
Údaje zo skenerov a cenové indexy

Helena Glaser-Opitzová¹

Abstrakt

Článok sa zaoberá metodologickými aj praktickými otázkami implementácie nových zdrojov údajov, tzv. údajov zo skenerov do produkcie cenovej štatistiky a zároveň prezentuje výsledky experimentálnej štúdie, ktorá bola vykonaná na reálnych transakčných údajoch piatich obchodných reťazcov. Poukazuje na výhody ich použitia, ale zároveň upozorňuje aj na možné riziká a nezodpovedané otázky najmä vo vzťahu k výberu vhodného indexu na úrovni elementárneho agregátu.

Kľúčové slová

index spotrebiteľských cien, údaje zo skenerov, elementárne cenové indexy

Abstract

The article deals with methodological and practical issues of implementation of new data sources, so-called scanner data, into the production of price statistics and at the same time presents the results of an experimental study that was conducted on real transaction data of five retail chains. It highlights the advantages of their use, but also draws attention to possible risks and unanswered questions, particularly in relation to the selection of an appropriate index at the elementary aggregate level.

Key words

consumer price index, scanner data, elementary price indices

JEL classification

E 43, E 31

1 Úvod

Inflácia a presnosť merania inflácie významne ovplyvňuje nielen oblasť podnikania, verejnú a štátnu správu, ale zohráva významnú rolu aj v každodennom živote, pretože inflácia ovplyvňuje hodnotu peňazí každého z nás a znižuje životnú úroveň obyvateľstva. Spôsobuje tiež ekonomickú nerovnováhu krajiny, vyvoláva špekulačné investície a v neposlednom rade vyvoláva výkyvy v menových kurzoch.

Okrem indexu cien výrobcov a deflátoru HDP sa za najznámejší ukazovateľ merania inflácie považuje index spotrebiteľských cien (CPI), ktorý v podmienkach Slovenskej republiky zostavuje a publikuje Štatistický úrad SR (ŠÚ SR) v mesačnej periodicite. Pri meraní inflácie prostredníctvom CPI sa zohľadňujú všetky tovary a služby, ktoré domácnosti spotrebúvajú, t. j. predmety bežnej spotreby, predmety dlhodobej spotreby a služby. V súčasnosti zostavenie tohto ukazovateľa postupne prechádza významnou inovačnou zmenou z pohľadu získavania nových zdrojov údajov a metód ich spracovania tak, aby informácie o cenách a následne štatistika merania inflácie lepšie odrážali skutočné zmeny v hospodárstve, boli spoľahlivejšie, prípadne včasnejšie a podrobnejšie.

¹ Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky, Katedra štatistiky, Dolnozemska cesta 1, 852 35 Bratislava, helena.glaser-opitzova@statistics.sk.

Keď hovoríme o nových zdrojoch údajov pre cenovú štatistiku, máme na mysli transakčné údaje obchodných reťazcov tzv. údaje zo skenerov a údaje získané z webových stránok (web-scraped data).

Údaje získané z webových stránok maloobchodných predajcov nám môžu okrem ceny poskytnúť aj množstvo doplňujúcich informácií o kvalitatívnych vlastnostiach produktov. Tieto informácie nám môžu pomôcť lepšie oddeliť zmenu ceny, ktorá je založená len na zmene kvality výrobku od čistej cenovej zmeny a získať tak ceny porovnateľné v čase.

Údaje zo skenerov sú podrobné údaje o predaji spotrebného tovaru a poskytujú informácie na úrovni čiarového kódu alebo presnejšie na úrovni primárneho identifikačného čísla GS1² pre výrobky alebo služby označovaného ako GTIN (global trade item number) predtým známeho ako EAN kód. Údaje zo skenerov od konkrétneho predajcu a za dané časové obdobie predstavujú vyčerpávajúci zoznam všetkých kódov položiek³, ich tržieb a predaných množstiev. Umožňujú zostaviť index zo všetkých transakcií predajcu alebo predajne.

Údaje zo skenerov existujú už niekoľko desaťročí⁴ a ich praktický význam pre oficiálnu štatistiku je v súčasnej dobe už evidentný. Údaje zo skenerov ponúkajú podrobnejšie a presnejšie informácie, pretože ide o veľký počet cenových pozorovaní, môžu napomôcť zvýšeniu efektívnosti a zníženiu zaťaženia respondentov štatistických zisťovaní. Údaje zo skenerov môžu dopĺňať a v niektorých prípadoch aj nahrádzať tradičné prístupy k zberu údajov za účelom zostavovania CPI. Používanie údajov zo skenerov spochybňuje niektoré tradičné interpretácie a aplikácie pojmov a vyvoláva metodologické aj praktické otázky súvisiace s tvorbou CPI, na ktoré sa v článku zameriavame a vybrané teoretické úvahy overujeme v praxi na transakčných údajoch piatich najväčších obchodných reťazcov z oblasti potravín a nealkoholických nápojov, ktoré sú ŠÚ SR poskytované v týždennej periodicite na základe uzatvorených dohôd.

Cieľom príspevku bolo syntetizovať poznatky súvisiace s problematikou implementácie údajov zo skenerov do produkcie cenovej štatistiky a následne porovnať získané poznatky, koncepty a metódy na reálnych údajoch.

2 Spotrebiteľské ceny – súčasný stav

Ako vyplýva zo správy o kvalite, ktorá sa týka zostavovania CPI, a ktorú ŠÚ SR zverejňuje na svojej internetovej stránke (www.statistics.sk), sledovanie vývoja spotrebiteľských cien a zostavovanie CPI/HICP⁵ vykonáva ŠÚ SR z údajov získavaných prostredníctvom terénneho zberu údajov o cenách tovarov a služieb. Zber údajov sa realizuje priamo v prevádzkach a obchodoch na celom území SR, kde obyvatelia obvykle nakupujú a ceny, ktoré sa zisťujú, sú tzv. pultové ceny. Ceny sa zisťujú v spolupráci so zamestnancami spravodajských jednotiek⁶. Ceny niektorých tovarov a služieb je možné zisťovať aj telefonicky, prostredníctvom internetu alebo centrálné, ak sa jedná o ceny administratívne, regulované, sadzobníkové, ale aj položky s trhovými cenami, ktoré sú rovnaké pre celé územie SR. Predané množstvá jednotlivých tovarov a služieb nie sú k dispozícii a tradičný spotrebný kôš, na ktorom

² Čiarový kód GS1 spustil digitálnu revolúciu, ktorá navždy zmenila spôsob, akým svet podniká. Pomocou jednoduchého skenovania bolo možné identifikovať produkt a pripojiť ho k počítačovému systému

(<https://www.gs1.org/about>)

³ V praxi môžu byť používané aj iné kódy ako GTIN, preto sme v tomto kontexte zaviedli všeobecný pojem "kód položky".

⁴ Prvý snímač čiarových kódov bol zavedený v USA v roku 1974.

⁵ V rámci Európskej únie (EÚ) bol zavedený osobitný index spotrebiteľských cien - harmonizovaný index spotrebiteľských cien (HICP). HICP sa zostavuje podľa harmonizovaného prístupu a jednotného súboru definícií. <https://ec.europa.eu/eurostat/web/hicp/legislation>.

⁶ Spravodajskou jednotkou je každý, od ktorého sa požaduje poskytnutie údajov na štatistický účel.

je založený výpočet indexu spotrebiteľských cien je relatívne malá vzorka z celej množiny tovarov a služieb (tzv. cenové reprezentanty). Ceny sú zisťované počas prvých 20 dní sledovaného mesiaca. Za cenové reprezentanty boli vybrané výrobky a služby, ktoré sa významne podieľajú na výdavkoch obyvateľstva a svojim rozsahom reprezentujú celú sféru spotreby. V súlade s metodikou Eurostatu sa univerzálny spotrebný kôš člení na 12 odborov a 44 skupín podľa klasifikácie ECOICOP⁷. Od januára 2022 univerzálny spotrebný kôš predstavuje súbor 738 reprezentantov. Z uvedeného počtu patrí do odboru „Potraviny a nealkoholické nápoje“ 146 reprezentantov. Váhy jednotlivých reprezentantov sa počítajú z údajov zistených štatistikou rodinných účtov, z podkladov štatistiky národných účtov a rôznych administratívnych zdrojov údajov. Index spotrebiteľských cien je indexom Laspeyresovho typu (ŠÚ SR, 2021):

$$I = \frac{\sum_{i \in N} \frac{p_{i1}}{p_{i0}} p_{i0} q_{i0}}{\sum_{i \in N} p_{i0} q_{i0}} 100 \quad (1)$$

kde:

p_{i1} je cena tovaru alebo služby v sledovanom období,

p_{i0} je cena tovaru alebo služby v základnom období,

q_{i0} predstavuje predané množstvo tovaru alebo služby v základnom období,

$\frac{p_{i1}}{p_{i0}}$ je individuálny index ceny určitého tovaru alebo služby a

N je spotrebný kôš produktov a služieb.

S výnimkou výrobkov s volatílnymi cenami, ceny získavané tradičným spôsobom sa zbierajú raz mesačne a len pre výber predajní. Aktuálny spotrebný kôš (popis a váhy jednotlivých odborov, skupín, tried, podtried medzinárodnej klasifikácie individuálnej spotreby podľa účelu (COICOP); pod tým popis a váhy jednotlivých reprezentantov) je publikovaný na stránke ŠÚ SR: www.statistics.sk. V rámci ponuky hlavnej hornej lišty Štatistiky > Makroekonomické štatistiky > Spotrebiteľské ceny a ceny produkčných štatistik > Ukazovatele > Spotrebiteľské ceny (inflácia) je možné nájsť spotrebné koše od roku 2013 vrátane súčasného.

3 Využitie údajov zo skenerov a súvisiace metodické zmeny v zostavovaní CPI

3.1 Vlastnosti údajov zo skenerov a ich dopad na produkciu CPI/HICP

Údaje zo skenerov od konkrétneho predajcu a za dané časové obdobie predstavujú vyčerpávajúci zoznam všetkých kódov predaných položiek, ich tržieb a predaných množstiev. Umožňujú zostaviť index zo všetkých transakcií predajcu alebo predajne. Napríklad sortiment potravín a nealkoholických nápojov, ktorý bol predmetom nášho skúmania je pokrytý v podmienkach Slovenska počtom 7000 až 29000 položiek, v závislosti od maloobchodného reťazca. Na základe metodických vysvetliviek CPI zverejnených na stránke ŠÚ SR (ŠÚ SR 2021) je pri súčasnom spôsobe zisťovania cien každá z vybraných predajní navštívená jedenkrát v mesiaci a za každého reprezentanta vybraný v predajni len jeden produkt (spolu cca 142 konkrétnych cien). Oproti tomu pri údajoch zo skenerov získame informáciu za celý predaj daného produktu v týždennom agregáte. Na základe tržieb vieme jednotlivým položkám priradiť váhu.

Údaje zo skenerov ponúkajú teda väčšie časové a produktové pokrytie v porovnaní s tradičnými metódami zberu údajov, poskytujú informácie o transakciách pre všetky výrobky

⁷ Európska klasifikácia individuálnej spotreby podľa účelu.

vo všetkých časových obdobiach. Poskytujú informácie o transakčných cenách, a nie o pultových cenách. To znamená, že môžeme sledovať priemernú cenu, ktorú spotrebiteľia skutočne zaplatili za každý výrobok. V porovnaní so súčasným tradičným zberom ponúkajú aj väčšie regionálne a produktové pokrytie a z hľadiska tržieb v skúmanej oblasti v podmienkach SR predstavujú pokrytie približne 80%. Po počiatkových investíciách sa v porovnaní s tradičným zberom údaje náklady štatistických úradov znížia, pretože odosielanie transakčných údajov môže byť plne automatizované.

Ďalšou výhodou týchto údajov je veľká miera detailných informácií o jednotlivých produktoch⁸, ktorá umožňuje definovať tzv. homogénne skupiny produktov a vypočítať cenové indexy elementárnych agregátov na nižšej úrovni agregácie ako je ECOICOP5 (viď nasledujúca kapitola). V údajoch je zaznamenaná každá transakcia a viditeľný je vznik nových kódov položiek, zánik položiek a zmeny ich relatívnej dôležitosti.

V Praktickej príručke Eurostatu na spracovanie údajov zo skenerov (Eurostat, 2017, str. 9) sa môžeme dočítať: „*Pri tradičnom zisťovaní cien musia cenári dôverovať intuícii a zdravému rozumu a môže sa stať, že ceny sa zbierajú dovtedy, kým je položka dostupná, aj keď už nie je reprezentatívna. V transakčných údajoch od obchodných reťazcov (údaje zo skenerov) je zaručená reprezentatívnosť*“.

Podľa Eurostat (2017) v priebehu roka zaniká 25% až 60% kódov položiek, samozrejme v závislosti od krajiny a sledovanej oblasti spotreby. Miznúce položky môžu byť nahradené novými, pričom úplne nové položky môžu v priebehu roka rozšíriť predávaný sortiment výrobkov. Položky, ktoré sa predávajú počas celého roka, môžu tvoriť iba podmnožinou celkovej populácie. Môžu tiež zmiznúť navždy alebo dočasne (napr. v prípade sezónneho tovaru). Okrem vzniku skutočne nových tovarových položiek sú položky často nahrádzané novými, obnovenými verziami tovarov. Jedná sa v podstate o rovnaké tovarové položky s nejakým povrchným rozdielom, ako je napríklad nový obal. Vzniká takto nielen nový kód položky bez zmeny kvality, ale aj problém týkajúci sa spracovania údajov.

Aj keď existuje veľa výhod v súvislosti s používaním údajov zo skenerov pre zostavovanie CPI/HICP, je potrebné spomenúť aj niekoľko problémov, ktoré musel ŠÚ SR v praxi riešiť. Použitie týchto údajov predstavuje dodatočnú pracovnú záťaž, ako aj počiatkové náklady na vybudovanie IT systému na spracovanie údajov. Bolo potrebné vybudovať kanál pre prenos údajov do informačného systému ŠÚ SR, čo v počiatkovej fáze tiež vyžadovalo finančné a ľudské zdroje. Implementáciou týchto zdrojov údajov do produkcie cenovej štatistiky sa zvýši závislosť ŠÚ SR na jednotlivých maloobchodníkoch, pretože nedodanie údajov by mohlo mať významný dopad na kvalitu štatistických produktov. Toto potenciálne riziko sa zmiernilo uzatvorením kvalitných dohôd a udržiavaním dobrých vzťahov s poskytovateľmi údajov avšak v praxi aj tak občas dochádza k výpadkom buď vo forme poskytnutia chybných súborov alebo nedodania údajov v požadovanom termíne. V neposlednom rade, implementácia nových zdrojov údajov komplikuje proces produkcie CPI/HICP z pohľadu ľudských zdrojov, vyžaduje si pokročilejšie IT kompetencie zamestnancov.

3.2 Spotrebný kôš a elementárny agregát

Spotrebný kôš vo všeobecnosti pozostáva z dvoch úrovní. Úroveň I predstavuje skupiny produktov a ich príslušných váh, ktoré sú usporiadané podľa klasifikácie ECOICOP. Úroveň II (pod 5-miestnou úrovňou klasifikácie), naopak, nepoužíva žiadny vopred definovaný systém klasifikácie. Spôsob, akým sa jednotlivé krajiny rozhodnú organizovať svoj klasifikačný

⁸ Uvedená informácia platí ak maloobchodný reťazec disponuje kvalitnou internou klasifikáciou a je ochotný ju zdieľať.

system pod päťmiestnou úrovňou je často určený dostupnosťou podrobných údajov o výdavkoch podľa produktov, výberom vzorky a dostupnosťou zdrojov.

Bez ohľadu na to, ako je úroveň II organizovaná, musí byť vytvorený rad elementárnych agregátov, ktoré existujú niekde v rámci tejto štruktúry II. úrovne a ktoré sú základnými stavebnými prvkami CPI/HICP.

Elementárny agregát je najmenší a relatívne homogénny súbor tovarov alebo služieb, pre ktorý definujeme a aj používame údaje o výdavkoch na účely zostavenia CPI. Tovary alebo služby, ktoré sú súčasťou určitého elementárnemu agregátu by si mali byť podobné z pohľadu konečného použitia a tiež cenového vývoja. Môžu byť definované nielen na základe ich charakteristických vlastností, ale aj s ohľadom na región a typ odbytiska, v ktorom sa nachádzajú a predávajú. Dosiahnutý stupeň ich homogenity závisí v praxi od dostupnosti príslušných údajov o výdavkoch. Elementárne agregáty môžu byť preto v rôznych krajinách definované odlišne. Je však potrebné dodržať niekoľko kľúčových pravidiel. Elementárne agregáty by mali:

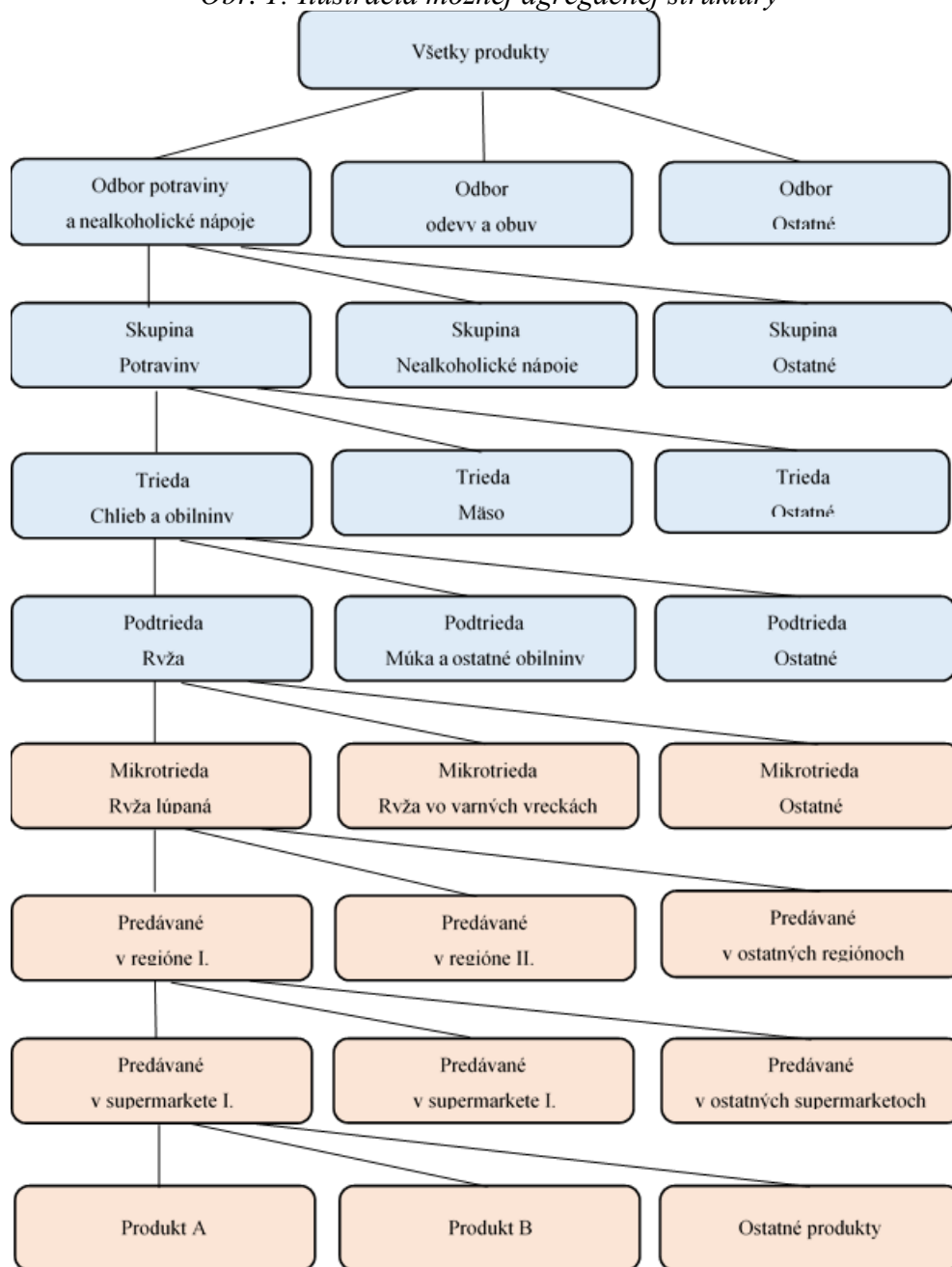
- pozostávať zo skupín tovarov alebo služieb, ktoré sú si čo najviac podobné a predovšetkým dostatočne homogénne z hľadiska konštrukcie a obsahu;
- pozostávať z produktov, u ktorých môžeme očakávať podobné cenové pohyby, pričom cieľom by mala byť minimalizácia rozptylu cenových pohybov v rámci agregátu;
- byť definované tak, aby mohli prípadne slúžiť ako straty v prípade plánovaného výberového zisťovania pre zber údajov.

Štruktúra agregácie pre CPI je znázornená na obr. 1. Použijúc klasifikáciu COICOP je celý súbor spotrebných výdavkov a služieb rozdelený do odborov ako napr. „Potraviny a nealkoholické nápoje“. Každý odbor sa ďalej delí na skupiny, ako napr. „Potraviny“, Skupiny sú ďalej rozdelené do tried, ako napr. „Chlieb a obilniny“. Triedy sa delia na podtriedy, ako napríklad „ryža“. Mnohé krajiny používajú ešte jemnejšiu klasifikáciu ďalšou dezagregáciou pod úroveň podtriedy na homogénnejšie mikrotriedy ako napríklad „ryža lúpaná“ a tie môžu byť ďalej členené podľa regiónu alebo typu predajne. Mikrotrieda nemusí byť ďalej členená a v tom prípade sa stáva elementárnym agregátom. V prípade detailnejšej dezagregácie, ktorá je zobrazená na obr. 1 odlišnou farbou, je príkladom elementárneho agregátu ryža lúpaná predávaná v regióne I, v supermarkete I.

Na základe horeuvedeného sme prostredníctvom detailných informácií o jednotlivých produktoch z údajov zo skenerov definovali 354 homogénnych skupín produktov pre odbor 01 – Potraviny a nealkoholické nápoje klasifikácie ECOICOP, t. j. definovali sme II. (národnú) úroveň klasifikácie ECOICOP (t. j. ECOICOP6) pre odbor 01, ktorá je spoločná pre údaje všetkých obchodných reťazcov, ktoré v tejto oblasti v súčasnosti spolupracujú so ŠÚ SR. Elementárny agregát je potom možné definovať na úrovni SR alebo aj na úrovni obchodného reťazca (vid' obr. 1). Vzhľadom na charakter údajov, ktoré ŠÚ SR poskytujú obchodné reťazce, nie je možné definovať elementárny agregát na regionálnej úrovni.

Údaje, ktoré na ŠÚ SR prichádzajú na týždennej báze je potrebné klasifikovať, t. j. každé pozorovanie musí byť namapované na klasifikačnú štruktúru ECOICOP6. Existujú rôzne techniky ako to urobiť. V súčasnosti napríklad na ŠÚ SR prebieha už fáza testovania vybraných metód strojového učenia na prepájanie údajov internej klasifikácie jednotlivých obchodných reťazcov na klasifikáciu ECOICOP6 a výsledky sú povzbudivé.

Obr. 1: Ilustrácia možnej agregačnej štruktúry



Zdroj: vlastné spracovanie

3.3 Výber tovarových položiek do výpočtu CPI/HICP

Eurostat (2017) odporúča dva možné prístupy k výberu položiek, ktoré by mali byť zahrnuté do výpočtu a to statický a dynamický prístup.

Statický prístup imituje tradičný fixný výber (spotrebný kôš) s tým rozdielom, že dochádza k zmene cenového konceptu. Ceny získavané tradičným spôsobom sú nahradené cenami za jednotku tovaru z údajov zo skenerov. Všetko sa riadi tradičnou metodikou, ale s výhodou úplných informácií o skutočných transakciách, na základe ktorých sa realizuje počítačový výber kódov položiek a v prípade potreby ich nahradenie počas roka. Je pracovne náročný a dostupné údaje využíva len v obmedzenej miere.

Pri dynamickej metóde ide o automatický reprezentatívny výber kódov položiek pre každé dva po sebe nasledujúce mesiace (t a $t + 1$, $t + 1$ a $t + 2$, $t + 2$ a $t + 3$ a tak ďalej), t. j.

výber všetkých spárovaných položiek s tržbami nad určitú hranicu, ktoré budú obsahovať nové a dostatočne dôležité položky, pričom položky, ktoré sú menej dôležité, sa vypustia.

Problémom pri sledovaní cenových zmien cez kódy položiek môže byť opätovné uvedenie výrobku na trh. Preto je niekedy vhodnejšie zaviesť postup, ktorý identifikuje opätovné uvedenie na trh a spája dva kódy položiek, ktoré sa vzťahujú na "rovnaký" výrobok. Opätovné uvedenie na trh môže byť spojené aj so zmenami veľkosti balenia, ktoré by sa tiež mali upraviť pri výpočte priemernej ceny. Napríklad 100 g čokoládová tyčinka, ktorá bola na trhu predtým sa teraz predáva v balení s hmotnosťou len 80 g.

Index elementárneho agregátu sa vypočíta na základe súboru spárovaných reprezentatívnych kódov položiek pre položky, ktoré sa skutočne predávajú v dvoch po sebe nasledujúcich obdobiach. Každý mesiac sa teda súbor jednotlivých výrobkov, ktoré vstupujú do výpočtu indexu vyberá nanovo. V praxi sa uplatňuje „cut-off“ výber, pri ktorom sa vyberajú najpredávanejšie výrobky v dvoch po sebe nasledujúcich obdobiach a ako indexová metóda sa často aplikuje Jevonsov index⁹.

Metóda dynamického koša bola pôvodne navrhnutá v práci Van der Grient, H. a de Haan, J. (2010). Tento prístup môžeme považovať za pragmatický kompromis, ktorý na jednej strane dovoľuje aktualizovať spotrebný kôš v porovnaní so statickým prístupom fixného koša, na druhej strane sa do určitej miery vyhneme problému s prípadným driftom¹⁰ reťazového časového radu lebo Jevons index je nevážený. Váhy sa použijú len implicitne, v rámci výberovej procedúry. Do výpočtu vstupujú len výrobky, ktoré prekročia určitú hranicu. Výrobky s nízkym predajom ignorujeme.

Hlavnou nevýhodou tejto metódy je, že sledované váhy nie sú explicitne zahrnuté do výpočtu indexu a predstavujú len 0 – 1 premennú. Hoci v porovnaní s reťazovým Törnqvistovým indexom je pravdepodobnosť reťazového driftu výrazne nižšia, nedá sa úplne vylúčiť. Určité skreslenie smerom nadol sa môže objaviť najmä vtedy, keď výrobky opúšťajú trh za znížené ceny (Lamboray, 2021). Na zmiernenie tohto problému sa môže použiť dumpingový filter, ktorý z výpočtov vylúči tie jednotlivé výrobky, ktoré vykazujú veľký pokles v cenách aj predaných množstvách.

Na základe horeuvedeného sme v rámci vykonaných experimentov v závislosti od použitej metódy zostavenia indexu aplikovali filtre, ktorými sme niektoré produkty vylúčili z výpočtu. Filtrovanie údajov sme vykonávali na množine údajov, ktorá by potenciálne vstupovala do výpočtu medzimesačného indexu. Filtrovaním sa odstránili extrémne zmeny ceny s minimálnou a maximálnou cenovou zmenou na základe experimentov nastavené na hodnoty 0,3 a 3 (filter na odľahlé hodnoty - outlier filter) a výpredajové produkty, ak je cenová zmena $\leq 0,8$ a súčasne zmena tržieb $\leq 0,2$ (dumping filter). V prípade dynamického prístupu (metódy dynamického spotrebného koša, ktorý sa vytvára vždy aktuálne pre dve po sebe idúce obdobia) sme následne odstránili z výpočtu aj produkty s nízkym predajom na základe vzťahu uvedeného v Eurostat (2017):

$$\frac{S_t^{t-1} + S_t^t}{2} > \frac{1}{n\lambda} \quad (2)$$

⁹ Cenový index definovaný ako nevážený geometrický priemer relatívnych cien bežného obdobia k základnému obdobiu (IMF a kol., 2020).

¹⁰ O reťazovom indexe sa hovorí, že driftuje, ak sa nevráti k jednotke, keď sa ceny v bežnom období vrátia na úroveň, ktorú dosahovali v základnom období. Reťazové indexy sú náchylné na drift, keď ceny počas období, ktoré pokrývajú, kolíšu.

To znamená, že produkt bol zaradený do vzorky pre výpočet indexu ak podiel S položky i na výdavkoch v mesiacoch t a $t-1$ prekročil prahovú hodnotu $\frac{1}{n\lambda}$, kde n je počet uvažovaných produktov a λ je fixný parameter a zvyčajne $\lambda = 1,25$ (Eurostat 2017).

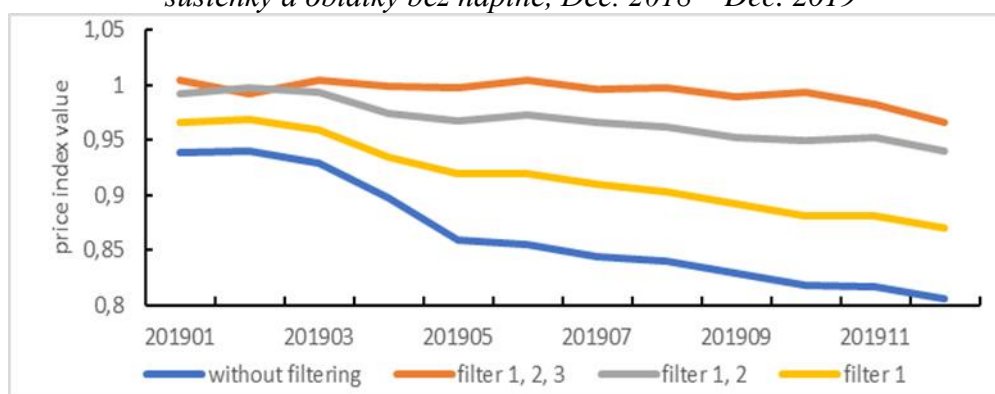
Vplyv filtrovania údajov na počty produktov, ktoré vstupujú do výpočtu indexov vidíme v tabuľke 2.

Tab. 2: Vplyv filtrovania na redukciiu súboru údajov (COICOP 01, priemerný mesiac 2019)

Typ filtra	Popis	Počet záznamov po filtrovaní
Bez filtrovania		60 239
Filter 1	filter odľahlých hodnôt	60 070
Filter 1, 2	filter odľahlých hodnôt + dumping filter	59 638
Filter 1, 2, 3	filter odľahlých hodnôt + dumping filter + filter na nízky predaj	17 057

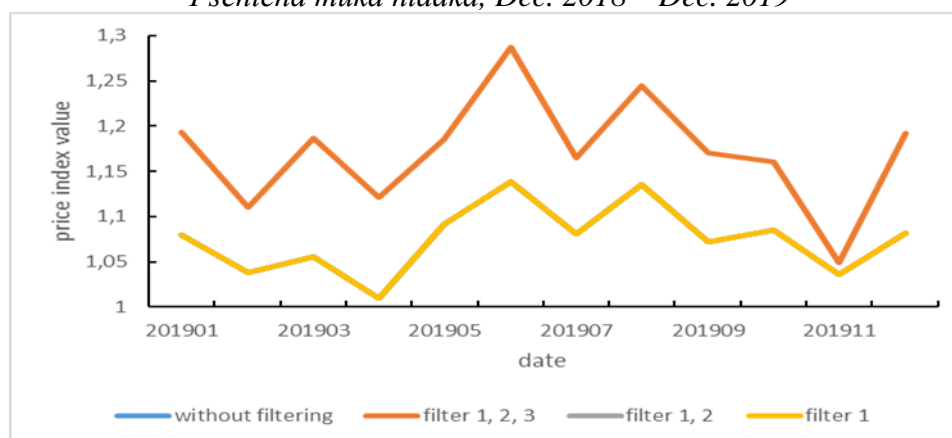
Zdroj: Glaser-Opitzová (2022)

Obr. 2: Vplyv filtra odľahlých hodnôt (filter 1), filtra dampingových cien (filter 2) a filtra nízkych predajov (filter 3) na hodnotu reťazového Jevonsovho indexu, ECOICOP6 – sušienky a obľátky bez náplne, Dec. 2018 – Dec. 2019



Zdroj: Glaser-Opitzová (2022)

Obr. 3: Vplyv filtra odľahlých hodnôt (filter 1), filtra dampingových cien (filter 2) a filtra nízkych predajov (filter 3) na hodnotu reťazového Jevonsovho indexu, ECOICOP6 – Pšeničná múka hladká, Dec. 2018 – Dec. 2019



Zdroj: vlastné spracovanie

Z obrázkov 2 a 3 je zrejmé, že reťazový Jevonsov index je citlivý na výber filtra, čo potvrdzujú aj výsledky experimentálnej štúdie (Bialek, 2020). V uvedených konkrétnych

prípadoch Jevonsov index nadobúda najnižšie hodnoty vtedy, ak ho počítame nad nefiltrovanou databázou produktov. Najviac produktov je z výpočtu odfiltrovaných z dôvodu nízkeho predaja. Práve tieto produkty by nevhodným spôsobom ovplyvnili vývoj, ale hlavne úroveň Jevonsovho cenového indexu, ktorý je počítaný bez použitia váh.

3.4 Definovanie produktu

Podľa (IMF a iní, 2020) pred použitím akejkoľvek metódy výpočtu indexu je potrebné definovať jednotlivé porovnávané produkty. Základným princípom je porovnávať podobné s podobným a sledovať ceny toho istého produktu v priebehu času. V kontexte údajov zo skenerov najpodrobnejšiu úroveň homogenity v údajoch zvyčajne predstavuje úroveň kódu položky. Okrem tejto dimenzie produktu je potrebné brať do úvahy typ predajne alebo predajcu a dimenziu času. Často sa ten istý výrobok predávaný v rôznych časových okamihoch, v tej istej alebo podobných predajniach môže považovať za definíciu produktu, ktorý je dostatočne homogénny, takže pre tento produkt môžeme vypočítať priemernú cenu transakcie. V niektorých kategóriách hlavne vysokoobrátkového tovaru, ako napríklad odevy alebo obuv, kódy tovarov veľmi často vznikajú a zanikajú čo výrazne komplikuje meranie cenových zmien. Riešením môže byť definovať produkt ako zoskupenie jednotlivých tovarových položiek s podobnými charakteristikami. Definíciu je možné poňať širšie aj užšie, ale tak, aby z pohľadu spotrebiteľa boli rozdiely medzi zoskupenými položkami nepodstatné. Výpočet ceny za jednotku tovaru na tejto úrovni umožňuje nielen zachytiť substitučné efekty medzi porovnateľnými výrobkami, ale uľahčuje aj zahrnutie nových výrobkov vstupujúcich na trh. Príliš široko definované zoskupenie položiek však môže viesť k skresleniu ceny za jednotku tovaru a vysokej volatilité. Rozhodnutia týkajúce sa definovania produktov môžu významne ovplyvniť výsledné cenové indexy a preto vplyv definície produktu na výsledky musí byť dôsledne otestovaný.

V prípade potravín a nealkoholických nápojov bol pojem produkt definovaný na úrovni kódu položky a obchodného reťazca, čo v našom prípade predstavuje najpodrobnejšiu úroveň homogenity v údajoch.

3.5 Časové pokrytie

Jednotková cena produktu presnejšie odráža ceny, ktoré platia spotrebiteľia v rámci celého sledovaného obdobia ako zistenie ceny v konkrétnom čase v prípade tradičného zberu (Balk, 1998). Jednotkové ceny zohľadňujú zľavy a vplyv týchto zliav na množstvo predaného tovaru. Obdobie, za ktoré sú jednotkové ceny vypočítané, je dôležité vzhľadom na presnosť jednotkovej ceny. Diewert a iní (2016) tvrdia, že jednotkové ceny používané na konštrukciu CPI by mali byť vypočítané za rovnaké obdobie ako je obdobie, za ktoré sa zostavuje index (napr. mesiac), a nie za čiastkové obdobie. V praxi sa však štandardne používa čiastkové obdobie referenčného obdobia z dôvodu včasnosti a časovej presnosti poskytovania štatistických produktov. Rovnako k tomu bude musieť pristúpiť aj ŠÚ SR a nájsť vhodný kompromis. Či bude navrhnuté obdobie dvoch alebo troch týždňov je momentálne ešte v štádiu experimentovania.

4 Cenové indexy pre elementárny agregát

Ideálny vzorec pre výpočet indexu, ktorý sa má použiť pre výpočet CPI by mal mať váhy pre každé cenové pozorovanie, ktoré sa použije na zostavenie elementárneho cenového indexu a tiež váhy pre agregovanie elementárnych indexov na vyššiu úroveň. Zdrojom týchto informácií môžu byť údaje zo skenerov.

Napriek tomu sa v praxi vo väčšine prípadov cenové indexy pre elementárne agregáty počítajú bez použitia explicitných váh. V tomto prípade sa musí elementárny index vypočítať

ako nevážený priemer cien, z ktorých pozostáva. V prípade výberu reprezentantov s pravdepodobnosťami proporcionálnymi k veľkosti niektorých relevantných premenných ako napr. tržby sú váhy implicitne zavedené výberovou procedúrou.

Jedným zo spôsobov ako sa rozhodnúť pre vhodný vzorec indexu je požadovať, aby spĺňal určité špecifické axiomy alebo testy. Tieto testy objasňujú vlastnosti indexov, ktoré nemusia byť na prvý pohľad zrejmé. Medzi štyri základné testy, ktorými môžeme ilustrovať axiomatický prístup patria (IMF a kol., 2020): test proporcionality, test súmernosti, časovo reverzný test a test tranzitivity. Jevonsov index na rozdiel od iných indexov spĺňa všetky uvedené testy a je teda z axiomatického hľadiska jednoznačne index s najlepšimi vlastnosťami. Aj keď sa donedávna veľmi nepoužíval, trend jeho využívania štatistickými úradmi má rastúcu tendenciu (IMF a kol., (2020)). Môže byť napísaný pre bázičné (0) a aktuálne obdobie (t) nasledovne:

$$P_J^{0,t} = \left(\prod_{i \in S} \frac{p_i^t}{p_i^0} \right)^{1/N_{0,t}} = \frac{(\prod_{i \in S} p_i^t)^{1/N_{0,t}}}{(\prod_{i \in S} p_i^0)^{1/N_{0,t}}} \quad (3)$$

S predstavuje množinu zhodných tovarových položiek patriacich do určitej kategórie a $N_{0,t}$ počet zhodných tovarových položiek. p_i^0 a p_i^t sú ceny (ceny za jednotku tovaru) každej položky $i \in S$ v období 0 a t .

Pri výbere vhodného indexu na meranie CPI/HICP je možné použiť okrem axiomatického prístupu aj prístup ekonomický. Ekonomický prístup je založený na ekonomickej teórii správania sa spotrebiteľa. Ekonomický prístup k teórii indexových čísel predpokladá, že množstvo je funkciou ceny, pričom pozorované údaje sa generujú ako riešenia rôznych problémov ekonomickej optimalizácie (IMF a iní, 2020). Najpreferovanejšími indexami pre účely merania CPI sú superlatívne cenové indexy (Diewert, 1976). Môžu byť napísané ako:

Fischerov cenový index (Fischer, 1922):

$$P_F^{0,t} = \sqrt{P_{La}^{0,t} P_{Pa}^{0,t}} \quad (4)$$

ktorý je geometrickým priemerom Laspeyresovho a Paascheho indexu a Törnqvistov cenový index (Törnqvist, 1936):

$$P_T^{0,t} = \left(\prod_{i \in S} \frac{p_i^t}{p_i^0} \right)^{s_i^0 + s_i^t / 2} \quad (5)$$

kde $s_i^0 = p_i^0 q_i^0 / \sum_{i \in S} p_i^0 q_i^0$ a $s_i^t = p_i^t q_i^t / \sum_{i \in S} p_i^t q_i^t$ označujú podiely výdavkov v období 0 a t ; q_i^0 a q_i^t sú predané množstvá.

Superlatívne indexy majú základ v ekonomickej teórii a použité váhy odrážajú zmeny v spotrebiteľskom správaní, ktoré ilustruje obrázok 4.

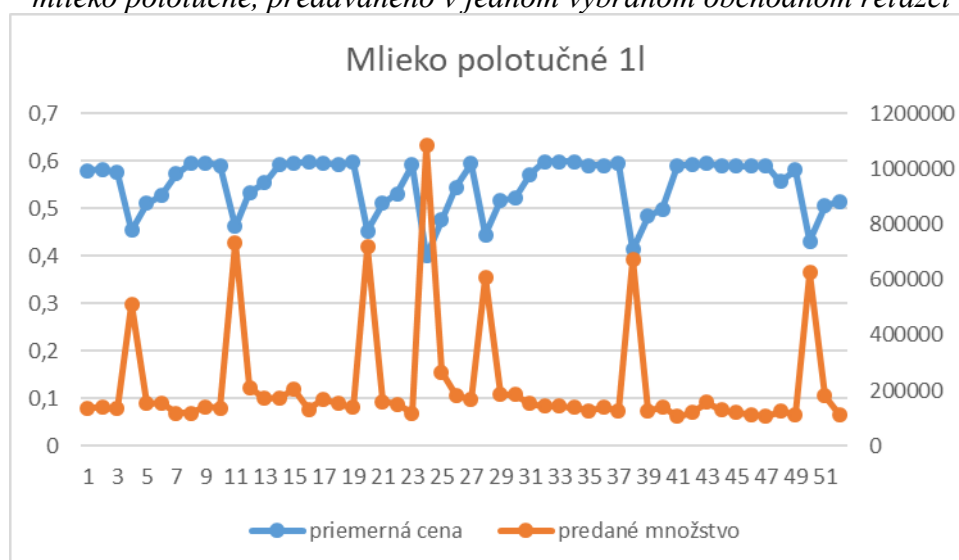
Podľa (Diewert, 1976) sa v prípade výpočtu indexov nad fixným košom (statický prístup k výberu produktov) superlatívne indexy (Fischer, Törnqvist) navzájom aproximujú. V našom prípade, ktorý je ilustrovaný na obr. 5 (dynamický kôš, reťazené verzie indexov) sa správajú rovnako. Ich vývoj je v súlade s ekonomickým prístupom v teórii cenových indexov, podľa ktorej superlatívne indexy aproximujú index životných nákladov a Laspeyres index vymedzuje jeho hornú hranicu a Paasche index jeho dolnú hranicu (Labudová a kol., 2021)

Okrem otázky vážené indexy vs nevážené indexy je potrebné sa zaoberať aj otázkou bázičné indexy vs reťazové indexy.

Bázické indexy majú tú výhodu, že neobsahujú drift v časovom rade indexov, ale majú aj nevýhodu. V priebehu času sa objavujú nové výrobky a staré výrobky miznú a je stále ťažšie porovnávať položky, ktoré sú k dispozícii v aktuálnom období, s položkami, ktoré boli dostupné v bázičkom období.

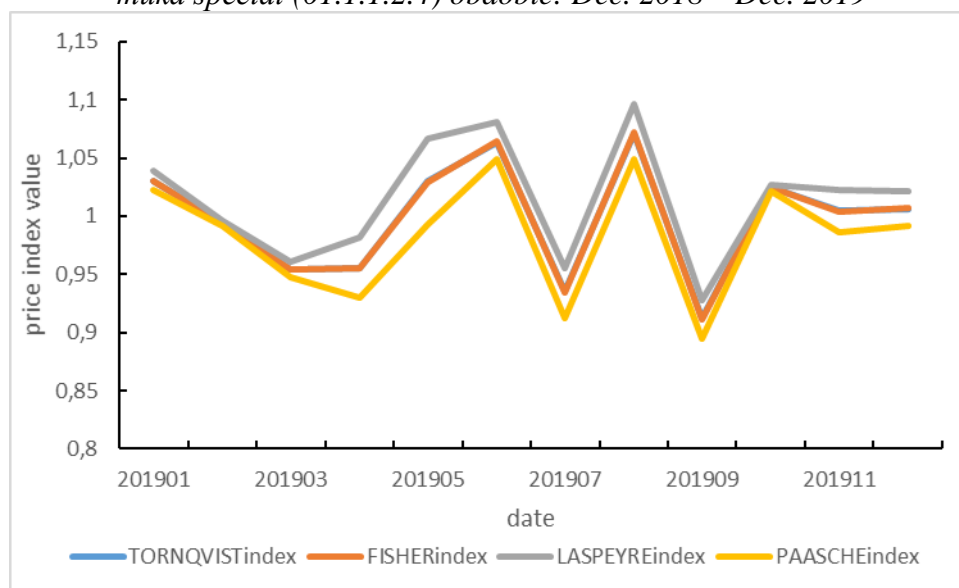
Všeobecne sa odporúča použiť reťazové superlatívne indexy a to z dôvodu vyššej miery spárovanania jednotlivých kódov položiek medzi dvoma po sebe idúcimi obdobiami a predpokladu menších rozdielov v cene a kvantite. Tento predpoklad však nezohľadňuje existenciu výpredajov a zliav, pri ktorých sa môže výrazne zvýšiť kvantita predaného tovaru a to až niekoľkonásobne, čo môže spôsobiť drift v časovom rade.

Obr. 4: Porovnanie vývoja ceny a predaného množstva produktu - ECOICOP6 – mlieko polotučné, predávaného v jednom vybranom obchodnom reťazci



Zdroj: vlastné spracovanie

Obr. 5: Porovnanie vývoja vážených bilaterálnych reťazených indexov, ECOICOP6 – múka špeciál (01.1.1.2.4) obdobie: Dec. 2018 – Dec. 2019



Zdroj: vlastné spracovanie

Ivancic, Diewert a Fox (2009) navrhli postup, ktorý poskytuje indexy superlatívneho typu bez driftu, prostredníctvom adaptácie teórie multilaterálnych indexových čísel. Tento postup maximalizuje množstvo spárovaných položiek v údajoch bez rizika driftu.

Metódy multilaterálnych cenových indexov sa zvyčajne používajú pre porovnanie cenových úrovní medzi krajinami a regiónmi tak, aby porovnania boli tranzitívne, t.j. výsledky sú nezávislé na výbere bázičky krajiny. Najznámejšie metódy sú GEKS (Gini, 1931), (Eltető a Köves, 1964), (Szulc, 1964), Geary-Khamis metóda (Geary, 1958), (Khamis 1972) a Country-Product Dummy (CPD) metóda navrhnutá v (Summers, 1973).

Multilaterálne priestorové porovnanie cien môže byť jednoducho prispôbené pre porovnanie v čase. Multilaterálne indexy spĺňajú požiadavku cirkularity alebo tranzitivity, to znamená, že sa dopracujeme k rovnakým výsledkom či už porovnáваме entity navzájom priamo alebo prostredníctvom vzťahov s inými entitami.

4.1 Metóda GEKS

Tranzitívny multilaterálny index GEKS je geometrický priemer pomerov všetkých bilaterálnych (v štandardnej verzii Fisherových) indexov kde každá entita je uvažovaná ako báza. Kontextu časových radov metódu prispôbili a ďalej rozvinuli Ivancic, Diewert a Fox (2011).

Pri aplikovaní metódy treba v prvom rade určiť dĺžku intervalu (okna), na ktoré má byť metóda aplikovaná. Spravidla sa metóda aplikuje na dĺžku okna jeden alebo dva roky plus jedno obdobie, aby sa zohľadnili aj produkty zaťažené sezónnosťou. Dĺžku okna označíme w . Vyberieme bázičné obdobie v rámci okna, budeme ho označovať ako k , a vypočítame bilaterálny index medzi obdobím k a každým ďalším obdobím v okne. Postup opakujeme pre všetky možné voľby k . Takto dostaneme maticu bilaterálnych indexov veľkosti $w \times w$ pre všetky možné páry jednotlivých období v rámci okna. Potom vypočítame GEKS indexy pre prvých w období ako:

$$[\prod_{k=1}^w P^{k,1}]^{1/w}, [\prod_{k=1}^w P^{k,2}]^{1/w}, \dots, [\prod_{k=1}^w P^{k,w}]^{1/w}, \quad (6)$$

kde výraz $P^{k,t}$ je bilaterálny index medzi obdobím t a bázičným obdobím k . V pôvodnej verzii, P označoval Fischerov index, v súčasnosti sa v metóde používa napr. aj Törnqvist, Walsh alebo Jevons index. V podstate je možné použiť akýkoľvek iný bilaterálny index, ak spĺňa podmienku časovej reverzibility. Táto vlastnosť vyžaduje, aby sa index medzi obdobiami a a b rovnal inverznej hodnote rovnakého indexu medzi obdobiami b a a . Matica bilaterálnych indexov $w \times w$ vyzera nasledovne:

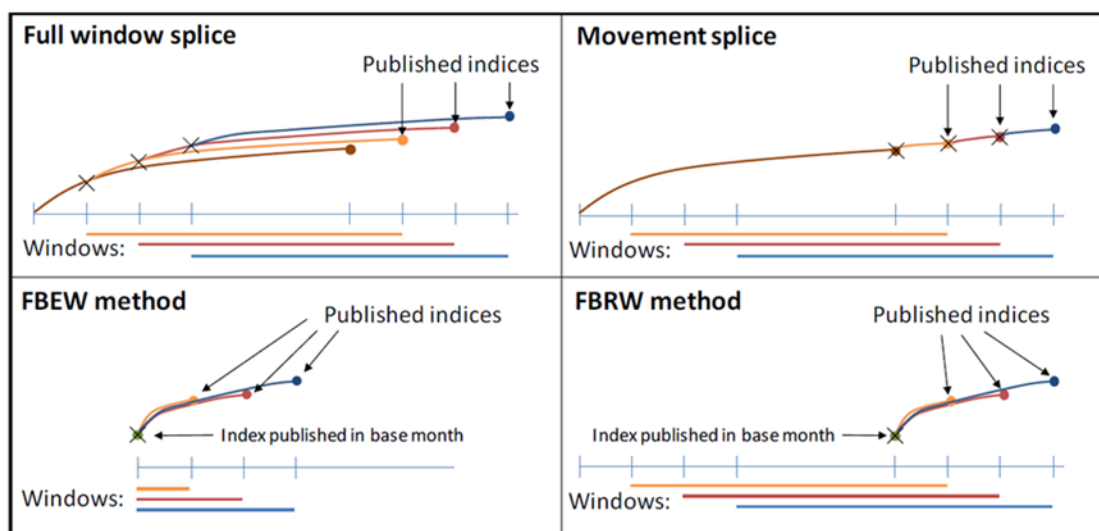
$$P = \begin{pmatrix} P^{1,1} & \dots & P^{1,w} \\ \vdots & \ddots & \vdots \\ P^{w,1} & \dots & P^{w,w} \end{pmatrix} \quad (7)$$

Potom prvá hodnota časového radu GEKS indexov pre prvých w období je geometrickým priemerom prvkov v prvom stĺpci vyššie uvedenej matice, druhý člen je geometrickým priemerom druhého stĺpca atď. V prípade multilaterálnych indexov vzniká problém pridaním informácií z nového mesiaca pretože v prostredí štatistickej produkcie sa musí časové okno každý mesiac upraviť tak, aby zahŕňalo údaje z posledného mesiaca. Časové okno môžeme upraviť rôznym spôsobom:

- Každý mesiac posunieme časové okno o jeden mesiac dopredu, dĺžka časového okna je konštantná, posledný najaktuálnejší mesiac zahrnieme, zatiaľ čo najstarší mesiac odstránime.
- Každý mesiac predĺžime časové okno o jeden mesiac. Časové okno sa každý mesiac predlžuje. Po uplynutí jedného roku sa časové okno nastaví na pôvodnú dĺžku.

Pri každom použití nového časového okna sa predtým vypočítané indexy môžu zmeniť. Aby sa predišlo revíziám už publikovaných výsledkov, používajú sa rôzne techniky spájania, ktoré spoja najnovší multilaterálny index s predchádzajúcimi výsledkami. Technicky funguje spojenie dvoch časových radov indexov cez spojku t. j. obdobie, cez ktoré vytvoríme prepojenie. Pre kompiláciu výsledkov v mesiaci t existujú rôzne metódy ako toto prepojenie vytvoriť a ku každej metóde ešte dva hlavné varianty spájania (Chessa, 2019). Dôvodom sú postupné posuny okna, ktoré generujú sekvenciu prepočítaných alebo „revidovaných“ indexov popri už publikovanom indexe v rovnakom období. Obidva, prepočítaný aj zverejnený index sú kandidátmi na index, na ktorý je možné naviazať nový rad indexov. Preto budeme rozlišovať dva varianty spájania (okrem metódy „Movement splice“, ktorá má index publikovaný v predchádzajúcom období ako jedinou možnosť prepojenia) a to prepojiť multilaterálny index zostavený v období t s multilaterálnym indexom zostaveným v období $t-1$ alebo prepojiť multilaterálny index zostavený v období t s publikovaným indexom (viď obrázok 6).

Obr. 6: Ilustrácia extenzívnych metód, kde \times znamená mesiac spájania a index



Zdroj: Chessa (2019, str. 5)

V prípade multilaterálnych metód sme sa pri experimentovaní zamerali hlavne na porovnanie GEKS_Jevons a GEKS_Törnqvist v kombinácii s rôznymi extenzívnymi metódami. Dĺžku okna sme zvolili 13 mesiacov. Mesačné GEKS multilaterálne indexy sme kompilovali na najnižšej úrovni agregácie, nad spoločnou databázou produktov. To znamená, že sme aplikovali len časovú a produktovú agregáciu, základná úroveň nebola špecifická pre obchodný reťazec. Elementárnou úrovňou, na ktorej sa zostavovali cenové indexy bola úroveň ECOICOP6. Indexy GEKS pre 354 skupín výrobkov boli agregované na vyššie úrovne ECOICOP s použitím štandardných ročne reťazených indexov Laspeyresovho typu s pevnými váhami, ktoré sa vzťahujú na predchádzajúci rok. Rovnaký postup sa používa aj na agregáciu "oficiálnych" indexov. Cenové indexy získané aplikovaním rôznych extenzívnych metód sa porovnávali s indexami, ktoré sú tranzitívne na celej dĺžke časového obdobia, ku ktorej sa vzťahujú súbory údajov. Výsledky prezentované v článku Glaser-Opitzová (2021) ukazujú nesignifikantné rozdiely medzi extenzívnymi metódami v prípade GEKS_Törnqvist aj GEKS_Jevons. Situácia je však odlišná pri vzájomnom porovnaní oboch multilaterálnych metód, kde bol rozdiel signifikantný a to hlavne z dôvodu, že GEKS_Jevons plne nevyužil potenciál údajov zo skenerov, pretože ide o nevážený index a možnosť konštruovať váhy na elementárnej úrovni je jednou z ich fundamentálnych výhod.

5 Záver

Transakcie z maloobchodnej siete sú pre cenovú štatistiku čoraz dostupnejším dátovým zdrojom. Kompilácia cenových indexov z takýchto zdrojov však nie je jednoduchá a vyvoláva mnohé metodologické aj praktické otázky.

Na základe detailných informácií o jednotlivých produktoch z údajov zo skenerov sme pre odbor 01 – Potraviny a nealkoholické nápoje klasifikácie ECOICOP definovali 354 homogénnych skupín produktov (mikrotried), t.j. definovali sme II. (národnú) úroveň klasifikácie, ECOICOP6. Podarilo sa nám definovať rovnaké homogénne skupiny produktov pre všetky spolupracujúce obchodné reťazce. Elementárny agregát je takto možné definovať na úrovni SR alebo aj na úrovni obchodného reťazca. Oba návrhy už boli predmetom empirického skúmania a z výsledkov vyplynulo, že je v podmienkach SR možné definovať elementárny agregát aj na úrovni SR, čo môže celý proces kompilácie CPI/HICP zjednodušiť.

Na základe ďalšieho empirického skúmania sme pre odbor potravín a nealkoholických nápojov definovali produkt na úrovni kódu položky a stanovili prahové hodnoty, ktoré z výberu tovarov, ktoré vstupujú do výpočtu cenových indexov na elementárnej úrovni vylučujú extrémne zmeny cien, výpredajové tovary alebo málo predávané tovary.

Ukončili sme experimentálne skúmanie nad fixným spotrebným košom (statický prístup k výberu tovarových položiek). Predmetom ďalšieho skúmania bude výlučne dynamický prístup.

Problematika implementácie údajov zo skenerov do štatistickej produkcie CPI je však veľmi široká a mnohostranná a mnohé metodologické otázky stále nie sú zodpovedané. Veľkou výzvou zostáva výber vhodného indexu pre elementárny agregát. Budúce empirické skúmanie bude zamerané najmä na možnú implementáciu Jevonsovho indexu a superlatívnych reťazových indexov, aj keď majú svoje limity, ale najmä na multilaterálne indexy.

Oznam

Príspevok bol spracovaný v rámci riešenia grantovej úlohy **VEGA 1/0561/21: Vplyv krízy COVID-19 na demografiu podnikov a zamestnanosť v SR a EÚ**.

Literatúra

- [1] Balk, B. M. (1998). On the Use of Unit Value Indices as Consumer Price Subindices. Paper presented at the Fourth Meeting of the International Working Group on Price Indices, Washington, DC, April 22–24. <http://www.ottawagroup.org>.
- [2] Bialek, J. (2020). PriceIndices – A New R Package for Bilateral and Multilateral Price Index Calculations. *Statistica: Statistics and economy journal* 101 (2), 54–69
- [3] Diewert, W. E. (1976). Exact and Superlative Index Numbers. *Journal of Econometrics* 4: 114–45.
- [4] Diewert, W. E., Fox, K. J. & de Haan, J. (2016). A Newly Identified Source of Potential CPI Bias: Weekly versus Monthly Unit Value Price Indexes. *Economics Letters* 141, 169–72.
- [5] Eltető, Ö. - Köves, P. (1964): On a Problem of Index Number Computation Relating to International Comparisons (in Hungarian). In.: *Statisztikai Szemle* 42, s. 507-518.
- [6] Eurostat (2017). Practical Guide for Processing Supermarket Scanner Data. Retrieved 13.05.2019 from <https://circabc.europa.eu/sd/a/8e1333df-ca16-40fc-bc6a-1ce1be37247c/Practical-Guide-Supermarket-Scanner-Data-September-2017.pdf>.
- [7] Fisher, I. (1922). *The Making of Index Numbers*. Boston, MA: Houghton-Mifflin.
- [8] Geary, R.C.: A Note on the Comparison of Exchange Rates and Purchasing Power between Countries. In.: *Journal of the Royal Statistical Society A* 121, 1958, s. 97-99.

- [9] Gini, C. (1931). On the Circular Test of Index Numbers. In: International Review of Statistics 9, s.3-25.
- [10] Glaser-Opitzová, H. (2021, December). The Multilateral Methods for the Consumer Price Index Measurement using Scanner Data in the Conditions of the Slovak Republic. Economic and Social Development 76th International Scientific Conference on Economic and Social Development – "Building Resilient Society". Retrieved November 13, 2022, from https://www.researchgate.net/publication/363845644_1169549Book_of_Proceedings_esdZagreb2021_Online_1_1
- [11] Glaser-Opitzová, H. (2022). Use of Scanner Data in Measuring the Consumer Price Index in the Conditions of the Slovak Republic. Proceedings from the EDAMBA 2021 conference, 92 – 102. <https://doi.org/10.53465/EDAMBA.2021.9788022549301.92-102>
- [12] Chessa, A.G. (2019). A comparison of index extension methods for multilateral methods. Paper presented at the 16th Meeting of the Ottawa Group on Price Indices, Rio de Janeiro, Brazil, 8 - 10 May 2019. Retrieved 11.10.2021 from [https://www.ottawagroup.org/Ottawa/ottawagroup.nsf/home/Meeting+16/\\$FILE/A%20comparison%20of%20index%20extension%20methods%20paper.pdf](https://www.ottawagroup.org/Ottawa/ottawagroup.nsf/home/Meeting+16/$FILE/A%20comparison%20of%20index%20extension%20methods%20paper.pdf).
- [13] IMF, ILO, EUROSTAT, UNECE, The World Bank (2020). Consumer price index manual: concepts and methods. Washington, DC : International Monetary Fund, 2020. Retrieved 11.10.2021 from <https://www.imf.org/en/Data/Statistics/cpi-manual>
- [14] Ivancic, L. Diewert, W. E. & Fox, K. J. (2009). Scanner data, time aggregation and the construction of price indexes. Discussion paper 09-09, University of British Columbia, Vancouver, Canada.
- [15] Ivancic, L. Diewert, W.E. & Fox, K.J. (2011). Scanner data, time aggregation and the construction of price indexes. Journal of Econometrics, 161, 24-35.
- [16] Khamis, S.H. (1972). A New System of Index Numbers for National and International Purposes, In.: Journal of the Royal Statistical Society A 135, s. 96-121.
- [17] Labudová, V., Pacáková, V., Sipková, L., Šoltés, E., Vojtková, M. (2021). Štatistické metódy pre ekonómov a manažérov, Bratislava: Wolters Kluwer.
- [18] Lamboray, C. (2021). Index Compilation Techniques for Scanner Data - Paper prepared for the Group of Experts on Consumer Price Indices UNECE, Online meeting, June 2021.
- [19] Sásiková, I. (2022, April 28). Index spotrebiteľských cien. Retrieved November 13, 2022, from <http://www.statistics.sk/>
- [20] Summers, R. (1973). International Price Comparisons Based Upon Incomplete Data. In.: Review of Income and Wealth, 19, s. 1-16.
- [21] Szulc, B. (1964). Indices for Multiregional Comparisons (in Polish). In.: Przegląd Statystyczny 3, , s. 239-254.
- [22] ŠÚ SR (2021). Metodické vysvetlivky – HICP / CPI. Dostupné 25.01.2021 na www.statistics.sk
- [23] Van der Grient, H. & de Haan, J. (2010). The use of supermarket scanner data in the Dutch CPI. CBS. Dostupné na [The use of supermarket scanner data in the Dutch CPI \(cbs.nl\)](http://www.cbs.nl), prístup 13.05.2019.
- [24] Törnqvist, L. (1936). The Bank of Finland's Consumption Price Index. Bank of Finland Monthly Bulletin 10: 1–8.