

## Idé om ny design för KPI kläder

### *För diskussion*

---

Syftet med denna PM är att presentera en ny idé för prismätningar av kläder. Hittills har SCB varit noga med att använda vetenskapliga metoder för att generera urval av butiker. Det visar sig emellertid att butiker inte styr prisbilden lika mycket som varumärken. SCB introducerar därför en ny design idé som utgår från varumärke. Undersökningens totala osäkerhet bör utvärderas med avseende på alla felkällor enligt kvalitetsbegreppet i Sveriges officiella statistik.

För att troliggöra att de antaganden som ligger till grund för den nya design idén är korrekta föreslår vi en empirisk undersökning.

Nämnden inbjuds att komma med synpunkter på den nya designen samt den föreslagna provmätningen.

---

### INNEHÅLL

<b>IDÉ OM NY DESIGN FÖR KPI KLÄDER</b> .....	<b>1</b>
<b>1 BAKGRUND</b> .....	<b>2</b>
<b>2 SYFTE</b> .....	<b>2</b>
<b>3 DAGENS METOD</b> .....	<b>2</b>
<b>4 IDÉ OM NY METOD</b> .....	<b>3</b>
<b>5 EMPIRISKA BERÄKNINGAR</b> .....	<b>4</b>
5.1 MÄRKESKVALITETSKODBYTEN .....	4
5.2 VARUMÄRKENAS MARKNADSANDELAR.....	4
5.3 VARUMÄRKEN I SMÅ OCH STORA KLÄDEFÖRETAG .....	6
5.4 HEDONISKA ANALYSER.....	7
<b>6 DISKUSSION</b> .....	<b>8</b>
6.1 OSÄKERHETSKOMPONENT URVAL .....	9
6.2 OSÄKERHETSKOMPONENT RAMTÄCKNING.....	10
6.3 OSÄKERHETSKOMPONENT MÄTNING .....	10
6.4 OSÄKERHETSKOMPONENT BORTFALL.....	10
6.5 OSÄKERHETSKOMPONENT BEARBETNING .....	10
6.6 OSÄKERHETSKOMPONENT MODELLANTAGANDEN.....	11
<b>7 FÖRSLAG</b> .....	<b>11</b>
<b>8 REFERENSER</b> .....	<b>11</b>
<b>APPENDIX</b> .....	<b>12</b>
A1: KVALITETSBEGREPPET I SVERIGES OFFICIELLA STATISTIK .....	12



## 1 Bakgrund

Inom varugruppen kläder är livslängden för produkterbjudanden (pe) relativt kort och flera ersättningar behöver normalt göras under en tolv månadersperiod för att skapa en serie av månatliga priser.

För kläder infördes en ny mätmetod fr.o.m. 1991 och en ny metod för värdering av kvalitetsskillnader mellan utgångna pe och ersättningar fr.o.m. 1994. Med en statistisk, s.k. hedonisk modell, applicerad på KPI-data för en begränsad tidsperiod av ett år skattas marknadens värderingar av olika egenskaper hos plaggen. Modellen använder uppgifter om varornas material, utförande, varumärke m.m. Förutom priser samlar SCB därför även in dessa egenskapsvariabler månadsvis.

När plagg utgår ur butikernas sortiment och därför måste ersättas med ett annat plagg används modellernas skattade koefficienter för att räkna om priset i basen. På så sätt blir prisnivån på ersättningsplaggen jämförbara med basvarianterna i kvalitetshänseende. Det är SCB:s erfarenhet att den hedoniska kvalitetsjusteringen generellt sett fungerar bra, men att de fysiska egenskaperna hos plaggen har relativt låg förklaringsgrad i de hedoniska modellerna.

Priskvoter för kläder har i jämförelse med många andra produktgrupper stor varians. SCB belastar idag inte enstaka butiker med stora urval av pe:n. Istället har urvalet av klädesbutiker utökats, vilket är en förklaring till undersökningens relativt sett höga kostnad, många butiker innebär en höga reskostnader. Relativt andra KPI-undersökningar är även prisinsamlingen dyr för kläder beroende på den tid det tar att notera alla egenskaper vid frekventa byten av pe:n.

Marknaden för kläder har under senare år ändrats. Dels har en tydlig förflyttning gjorts mot "handel på nätet" och dels tillämpas så kallad central prissättningen för flera stora företag i urvalet. Central prissättning innebär att ett mångbutiksföretag har samma priser i alla sina butiker (även nät) för en specifik produkt. Sortimenten och troligen säsonger varierar dock i de olika butikerna. Vinterkläder kan exponeras tidigare i norr än i söder, en butik på "fina gatan" har fler exklusiva produkter än en butik i förort och realisationspriser introduceras ofta först på nätet. Generellt kan man se att en webbutik har ett större sortiment än en fysisk butik. SCB har de senaste åren flyttat en del av prismätningen från fysiska butiker till hemsidor. För stora mångbutiksföretag har i dessa fall pe:n i flera butiker ersatts med motsvarande totala mängd pe:n från en webbutik.

Mot bakgrund av ovan ser SCB ett behov av att se över undersökningsdesignen för produktgruppen kläder.

## 2 Syfte

Syftet med denna PM är att presentera en idé till ny design för kläder som ger mer "value for money". Baserat på den felmodell som beskrevs i PM "En modell för resursallokering inom KPI" vid nämndens möte nr 4 föreslås att en alternativ design utvärderas m.a.p. "total survey error", d.v.s. att alla osäkerhetskällor beaktas.

## 3 Dagens metod

Dagens metod kan karaktäriseras på följande sätt:

- Ett sannolikhetsurval av butiker dras där inverterade inklusionssannolikheter används för uppräknig av prisobservationer på ett vetenskapligt sätt.

- SCB mäter även priser på webben i ett antal som ungefär motsvarar webb-handelns andel. Per företag är dock antalet inte alltid proportionellt mot marknadsandelen. Webbinsamling görs enbart för stora mångbutiksföretag.
- Undersökningen har en god täckning av klädtyper. Dessa mäts med 25 medvetet specificerade produktgrupper. I undersökningen ingår exempelvis ej herrunderkläder (strumpor, undertröja) som representeras av kalsonger, herrkostym som representeras av kavaj, festklänningar (aftonklänning, brudklänning) som representeras av vanliga klänningar.
- En medveten fördelning av produktgrupper görs per butik så att cirka 25 pe:n skall mätas. I stora försäljningsställen ”har det blivit så” att fler pe:n väljs. Av effektivitets-skäl mäts gärna fyra produkter av varje produktgrupp i en butik men alltså inte alla 25 produktgrupper även om de skulle finnas.
- Valet av pe som prisinsamlarna gör i butik utgår från den som bedöms populär och mycket såld. Ersättningarna under året skall göras enligt ”mycket sålt och ganska likt” föregående.
- Kvalitetsvärdering med hedonisk analys bedöms inte leda till bias men minskar inte heller medelfelet för prisindex vilket beror på att dagens modell har ganska låg förklaringsgrad.
- Varumärket har stor betydelse vid kvalitetsjusteringen. Under 90-talet nyttjade SCB tjänster från experter på en modetidskrift för att klassificera varumärkenas status i fem klasser. Klassificeringen gjordes oberoende av insamlade priser. Klassificering enligt prisnivå skulle leda till bias i prisindex. Idag görs klassificeringen i nio statusklasser av SCB. 90 % av data är dock koncentrerade till kvalitetsklass 3 – 6.

## 4 Idé om ny metod

Istället för urval av butiker och produkter utgår undersökningen från ett urval av de vanligast förekommande varumärkena. Med data från de senaste årens prismätning av kläder kan SCB ta fram uppgifter om antalet pe:n per varumärke. Cut-off-urval tillämpas sedan på ett lämpligt antal varumärken. Enligt analyser av damkläder svarar 46 märken för 80 % av alla pe:n. För herrkläder är motsvarande antal 56.

Urvalet av varumärken blir därmed inget problem det första året, då tidigare undersökningar kunde användas som ram. Kommande år kan förhoppningsvis information om alla importörer av kläder från Importprisindex (IMPI) användas.

I de fall central prissättning tillämpas hos företag bör prisinsamlingen för kläder ske via nätbutiken istället för i de fysiska butikerna. Det stora mångbutiksföretagen har i hög grad enbart sina egna varumärken.

Andra märken kan finnas i ett flertal webbutiker. Här krävs ett urvalsförfarande för att välja en eller flera sajter (butiker) där priset skall hämtas.

Utöver märke insamlas fysiska variabler på samma sätt som nu, d.v.s. material och utförande. Den hedoniska modellen utrustas med en dummyvariabel för varje märke, däremot ingår inte längre dummyvariabler för butiksgroup. Kvalitetsjustering av baspris görs nu enbart med avseende på fysiska egenskaper.

## 5 Empiriska beräkningar

### 5.1 Märkeskvalitetskodbyten

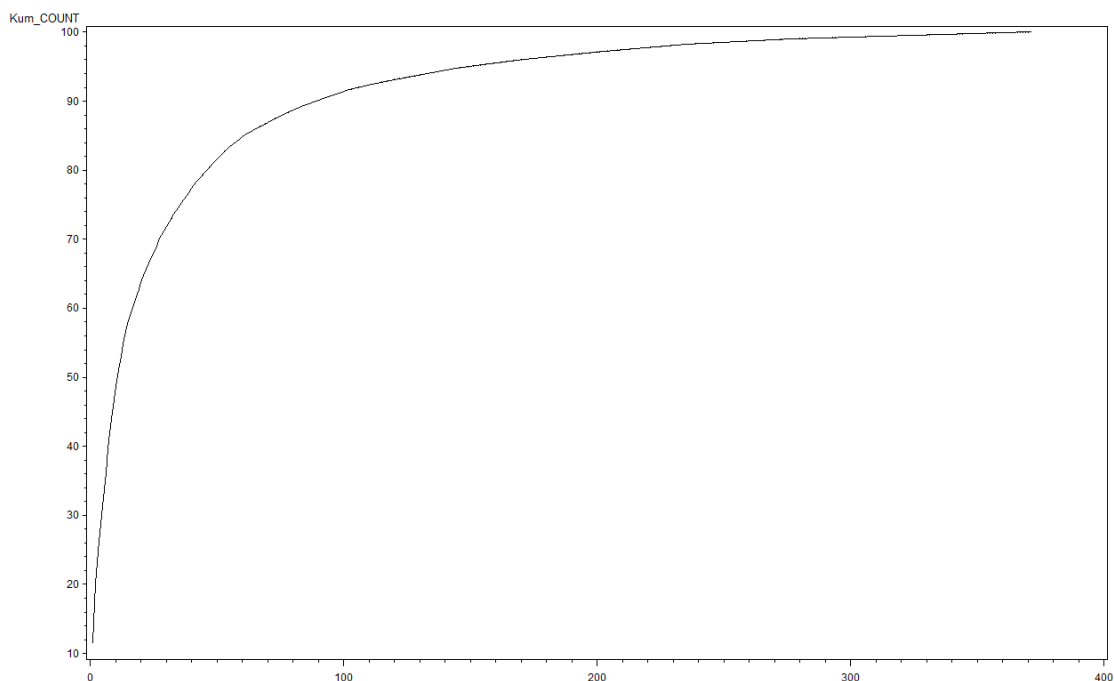
Antalet ändringar av märkenas kvalitetskoder mellan åren kan vara en indikation på kvaliteten i SCB:s kodning. Det skulle t.ex. inte vara bra om ett visst märkes kvalitetskod justeras uppåt ett år och nedåt nästa. Av 1 105 varumärken 2014 – 2018 så är det endast 14 som har fått ändra kvalitetskod. Fem gånger har kvalitetskoden sänkts och nio gånger har den höjts. Ingen gång har ett varumärke fått höjd kod och senare sänkt kod eller vice versa.

### 5.2 Varumärkenas marknadsandelar

Denna analys omfattar produktgrupperna kjol, klänning långbyxor, blus, tröja, kavaj och jacka.

Under 2016 – 2018 fanns 371 olika varumärken i insamlade data. Ett litet antal varumärken utgör en stor andel av observationerna, se figur 1 och tabell 1.

Figur 1 Kumulativ fördelning av observationer per varumärke för damkläder

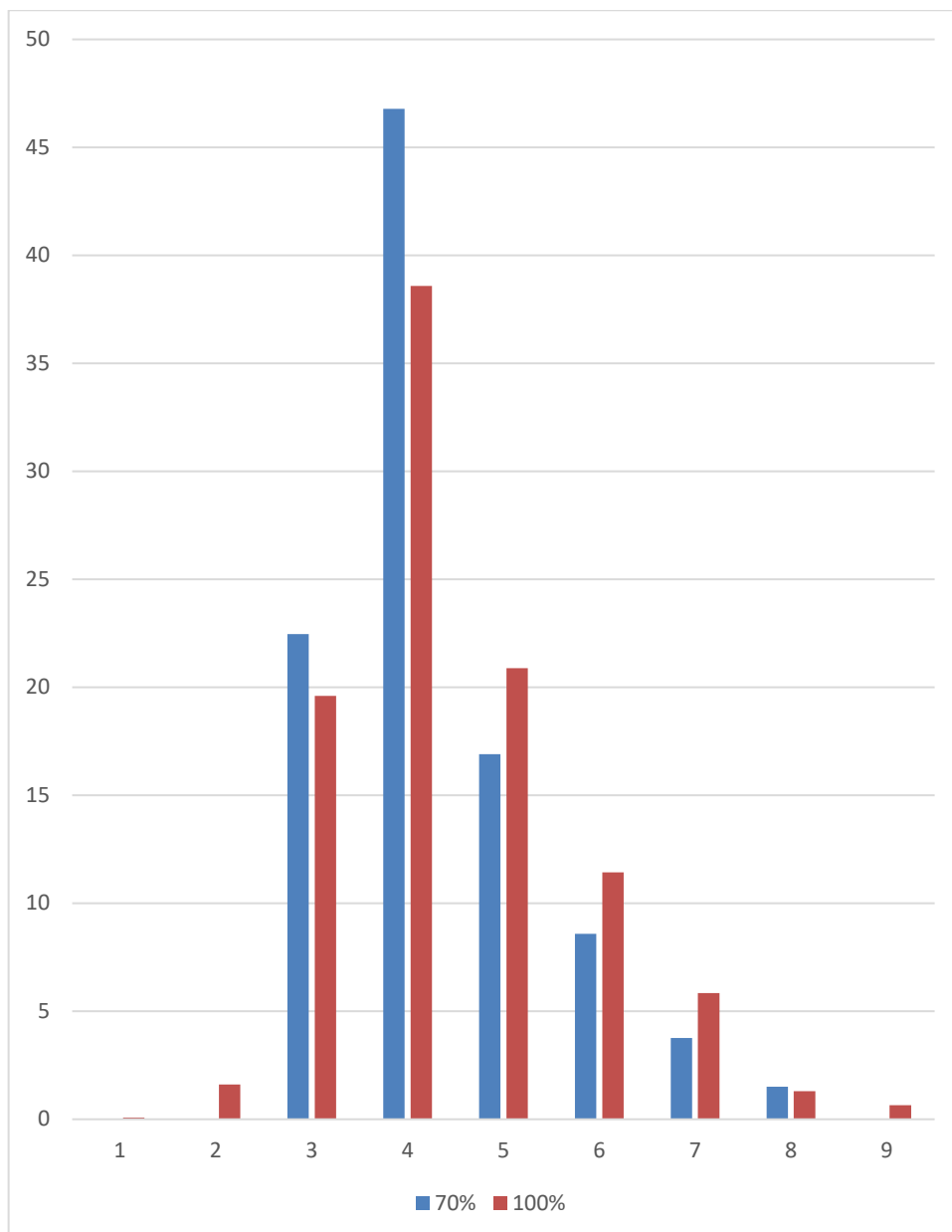


Tabell 1 Kumulativ fördelning av observationer per för damkläder

Antal varumärken	Andel av observationerna
11	50 %
28	70 %
46	80 %
89	90 %
150	95 %

Figur 2 visar att med 28 varumärken, vilka täcker 70 % av alla prisobservationer för damkläder, så har vi större andel varumärken i kvalitetsklass 4 och 3 men mindre andel märken i klasserna 5 – 7. Skillnaderna är dock inte dramatiska.

Figur 2 Andel observationer per märkeskvalitetsklass för damkläder



### 5.3 Varumärken i små och stora klädesföretag

I tabell 2 visas varumärkenas fördelning med avseende på den s.k. kvalitetskod som SCB sätter på märket för fyra grupper av butiker med avseende på företagets totala antal butiker i Sverige. De stora mångbutiksföretagens marknadsandel är stor (56,7 %) och deras märken har SCB klassat i intervallet 3 – 6 i huvudsak.

Tabell 2 Fördelning av damkläder på varumärkets kvalitetsklass och butiksgrupp 2016 – 2018. Procent

Märkens kvalitetsklass	Enbutiksföretag	Företag med 2-9 butiker	Företag med 10-49 butiker	Företag med 50- butiker	Alla butiker
1	0,1	0,0	0,0	0,0	0,1
2	0,8	0,0	0,0	0,8	1,6
3	1,3	1,3	1,4	15,6	19,6
4	2,2	3,7	1,0	30,6	37,5
5	3,6	1,8	8,1	8,0	21,4
6	5,8	2,3	2,5	1,2	11,8
7	1,6	1,8	2,0	0,6	6,0
8	0,2	0,5	0,6	0,0	1,3
9	0,1	0,1	0,5	0,0	0,7
Summa	15,7	11,4	16,2	56,7	100,0

Tabell 3a – 3c visar i vilken grad som företag av olika storlek har de mest frekventa varumärkena eller många andra. De stora mångbutiksföretagen säljer i hög grad sina egna varumärken och dessa varumärken säljs inte av något annat företag. De 28 mest frekventa varumärkena utgör  $47,7 / 56,7 = 84 \%$  av all handel där. För enbutiksföretag är det tvärtom,  $12,8 / 15,7 = 82 \%$  av handeln avser andra varumärken än de på riksnivå mest sålda, allt enligt fördelningen av insamlade data för KPI.

Tabell 3a Fördelning av damkläder efter varumärkens frekvens och butiksgrupp 2016 – 2018. Procent

Frekventa resp icke frekventa märken	Enbutiksföretag	Företag med 2-9 butiker	Företag med 10-49 butiker	Företag med 50- butiker	Alla butiker
28 varumärken som ger 70 % täckning	2,9	8,9	10,7	47,7	70,2
Övriga mindre frekventa märken	12,8	2,5	5,5	9,0	29,8
Summa	15,7	11,4	16,2	56,7	100,0

Tabell 3b Fördelning av damkläder efter varumärkens frekvens och butiksgrupp 2016 – 2018. Procent

Frekventa resp icke frekventa märken	Enbutiksföretag	Företag med 2-9 butiker	Företag med 10-49 butiker	Företag med 50- butiker	Alla butiker
46 varumärken som ger 80 % täckning	5,8	9,4	12,7	52,5	80,3
Övriga mindre frekventa märken	9,9	2,0	3,6	4,2	19,7
Summa	15,7	11,4	16,2	56,7	100,0

Tabell 3c Fördelning av damkläder efter varumärkens frekvens och butiksgrupp 2016 – 2018. Procent

Frekventa resp icke frekventa märken	Enbutiksföretag	Företag med 2-9 butiker	Företag med 10-49 butiker	Företag med 50- butiker	Alla butiker
150 varumärken som ger 90 % täckning	9,6	10,2	14,8	55,4	90,1
Övriga mindre frekventa märken	6,1	1,2	1,4	1,3	9,9
Summa	15,7	11,4	16,2	56,7	100,0

#### 5.4 Hedoniska analyser

Regressionsmodeller har skattats med data för damkläder från prismätningarna 2016 – 2018. Den första observationen av varje enskilt plagg har använts. Modellen, som har logaritmen av ordinarie pris som beroende variabel, skattas för en grupp av produktgrupper simultant. Alla förklarande variabler är av dummy-typ.

De förklarande variablerna är:

- Produktgrupper: kjol, klänning långbyxor, blus, tröja, kavaj och jacka
- Fysiska egenskaper: material, foder, fickor etc
- Varumärken, dels de 9 s.k. kvalitetsklasserna och dels varje enskilt märke
- Butiksgrupp: nio grupper av butiker enligt en gruppering som skapades för 25 år sedan, med bl.a. kända lågprisbutiker (typ outlets), varuhus, stora mångbutiksföretag, kända exklusiva butiker samt övriga.
- Tid: dels en dummy för år för att fånga effekten av nya butiksurval per år och dels en variabel för månad för att skatta inflationen.

Indikatorn Inflation i tabellerna är skattningen av månadsvariabeln i modellen, vilket blir en tillämpning av time-dummy-metoden, skulle man kunna säga.

Tabell 4a Regressionsmodeller med data för damkläder 2016 – 2018. Endast de 28 varumärken som ger 70 % täckning av antalet prisobservationer.

Förklarande variabler	Adj.R <sup>2</sup>	Inflation
Alla	86,73	2,75
Alla utom butiker	86,71	2,79
Alla utom fysiska attribut	84,46	2,85
9 Märkesklasser i stället för 28 märken	82,45	1,06
Alla utom märken överhuvudtaget	66,86	-0,85

Tabell 4b Regressionsmodeller med data för damkläder 2016 – 2018. Endast de 46 varumärken som ger 80 % täckning av antalet prisobservationer.

Förklarande variabler	Adj.R <sup>2</sup>	Inflation
Alla	87,89	2,40
Alla utom butiker	87,87	2,44
Alla utom fysiska attribut	85,91	2,32
9 Märkesklasser i stället för 46 märken	79,44	0,26
Alla utom märken överhuvudtaget	65,99	-0,64

Tabell 4c Regressionsmodeller med data för damkläder 2016 – 2018. Endast de 89 varumärken som ger 90 % täckning av antalet prisobservationer.

Förklarande variabler	Adj.R <sup>2</sup>	Inflation
Alla	88,79	2,76
Alla utom butiker	88,75	2,87
Alla utom fysiska attribut	86,97	2,76
9 Märkesklasser i stället för 89 märken	83,01	1,20
Alla utom märken överhuvudtaget	65,17	0,35

Tabell 4d Regressionsmodeller med data för damkläder 2016 – 2018. Endast de 150 varumärken som ger 95 % täckning av antalet prisobservationer.

Förklarande variabler	Adj.R <sup>2</sup>	Inflation
Alla	89,53	2,63
Alla utom butiker	89,50	2,71
Alla utom fysiska attribut	87,85	2,66
9 Märkesklasser i stället för 150 märken	83,41	0,43
Alla utom märken överhuvudtaget	65,49	-1,12

Vi ser mycket tydligt följande:

1. Varumärke har hög förklaringsgrad på variabeln pris
2. De nio klasserna av varumärken som SCB idag använder ger mycket lägre förklaringsgrad än om varje vanligt förekommande märke får vara med i modellen.
3. Effekten av butiker i 8 grupper är helt försumbar. Detta tyder på att de flesta butiker sätter priser enligt någon tillgänglig rekommendation så att de för exakt samma plagg sätter samma pris. Det finns naturligtvis korrelation mellan butikgruppen stora mångbutiksförtag och deras egna varumärken.

Dessa iakttagelser är avgörande för att gå vidare med den skissade nya metoden.

## 6 Diskussion

Konsumenter väljer i allt större omfattning andra säljkanaler än fysiska butiken för inköp av kläder (Carlsson och Strandberg, 2018). Med anledning av detta utreder vi om en re-design av



undersökningen av kläder är möjlig eller rent av nödvändig. Preliminära analyser visar att det kan vara möjligt att frångå första stegets urval av försäljningsställen och i stället utgå från ett urval av de viktigaste varumärkena.

För att säkerställa att en ny design inte skulle innebära en försämring av kvaliteten i undersökningen jämför vi den statistiska tillförlitligheten i de båda metoderna. Vi utgår från kvalitetsbegreppet i Sveriges officiella statistik.

### 6.1 Osäkerhetskomponent Urval

Gammal metod	Ny metod
<ul style="list-style-type: none"> <li>• Sannolikhetsurval av butiker</li> <li>• Godtyckligt urval av produkter</li> <li>• Subjektivt urval av pe av intervjuare i butik.</li> <li>• Variansen i priskvoter genereras till största delen när intervjuaren väljer pe.</li> </ul>	<ul style="list-style-type: none"> <li>• Cutoff-urval av de mest förekommande varumärkena</li> <li>• Sannolikhetsbaserad urvalsmetod för webbutik</li> <li>• Eventuellt inslag av slumpmässigt val av plagg, alternativt medvetet val av s.k. basplagg</li> </ul>

Givet central påverkad prissättning för alla varumärken kan vi frångå prisinsamling i fysiska butiker och istället hämta priser via internet. Den besparing som kan göras i prisinsamlingen kan återinvesteras i undersökningen genom att utöka urvalsstorlekarna och på så sätt minska variansen.

## 6.2 Osäkerhetskomponent Ramtäckning

Gammal metod	Ny metod
<ul style="list-style-type: none"> <li>Urvalsramen för butiksurvalet har en medveten undertäckning av 7,5 %</li> <li>Specificerade produktgrupper lämnar en andel kläder utanför mätningen</li> </ul>	<ul style="list-style-type: none"> <li>Cutoff-urval av de 70 – 80 % mest frekventa varumärkena</li> <li>Specificerade produktgrupper lämnar en andel kläder utanför mätningen</li> </ul>

Enligt gällande metodstrategi skall alla metoder vila på vetenskaplig grund. Det är tveksamt om cutoff-urval kan tillskrivas den egenskapen. Metoden är ändå flitigt tillämpad av SCB, i de situationer informationen man utesluter bedöms sakna betydelse för det slutliga resultatet. Bedömningen är att ett cutoff-urval kan motiveras i detta fall.

## 6.3 Osäkerhetskomponent Mätning

Gammal metod	Ny metod
Intervjuarna noterar plaggets pris och egenskaper med handdator.	Hittar vi alla egenskaper på webben?

Mätningar på nätet bygger på att det är möjligt att hitta samma vara så länge den finns tillgänglig i webbutiken. I de fall då artikelnummer eller modellnamn finns angivet är det enkelt. I annat fall gäller att på samma sätt som i prismätningar i fysisk butik, noggrant beskriva plagget så att man säkert vet att man hittar rätt från månad till månad.

Det finns farhågor att alla egenskaper inte finns tillgängliga i de produktbeskrivningar som finns i webbutikerna. Det återstår att se i vilken utsträckning detta är ett problem.

## 6.4 Osäkerhetskomponent Bortfall

Gammal metod	Ny metod
Obetydligt	Obetydligt

Vi har ingen anledning att tro att bortfallet är större i någon av modellerna.

## 6.5 Osäkerhetskomponent Bearbetning

Gammal metod	Ny metod
Produktionssystemet Pi09 samt hedoniska analyser	Produktionssystemet Pi09 samt hedoniska analyser

Inga skillnader

## 6.6 Osäkerhetskomponent Modellantaganden

Gammal metod	Ny metod
<ul style="list-style-type: none"> <li>Klassificeringen av tusentalet varumärken i nio s.k. kvalitetsklasser är en svaghet. Enligt vilka kriterier kan detta göras och vilken expertis måste vi ha?</li> <li>De hedoniska modellerna har inte så hög förklaringsgrad för de fysiska egenskaperna.</li> </ul>	<ul style="list-style-type: none"> <li>De hedoniska modellerna utnyttjar variabel Varumärke fullt ut vilket ger bättre anpassning av modellen.</li> <li>Kvalitetsjustering vid byten av produkter behöver inte göras med avseende på varumärke då ersättning alltid har samma varumärke.</li> </ul>

Regressionen som används för hedonisk kvalitetsjustering skulle bli säkrare med den nya designen. I stället för en osäker indelning av varumärkena i nio grupper hanterar modellen varje märke var för sig. Effekten av den fysiska butiken behöver heller inte förklaras i regressionen. Vi förväntar oss därför att kvalitén i kvalitetsjusteringarna kommer bli högre.

## 7 Förslag

För att troliggöra att de antaganden som ligger till grund för den nya design idén är korrekta föreslår vi en empirisk undersökning. Med start i basen 2019 (december 2018) och sex månader framåt vill vi genomföra provmätningar enligt:

- Provmätningen innefattar fyra produktgrupper av olika karaktär: kjol, jeans, herrjacka och träningsbyxa
- För varje produktgrupp prismäts fyra varumärken
- För varje produktgrupp och varumärke väljs slumpmässigt fyra webbbutiker
- Fyra pe väljs per produktgrupp, varumärke och webbbutik, helst exakt samma.

Resursåtgång skall antecknas för framtida allokering av antalet varumärken per plagg, antalet webbbutiker per kombinationen varumärke och plagg, samt antal pe per plagg, varumärke och webbbutik.

Antagandet om hög grad av centralt påverkad/styrd prissättning är viktigt. Detta studeras genom att analysera de fyra priserna i de olika webbbutikerna när plaggen är identiska. Om antagandet stämmer är butiksdimensionen oviktig.

Nämnden inbjuds att komma med synpunkter på den föreslagna designen samt den föreslagna provmätningen.

## 8 Referenser

Carlsson och Strandberg (2018) ”KPI:s urval och prismätning i en digital ekonomi – nuläge, utmaningar och möjligheter” presenterat vid nämndens möte nr 5

Norberg, Strandberg och Ståhl (2018) ”En modell för resursallokering inom KPI” presenterat vid nämndens möte nr 4

## Appendix

### A1: Kvalitetsbegreppet i Sveriges officiella statistik

Av Statistiska centralbyråns föreskrifter (SCB-FS 2016:17) om kvalitet för den officiella statistiken följer att kvalitetsbegreppet har fem huvudkomponenter som i sin tur är uppdelade i underkomponenter och, i förekommande fall, delkomponenter enligt nedan.

- Relevans
  - Ändamål och informationsbehov
    - Statistikens ändamål
    - Statistikanvändares informationsbehov
  - Statistikens innehåll
    - Objekt och population
    - Variabler
    - Statistiska mått
    - Redovisningsgrupper
    - Referenstider
- Tillförlitlighet
  - Tillförlitlighet totalt
  - Osäkerhetskällor
    - Urval
    - Ramtäckning
    - Mätning
    - Bortfall
    - Bearbetning
    - Modellantaganden
  - Preliminär statistik jämförd med slutlig
- Aktualitet och punktlighet
  - Framställningstid
  - Frekvens
  - Punktlighet
- Tillgänglighet och tydlighet
  - Tillgång till statistiken
  - Möjlighet till ytterligare statistik
  - Presentation
  - Dokumentation
- Jämförbarhet och sam användbarhet
  - Jämförbarhet över tid
  - Jämförbarhet mellan grupper
  - Sam användbarhet i övrigt
  - Numerisk överensstämmelse