Ekonomisk statistik och analys
Konsumentpriser
Martina Sundström
Can Tongur

PM till nämnden för KPI
Sammanträde nr 20
2025-05-14

Fortsatt arbete med kläder och skor 2025

I föreliggande projektrapport redogörs för arbetet med kläder och skor i KPI/ HIKP inom ramen för det av Eurostat finansierade projekt som genomfördes 2023–2025 och nyligen färdigställdes. Fokus i projektet har varit på att identifiera förutsättningar för användning av transaktionsdata för beräkning av prisindex för kläder och skor utifrån flertalet tillgängliga datakällor.

Medfinansieras av Europeiska unionen

Om projektet

Sektionen för Konsumentpriser har under två års tid genomfört ett projektarbete inom kläder och skor i KPI/HIKP med avsikten att identifiera förutsättningar för användningen av transaktionsdata för prisindex.

Som en del i arbetet att modernisera prisinsamlingen av kläder och skor i KPI/HIKP presenterades till nämndens sammanträde våren 2024 en promemoria om förutsättningarna kring urvalsarbetet och övergången till helt digitaliserad insamling. I promemorian lyftes de kvarvarande frågorna vid förflyttningen av prisinsamlingen från de återstående fåtal fysiska butiksbesöken till insamling enbart via internet (Öhman och Tongur, 2024). Detta har varit en del i SCB:s effektivisering av prisinsamlingen som alltså inte längre sker fysiskt överhuvudtaget från och med 2025.

I föreliggande studie redogörs kortfattat för det arbete som genomförts inom ramen för det nämnda projektarbetet och som delvis finansierats av Eurostat. Avsikten har varit att utreda förutsättningarna för användandet av transaktionsdata för framställning av prisindex för kläder och skor i stället för nuvarande prisinsamling via internet, alltså helt utan förbehållet att den nuvarande ansatsen med hedoniska kvalitetsvärderingar och insamling av plaggens egenskaper nödvändigtvis måste vara tillämplig.

SCB

Som även redogjordes av Öhman och Tongur (2024) finns inom kläder och skor ett inte försumbart antal olika stora aktörer inom olika delar/segment av marknaden som erbjuder damkläder, barnkläder, herrkläder, olika typer av skor (i termer av användningsområden och prisklasser). Därtill förekommer omfattande försäljning av denna typ av varor i såväl sportaffärer som varuhus, vilket sammantaget ger bilden av att marknaden är tämligen fragmenterad.

Då SCB har tillgång till transaktionsdata från ett antal av kedjorna inom kläder och skor kunde dessa data utredas närmare i projektet. En förutsättning som kunde noteras tidigt, nära kopplat till klassificeringsarbetet, var det varierande (och ibland ganska begränsade) informationsinnehållet om de enskilda produkterna vilka vanligen kännetecknas av 1) kort livslängd på galgen och 2) skarpa inslag av både säsong och mode. Detta kan kontrasteras med den tydligt definierade manuella mätprocess som hittills tillämpats (må det vara fysiskt i butik eller via hemsidor, internet) genom att prisinsamlare följt enskilda plagg ett antal månader och vars egenskaper (märke och vissa attribut) manuellt identifierats för beaktande vid byten.

Utifrån de analyser av transaktionsdata som gjordes i projektarbetet identifierades bland annat följande huvudsakliga områden som väsentliga för användandet av dessa data:

- Transaktionsdata måste klassificeras grundligt utifrån varierande förutsättningar per datamängd och inte alltid utifrån typiska hierarkier/kategoriseringar som i enlighet med COICOP.

- Produktbyten måste hanteras (se avsnittet ***item turnover***) – detta förutsätter en effektiv metod för indexberäkning i stället för nuvarande hedoniska ansats.

- Kategoriförändringar måste hanteras (se avsnittet ***category turnover***) i den takt som bedöms effektivt och i nära koppling till den valda indexberäkningsmetoden.

- Olika datamängder (och produkter) kan innebära olika behov av filtrering och lagring (se avsnittet ***filtering and subsequent data storage***)

För att kunna använda transaktionsdata i projektet behövde det först klassificeras enligt COICOP. Detta var väsentligt och föregick i naturlig ordning övriga steg i arbetet, dvs. före eventuell filtrering och före indexberäkningar. I projektet prövades olika nivåer av hierarkin i respektive datamängd för beräkning av den befintliga ansatsen med Jevons index, utan explicita kvalitetsvärderingar som annars tillämpas genom den hedoniska ansatsen.

Utifrån nämnda arbete och analyserna däri kunde det inses att en implementering av transaktionsdata inom kläder och skor skulle innebära en betydande omställning från den nuvarande ansatsen med begränsade urval och manuell insamling av i förväg definierade egenskaper hos plaggen (kläder/skor) för de hedoniska kvalitetsvärderingar till en desto mer integrerad/datadriven ansats per informationsmängd (datakälla), i förekommande fall med en väsentlig och löpande klassificeringsprocess. Detta skulle behöva vara genomförbart jämte övrig (manuell) insamling med nuvarande kvalitetsvärderingansats, alltså med parallella produktionsmetoder för statistikframställan. Processen skulle dessutom behöva skalas upp i takt med ökad tillgång till data från fler kedjor.

Referenser

Öhman, S. och Tongur, C. (2024). *Prismätningar för kläder och skor*. Stockholm. Promemoria till nämnden för konsumentprisindex, sammanträde nr. 18. SCB. https://www.scb.se/contentassets/1b48f2064ebd46a78eda4d68d51c0403/18/7.-forutsattningar-for-prisinsamling-av-coicop-03.pdf [Hämtad 2025-05-12]

Modernizing the clothing and footwear component of the Swedish HICP

Eurostat grant project – Summary report

Project: 101103449 – 2022 – SE - PRS

Topic: SMP-ESS-2022-PRS-H1-5996-IBA

Project background

Over the last two decades the Swedish HICP has made a transition to include more and more transaction data. At the very beginning of this transition, in early 2010s, obtaining and implementing transaction data for COICOP 01 and 02 was prioritized. Following that, Statistics Sweden started to receive transaction data for pharmaceutical products, different types of traveling services such as railway travel and package holidays, and dental care as well as data on several other kinds of goods and services.

While no price observations in the Swedish HICP are collected in physical stores anymore as of 2025, clothing and footwear prices are still collected manually from the internet. One of the reasons for a still cumbersome price collection is the quality adjustment method for COICOP 03 through a hedonic regression.

The aim of this project has been to find a way to implement transaction data while at the same time safeguard or improve the quality of the indices for clothing and footwear.

Project work

Developing a hierarchical structure for transaction data

As one of the main tasks in the project work was to elaborate on the obtained transaction data on clothing and footwear from several retail chains, the first theoretical mission was to establish how the data could be used. Two main levels of aggregation have been elaborated on for the context, namely elementary product groups (EPG) and the micro aggregates (MA) in a suitable manner for each of the data sources used. These principles on aggregation necessary for the forthcoming adaptations to COICOP 2018 and the broader use of transaction data in the HICP have been established in a previous Eurostat granted project, "*Towards More Efficient Use of Transaction Data in the Swedish HICP"* (Ståhl, 2023a) and have been presented for the Swedish CPI Board, c.f. Ståhl (2021,2023b).

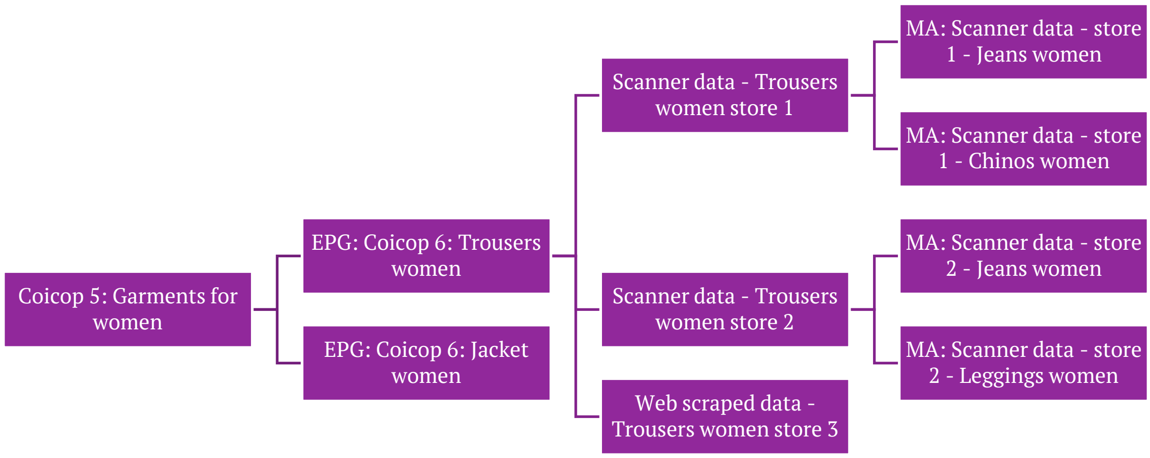
Elementary Product Groups

In creating a necessary hierarchical structure for COICOP 03 the upcoming COICOP 2018 had to be accounted for. As a first step, it was necessary to identify which EPG's that had to be added compared to the current structure in the CPI and HICP. These included dividing children's clothing into the two dimensions *boy* and *girl*.

Micro Aggregates

Definitionally, the Micro Aggregate (MA) is the assembly of similar items typically within a distinct retailer, i.e. an intuitive granularity. The aim has been to have conceptually similar MAs between the data sources, as described in an intuitive manner in Figure 1 below over a typical aggregation flow. Nevertheless, the reality of data requires to allow for both retailer-specific EPGs and, below that, also MAs that are specific to the retailer.

Figure 1: Example of a hierarchical structure covering the EPG and MA levels for aggregation.



Developing a database structure for storage of these transaction data

The initial step in the database work for the project was to design and create necessary tables in a dedicated database for this project. Number of variables (columns) and variable names were harmonized between data sources for the storage, whenever possible, while maintaining any integral differences necessary to preserve.

For this project work, the designated raw data was stored according to above mentioned structure as a first step, and then followed by a second storage step in which EPGs were assigned to all observations, based on several rules and manual review. Thus, after the raw data was stored in adapted tables, yearly and per data source, the necessary classification (EPG) per item as well as the necessary meta data information from the underlying physical files such as date and time stamps were stored in a secondary storage table. This process made cross-validations and re-classifications possible in a quality control loop.

Filtering and subsequent data storage

The data storage process for the project work has been set up in accordance with internal guidelines of Statistics Sweden regarding data storage for statistical purposes, with necessary adaptations for these specific transaction data on clothing and footwear.

In stages following initial storage, filtering and delimitations also come in as a necessary quality criterion and assessments had to be made on the extent of replicability needed at each phase of the statistics production process. This final stage of data storage, at which classifications and filtering have been made and are reproducible, is denoted “final observational register” in the terminology used by Statistics Sweden.

The current hedonic approach to quality adjustments

In clothing and footwear items, Statistics Sweden uses a hedonic quality adjustment method to estimate differences in quality between replacements and their base product. Necessary data on product characteristics is collected monthly, and brands are distributed into different quality classes for estimations.

Variables for physical characteristics tend to explain merely a small amount of the price variation in clothing, as explained by Norberg and Bäckström (2012). Typically, some 80 percent of total price variation in clothing data is explained by brand and outlet type, leaving some 20 percent variation unexplained. Product characteristics usually add merely 1-2 percentage points explanatory power to this, rarely more than so. In whole, this employed method for quality adjustments of clothing makes the cost for producing the corresponding indices relatively (and absolutely) high in comparison with other product groups.

The issue of the limited contribution from specific quality characteristics was also pointed out by Norberg and Strandberg (2018). In their study, they derived the effect from leaving out different parts of the hedonic model and re-estimated the inflation based on a varying set of most frequent brands, which altogether rendered a coverage of either 70, 80 or 90 percent of the CPI data for the years 2016, 2017 and 2018. It was observed that leaving out all quality characteristics while controlling for the outlet dimension (which is always given due to the fixed outlet sample) and estimating merely the effect from brand gave identical or almost identical inflation estimates for clothing in comparison with a full model including all quality characteristics, although with about 2 percentage points less explained variation (adjusted R2).

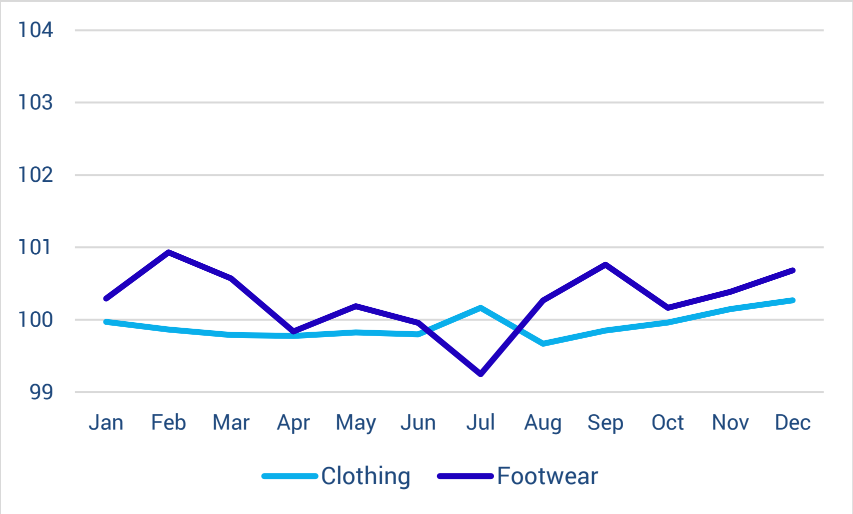
Implicit quality index

One way to measure the overall impact from quality adjustments is to calculate a so-called implicit quality index (IQI). The measure is a ratio between the price index without and with quality adjustments, and it formulates for any considered product group g some month m in year y as follows:

$$IQI_{m,y}^{g} = \left(\frac{price\ index\ without\ quality\ adjustments}{price\ indx\ with\ quality\ adjustments}\right)_{m,Y}^{g} \times 100.$$

The IQI estimate is communicated by Statistics Sweden in the annual quality assessment report, available from the CPI homepage (scb.se/KPI). An IQI value greater than 100 indicates that replacements made for the items have been estimated to correspond to an improvement in quality. The converse applies for values below 100. When equal to 100, the introduced replacement products have been estimated to be equivalent to the exchanged products.

Graph 1: Implicit Quality Index (IQI) for clothing and footwear, year 2023



Source: Swedish CPI IQI report for 2023

Clothes show a seasonal pattern where seasonally adapted products are replaced during the spring, which the hedonic model adjusts for as a deterioration in quality (i.e. padded jackets are replaced with un-padded jackets), causing the IQI to drop. Towards the autumn, the warmer jackets return to the stores which is estimated as quality improvements and the IQI rises towards 100.

Empirical findings in the project work

Data in the empirical analysis

Statistics Sweden has accumulated transaction data over the years for several product groups, also within clothing and footwear. For this project work, the years 2022, 2023 and 2024 were chosen, with December 2021 as base for 2022. Although the data stretches farther back than 2021, this specific period was deemed sufficient and supposedly not affected by the pandemic.

In the following, some aspects of the data are reported. A bilateral index approach was used to illustrate a simple application of these data under current methodology (Jevons).

Item turnover (churn)

In terms of item churn, the following table (Table 1a) shows the average share of newly introduced items per month, aggregated over clothing (03.1) and footwear (03.2). Items occurring in the very first month in the data from each retail chain were omitted to avoid a misleading peak from as mass of observations with potentially earlier entry.

Table 1a Average monthly "churn" in percent, i.e. the share of newly entered items per month

Month	**Jan.**	**Feb.**	**March**	**April**	**May**	**June**	**July**	**Aug.**	**Sept.**	**Oct.**	**Nov.**	**Dec.**
Churn*	5.15	14.34	11.39	10.0	9.44	7.07	6.66	8.41	8.39	7.81	6.78	5.6

*__Churn__ is computed as the share of newly introduced items a specific month related to all sold items, unweighted with respect to item turnover and COICOP. It is averaged over all included data (retail chains and years).

Table 1b Cumulative life span of items (in percent of all items)

Months	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Life*	12	20	28	35	44	52	60	66	71	76	79	82

*__Life (span)__ is computed as the share of items existing merely one month, two months, etc, reported here up to 12 months at which the series is truncated.

It is understood from the data, based on hundreds of thousands of items, as illustrated in Table 1b, that less than half the number of items have a shelf life of at most six months, with a concentration of one third between 1 and 4 months.

Category turnover

Additionally, there was observed a markable "churn" of categories, specifically at more detailed levels, such that categories had a "within-year" lifespan of a certain number of months – several of which were most likely seasonally affected. In the following Table 1c, category lifespan is shown conditional on the formation of the fixed basket with December base, i.e. the cumulative life span of categories during the basket year given existence in the base period.

Table 1c Cumulative life span in months expressed as percentage of categories given existence in the base period (December).

Months	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Life*	5	9	12	14	17	19	20	23	26	29	38	69

*__Life (span)__ is computed as the cumulative share of categories existing at most one month, two months,… and up to the cutting point at 12 months. Computations are conditional on the base period December preceding year. Results are averaged over all included data, i.e. all retail chains and years.

As can be interpreted from Table 1c, the mass of categories exists almost a year or longer (11-12 months and further), whereas a significant share (more than a third) enters and exits the data during the year, here conditional on the chosen base month December.

The following basket attrition is observed in Table 1d when employing a fixed basket on most detailed level (exact item matching) with base in December preceding year, i.e. the complete item set available in December *y-1*. Coverage is here expressed in terms of turnover used in the fixed basket in relation to all turnover in the month, reported as an average over retailers and all items (individually computed per Micro Aggregate and then averaged) and the years included in the study.

Table 1d Average basket coverage during the year from a fixed basked set in December from transaction data, aggregated over both COICOP classes (03.1, 03.2).

Month	**Jan.**	**Feb.**	**March**	**April**	**May**	**June**	**July**	**Aug.**	**Sept.**	**Oct.**	**Nov.**	**Dec.**
Coverage*	0.89	0.74	0.62	0.46	0.35	0.35	0.34	0.36	0.35	0.35	0.33	0.28

***Coverage** is the current month share of turnover for the remaining items as defined in the base period basket respectively to each of the twelve following month's total turnover, aggregated up from respective EPG per retail chain (i.e. Micro Aggregate). Subtracting the coverage from one (1) will render the cumulative attrition in terms of (concurrent) turnover. The reported averages in the table cover all included Micro Aggregates for the retail chains in the study and available years, respectively.

It can be understood from Table 1d that for clothing and footwear, the typical attrition is quite fast but flattens out after some six months – indicating that on average, one third of the sales comes from long-lasting items while two thirds of the turnover come from a fairly "perishable" set of items given that the starting point has been December for each year in the analysis.

Index computations

The bilateral Jevons l index was applied at three different homogenization levels for matching:

1. Item level, as given by the most detailed identificatory in each dataset,
2. Intermediate level, chosen appropriately for each retailer, and
3. EPG level

Levels 1 and 3 reflect the two possible extremes regarding homogeneity, from most specific to least specific (categorically) whereas level 2 is assumed to be somewhere in between those and necessarily adapted to the categorization structure per retailer. The employment of an intermediate level in aggregations still had to ascertain that individual articles therewithin would fall into correct EPG.

The finally chosen intermediate level in the respective categorization appropriate for each data source rendered the MARS scores in Table 2 according to the MARS procedure as suggested by Chessa (2019) and exemplified by Eurostat (2022).

Table 2 MARS scores (aggregated and averaged) for the most detailed and intermediate levels.

Month	**Jan.**	**Feb.**	**March**	**April**	**May**	**June**	**July**	**Aug.**	**Sept.**	**Oct.**	**Nov.**	**Dec.**
MARS 1	75.4	69.0	63.8	58.3	51.5	49.3	47.7	47.2	43.6	41.4	37.0	34.4
MARS 2	29.5	29.8	32.7	31.3	35.3	32.2	30.3	29.4	32.6	31.9	32.8	33.3

Nota bene: MARS 3 is EPG level, the least detailed level and has zero score by construction due to zero homogeneity while achieving full match (100%), rendering the outcome of 0*100=0, and is thus omitted from the table.
Intermediate level is MARS 2 and MARS 1 is the benchmark level 1, exact item matching.

For the three levels, both a weighted and an unweighted price index was computed and compared with the corresponding HICP price index through a distance measure, namely the mean absolute deviation (MAD). The Mean Absolute Deviation (MAD) formulates over a set *s* as the mean (1/s) of absolute deviations from a reference value *ref*:

$$\text{MAD} = \frac{1}{s}\sum_{i=1}^{S}\left|y_{i,t} - ref_i\right|$$

The MAD computations are reported in Table 3 below with corresponding HICP as reference value.

Table 3 Weighted (w) and Unweighted (u.w.) Mean Absolute Deviation **MAD(**w) and **MAD(**u.w)

COICOP	**Year**	Level 1		Level 2		Level 3	
		MAD(w)	**MAD(**u.w)	**MAD(**w)	**MAD(**u.w)	**MAD(**w)	**MAD(**u.w.)
03.1	2022	23.85	22.72	6.87	8.21	6.75	6.75
	2023	14.41	16.60	8.61	9.89	8.61	8.61
	2024	19.81	22.10	8.2	7.98	6.73	6.73
03.2	2022	18.48	22.16	10.76	15.59	11.94	11.94
	2023	10.13	12.59	7.15	5.38	3.72	3.72
	2024	8.13	14.04	8.24	7.15	9.68	9.68

The results in Table 3 indicate that the largest leap between aggregations comes when moving up form level 1, i.e. the decaying basket achieved with level 1 is undesirable in comparison with the less homogeneous but still relevant level 2, given the HICP reference value.

A showcase for the current hedonic approach with seasonal basket "separation"

In analogy with the intermediate matching approach (level 2) for the transaction data, a grouping approach applied to the price index data for three selected product groups for the years 2022, 2023 and 2024. The identification process for the showcase required text mining of stored information in the internal production databases to identify necessary item characteristics indicating seasonality. As a result, a stratified unit value was computed over two subsets (*strata*) of items based on a seasonality adaptation criterium:

Stratum 1 (seasonal): the replaced items that carry some seasonal adaptation and thus have a recomputed base price with respect to seasonality, and
Stratum 2 (standard): the remainder of items that are not seasonally adapted, i.e. items that do not comprising a change from/to a seasonal characteristic that affects the recomputed base price.

An implicit quality index (IQI) was computed from on the unit value index from the stratification and compared with an overall unit value index in which stratification was not employed for the product group in an attempt to illustrate seasonal categorization as stratification to achieve homogeneity.

For the analysis, the unit value approach was preferred over the existing mean of ratios (Jevons) as a means of homogenization with respect to the assessment made through counter seasonality while reducing the overall impact from all *other* characteristics that affect the hedonic recalculation per item at each replacement, assuming non-seasonal characteristics cancel out more readily.

Some results for the experimental calculations are summarized in Table 4 below.

Table 4 Quality adjustment parameters, IQI (December), and MAD with reference 100.

Product Group	**Quality adjustment parameters**	**Year**	**IQI (orig.)**	**IQI (u.v.)**	**MAD** (IQI)	**MAD** (IQI u.v.)
Women's skirt	Brand, fabric, lining/padding, length, others*	2022	96.6	99.8	2.40	0.73
		2023	100.7	100.2	0.55	0.46
		2024	98.5**	102.9	0.65	1.91
Women's dress	Brand, fabric, lining/padding, others*	2022	99.3	100	1.73	0.17
		2023	98.5	100.9	0.78	0.38
		2024	101.7**	100.2	0.63	0.33
Women's Jacket	Brand, fabric, down, lining/padding, length, others*	2022	99.6***	100	0.54	0.08
		2023	100	99.4	1.33	0.24
		2024	97.7**	100	1.97	0.14

Nota bene: *Others covers parameters that may be applicable in certain cases.
** Preliminary and embargoed until CPI Board Meeting in May 2025.
*** IQI values may differ slightly to preceding publications for accordance with this analysis.

Seen in Table 4, the stratification into a seasonally adapted basket and a standard basket may serve equally well as the hedonic adjustment, upon judging from the computed unit value implicit quality index IQI(u.v). The latter is in most cases closer to unit quality adjustments (100) than the original IQI as computed and published annually, with two exceptions: 1) 102.9 versus 98.5 for Women's skirt in 2024, and 2) 99.4 versus 100 for Women's jacket in 2023.

It is seen that, on average, the MAD obtained from the unit value IQI deviates less from 100 than the corresponding actual IQI, implying that the impact from stratification appears less intrusive than the (explicit) hedonic approach (c.f. MAD for the actual IQI).

Practical pitfalls of hedonic estimations

As the hedonic approach has been used for clothing since early 1990s at Statistics Sweden, much experience is gained on the operational implications of the method. The model is estimated every year to account for the price movements in clothing and footwear to incorporate new outlets and brands as well as relative changes between quality attributes.

One potential deficiency of the employed hedonic approach has been, since the very beginning, that the distribution of characteristics (or brands) has not been balanced in data collection. This could be thought of as a three-folded issue:

1) first, price collectors make replacements depending on what is available in the outlet and fits the product description, regardless of *potential* characteristics and regardless of *what was chosen in the base period*, thus allowing for an implicit "drift" both in terms of availability bias and potentially some level bias as outlets may opt to offer items based on profitability given that it fits their "product line" rather than maintaining specific quality attributes that are identified by Statistics Sweden.

2) second, brands may enter early in their "life span", when being promoted while crisp new on the market. This may cause their quality attributes to be downward biased in the course of time. This issue has been observed by Norberg and Bäckström (2012) and denoted *emerging brands* for which they also provide a solution that circumvents the brand grouping that is employed today.

3) third, given a certain characteristic, e.g. cotton or wool, there may exist different material qualities within each of these. Such unobservable effects pass unnoticed but are very likely affecting price, and, not the least, may be altered over time such that the same brand (or outlet) may switch quality but remain at a certain price, given both fabric (e.g. cotton) and brand. It is very common that there are typical “price points” in outlets (say 199 SEK or 499 SEK) rather than prices that exactly reflect a specific characteristic, hence quality may reflect price level instead of the opposite.

This latter point (3), that a nominally identical characteristic may in effect be completely differently valued between brands and over time, has been a consideration over the course of years. Statistics Sweden continuously monitors and reviews the automatic quality assessments whenever they are of significant value. This has become a concern over the years in terms of an unobservable component that may be correlated to some intrinsic quality level of the brand.

Project conclusions

The perhaps largest take-away from the empirical part of this project work is the necessity of finding an appropriate homogeneity level to group transaction data observations. This is especially the case if the statistical production process is to be automatized to any extent and will thus require elaboration specific to each data source and not is necessarily stable, as noticed in terms of category churn.

Although the issue of homogenization was merely applied in a bilateral context, similar challenges will exist in a multilateral approach. However, depending on context, the seasonality issue may then be resolved in an embedded manner through improved methodology.

Availability of transaction data in clothing and footwear implies the need of a future production environment capable of accommodating different data sources and computational methods, regardless of choice of index methodology. Such a mixed-mode approach needs to be implemented and supervised with precaution to maintain the balance of the collected items between retailers. The provided aggregation scheme in Figure 1 illustrates these issues.

The clothing and footwear markets are fairly fragmentized in Sweden, as was pointed out by Öhman and Tongur (2024) in a paper to the Swedish CPI Board. Consequently, implementation of transaction data will be beneficial in terms of replacing small sample manual price collection for a specific retail chain with full coverage data. But, due to this market fragmentation, having a limited set of retailers’ transaction data will imply that several production systems and/or facilities must be in use in parallel and, as experienced in this project, each implementation can be a cumbersome process.

Statistics Sweden has had a web scraping solution for price statistics during a few years and efforts have been made to use it for clothing. However, due to continuous technical obstacles of varying sorts, web scraping has turned into a very limited and unreliable approach and could not deliver data for exploration in this project.

Another assessment made in this project work has been the assessment of seasonal products versus other characteristics, i.e. the very fundamental parts that motivate the use of a hedonic method. As realized, both through the preceding papers referenced and the annually published IQI, there is perhaps little gain from specific characteristics to which price differences are attributable whereas seasonality can be resolved through stratification.

It can be seen in the HICP inventory that especially direct comparisons appear highly frequent, sometime accompanied by bridged overlap, supporting the idea in the empirical showcase with seasonal stratification to cancel out effects from specific characteristics.

References

Chessa, A. G. (2019). “MARS: a method for defining products and linking barcodes of item relaunches”, paper presented at the New Techniques and Technologies for Statistics 2019 conference, Brussels, Belgium, 12–14 March.

Eurostat. 2022. “Guide on Multilateral Methods in the Harmonised Index of Consumer Prices - 2022 edition.” Eurostat methodological guide. DOI: https://doi.org/10.2785/873932.

Hansson Bierbum, L. and Ottosson M. 2024, ”Kvalitetsvärderingsrapport 2023”. Memo written for Statistics Sweden’s CPI Advisory board, meeting no 18, 2024 [in Swedish]. Available at: https://www.scb.se/contentassets/1b48f2064ebd46a78eda4d68d51c0403/18/3.-kvalitetsvarderingsrapport-2023.pdf

Norberg, A. and Bäckström, P. 2012. ”Beklädnadsvaror i KPI: Varumärken i de hedoniska modellerna”. Memo written for Statistics Sweden’s CPI Advisory board, meeting no 246, 2012 [Ingress in Swedish, paper in English]. Available at: https://www.scb.se/contentassets/1b48f2064ebd46a78eda4d68d51c0403/bekladnadsvaror-kpi.pdf

Norberg, A. and Strandberg, K. 2018. ”Idé om ny design för KPI kläder”. Memo written for Statistics Sweden’s CPI Advisory board, meeting no 5, 2018 [in Swedish]. Available at: https://www.scb.se/contentassets/1b48f2064ebd46a78eda4d68d51c0403/klader-i-kpi.pdf

Ståhl, O. 2021. ”Aggregeringsprinciper inom COICOP 01 och 02”. Memo written for Statistics Sweden’s CPI Advisory board, meeting no 12, 2021 [in Swedish]. Available at: https://www.scb.se/contentassets/1b48f2064ebd46a78eda4d68d51c0403/aggregeringsprinciper-inom-coicop-01-och-02.pdf

Ståhl, O. 2023a. “Towards More Efficient Use of Transaction Data in the Swedish HICP”. Project summary report for the Eurostat grant work [101034458–2020–SE–HICP].

Ståhl, O. 2023b. “Metodutredning för COICOP 01 och 02 – Projektrapport”. Memo written for Statistics Sweden’s CPI Advisory board, meeting no 15, 2023 [in Swedish]. Available at: https://www.scb.se/contentassets/1b48f2064ebd46a78eda4d68d51c0403/16/metodutredning-coicop-01-och-02.pdf

Öhman, S. and Tongur, C. 2024. “Prismätningar för kläder och skor”. Memo written for Statistics Sweden’s CPI Advisory board, meeting no 18, 2024 [in Swedish]. Available at: https://www.scb.se/contentassets/1b48f2064ebd46a78eda4d68d51c0403/18/7.-forutsattningar-for-prisinsamling-av-coicop-03.pdf