokala-arbetsmarknader-LA/)

## 9 Appendix

### 9.1 A1: Rekommenderat insamlingsätt per bransch

Stratum/ BranschKod	Beskrivning	Urvals- storlek	Insamlingsätt
451100	HANDEL MED PERSONBILAR	5	Besök
452010	BILSERVICEVERKSTÄDER, EJ SPECIALISERADE	5	Telefon
452040	DÄCKSERVICEVERKSTÄDER	10	Telefon
453200	DETALJHANDEL MED BILRESERVDELAR & TILLBEHÖR	5	Besök
467300	Partihandel m virke, andra byggmat. & sanitetsgods	10	Besök
471110	STORMARKNADER, BRETT SORTIMENT	56	Besök
471900	VARUHUS, BRETT SORTIMENT, EJ LIVSMEDEL	7	Besök
471120	LIVSMEDELSBUTIKER MED BRETT SORTIMENT	32	Besök
472600	TOBAKSAFFÄRER	6	Telefon
472910	Hälsokostbutiker	5	Telefon
473000	BENSINSTATIONER	80	Telefongruppen i Örebro
474100	Butiker med datorer, programvara, data- och tv-spel	12	Besök
474200	TELEBUTIKER	26	Telefon
474300	RADIO- & TV-AFFÄRER	20	Besök
475100	GARN-, TYG- & SYBEHÖRSAFFÄRER	8	Telefon
475200	JÄRN-, BYGG- & VVS-VARUHANDEL	12	Besök
475230	FÄRGHANDEL	5	Telefon
475310	TAPET- & MATT OCH GOLVBELÄGGNINGSHANDEL	9	Telefon
475320	HEMTEXTILBUTIKER	23	Besök
475400	VITVARU- & HUSHÅLLSMASKINAFFÄRER	7	Besök
475910	ÖVRIGA MÖBEL- OCH MATTAFFÄRER	25	Besök
475930	BOSÄTTNINGSAFFÄRER	14	Besök
475940	BELYSNINGS- OCH ELAFFÄRER	6	Telefon
475950	MUSIKHANDEL	5	Telefon
476100	BOK- OCH PAPPERSHANDEL	7	Telefon
476220	PAPPERSHANDEL	7	Telefon
476300	SKIV- OCH VIDEOBUTIKER	12	Telefon
476400	SPORT- OCH CYKELAFFÄRER	30	Besök
476500	LEKSAKSAFFÄRER	20	Besök
477100	KLÄDESAFFÄRER	120	Besök
477210	SKOAFFÄRER	31	Besök
477220	VÄSKAFFÄRER	10	Telefon
477300	APOTEK	5	Besök
477500	PARFYMAFFÄRER	7	Telefon

Stratum/ BranschKod	Beskrivning	Urvals- storlek	Insamlingsätt
477610	BLOMSTER- & TRÄDGÅRDSVARUHANDEL	25	Telefon
477620	Zooaffärer	5	Telefon
477710	URAFFÄRER	9	Telefon
477720	GULDSMEDSAFFÄRER	12	Telefon
477810	OPTIKER	10	Telefon
477820	FOTOAFFÄRER	5	Telefon
477890	ÖVRIGA SPECIALBUTIKER	5	Telefon
479110	Postorder, brett sortiment	7	Web
479120	Postorder, textil och konfektion	9	Web
479130	Postorder/internet, böcker	9	Web
479140	Postorder/internet, datorer & elektronisk utrustning	5	Web
479150	Postorder/internet, sport och fritid	0	Web
479190	Postorder/internet	9	Web
551000	HOTELL EXKL KONFERENSANLÄGGNINGAR (KPI)	47	Telefon
561000	RESTAURANGER	85	Telefon
562910	PERSONALMATSAL	10	Telefon
750000	Veterinärkliniker	5	Telefon
772200	VIDEO- OCH DVDFILMUTHYRARE	5	Telefon
952300	SKOMAKERIER, KLACKBARER MM	5	Telefon
960120	KONSUMENTTVÄTTERIER	5	Telefon
960210	FRISÖRSALONGER	35	Telefon



## 9.2 A2: Variansanalys

Test av signifikans för produktgrupper och regioner enligt variansanalysmodellen  
 $\ln(\text{priskvot}_{p,b,e}) = \text{Produktgrupp}_p + \text{Region}_r + \varepsilon_e$

där  $\text{Produktgrupp}_p$  = genomsnittlig priskvot för produktgrupp p

$\text{Region}_r$  = genomsnittlig priskvot för regionen r.

och  $\varepsilon_e$  = en ”slumpeffekt”, vilken uppstår av intervjuarens val av produkterbjudande.

Produktområde	År	p-värde	
		Produktgrupp	Region
Kläder	2011	0,05	0,11
Kläder	2012	0,65	0,16
Kläder	2013	0,28	0,16
Kläder	2014	0,13	0,11
Skodon	2011	0,92	0,88
Skodon	2012	0,38	0,08
Skodon	2013	<b>0,00</b>	0,24
Skodon	2014	0,76	0,69
Möbler	2011	<b>0,01</b>	0,11
Möbler	2012	0,16	0,66
Möbler	2013	0,35	0,08
Möbler	2014	0,32	0,99
Husgeråd	2011	0,15	0,38
Husgeråd	2012	<b>0,00</b>	0,11
Husgeråd	2013	<b>0,00</b>	<b>0,01</b>
Husgeråd	2014	0,39	0,56
Vitvaror	2011	0,11	0,26
Vitvaror	2012	0,89	0,05
Vitvaror	2013	0,33	0,37
Vitvaror	2014	0,82	0,57
Hemelektronik	2011	<b>0,02</b>	0,60
Hemelektronik	2012	<b>0,03</b>	0,11
Hemelektronik	2013	<b>0,03</b>	0,39
Hemelektronik	2014	<b>0,00</b>	0,51
Verktyg	2011	<b>0,00</b>	0,30
Verktyg	2012	0,10	0,53
Verktyg	2013	0,29	0,37
Verktyg	2014	0,73	0,20
Leksaker	2011	<b>0,03</b>	0,69
Leksaker	2012	<b>0,00</b>	<b>0,04</b>
Leksaker	2013	0,18	<b>0,03</b>
Leksaker	2014	<b>0,01</b>	0,46
Kost & logi	2011	0,45	0,47
Kost & logi	2012	<b>0,00</b>	0,71
Kost & logi	2013	<b>0,03</b>	0,21
Kost & logi	2014	<b>0,03</b>	0,82

### 9.3 A3: Butiksvikten i urvalsstratum

Av Martin Ribe, utkast 2005-01-28

Följande tillämpas i lokalpris- och dagligvarusystemen från årsköret 2005.

Butiksvikten är ett storleksmått för vägning av varje butik i urvalsstratum, när butiker i totalstratum vägs med antal anställda:

$$W = \frac{\sum_{\substack{\text{butiker} \\ \text{i bransch, ram,} \\ \text{urvalsstratum}}} \#(\text{anställda})}{\sum_{\substack{\text{butiker} \\ \text{i bransch,} \\ \text{bruttourval}}} \#(\text{anställda})} \cdot \frac{\sum_{\substack{\text{butiker} \\ \text{i bransch,} \\ \text{nettourval}}} \#(\text{anställda})}{\sum_{\substack{\text{butiker} \\ \text{i bransch,} \\ \text{nettourval}}} 1}$$

”Antal anställda” avser i praktiken antalet anställda plus ett.

Butiksvikten används i medelvärdesestimation enligt:

$$\bar{X} = \frac{\sum_{\substack{\text{butiker} \\ \text{i bransch,} \\ \text{totalstratum}}} \#(\text{anställda}) \cdot X + W \cdot \sum_{\substack{\text{butiker} \\ \text{i bransch,} \\ \text{urval}}} X}{\sum_{\substack{\text{butiker} \\ \text{i bransch,} \\ \text{totalstratum}}} \#(\text{anställda}) + W \cdot \sum_{\substack{\text{butiker} \\ \text{i bransch,} \\ \text{urval}}} 1}$$

Vid indexberäkning enligt Jevons-formeln (”geometriskt index”) avser  $X$  logaritmerad priskvot och  $\bar{X}$  logaritmerat indextal.