

SLOVENSKÁ ŠTATISTIKA a DEMOGRAFIA

SLOVAK STATISTICS
and DEMOGRAPHY

4/2026

ročník/volume 36

Recenzovaný vedecký časopis so zameraním na prezentáciu moderných štatistických a demografických metód a postupov.

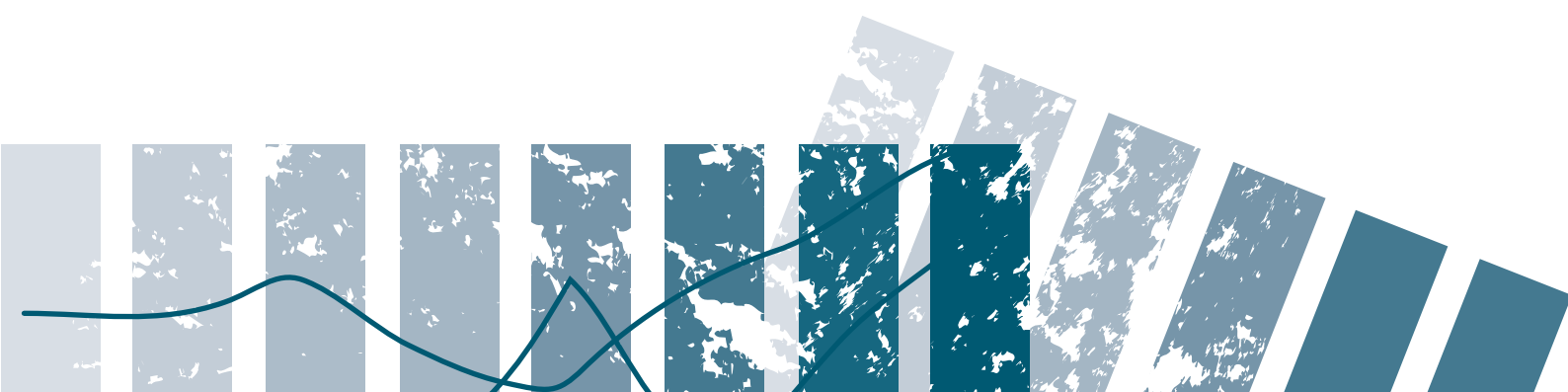
Scientific peer-reviewed journal focusing on the presentation of modern statistical and demographic methods and procedures.

Článok/Article: 3

Typ článku/Type of article: vedecký článok/scientific article

Strany/Pages: 27 – 44

Dátum vydania/Publication date: 15. október 2026/October 15, 2026



Peter KNÍŽAT

Štatistický úrad Slovenskej republiky, Fakulta managementu, Univerzita Komenského v Bratislave

**TRANSAKČNÉ ÚDAJE Z INTERNETOVÝCH OBCHODOV
AKO NOVÝ ZDROJ ÚDAJOV PRE CENOVÚ ŠTATISTIKU:
MOŽNOSTI ZOSTAVOVANIA INDEXOV SPOTREBITEĽSKÝCH CIEŇ**

**TRANSACTION DATA FROM ONLINE RETAILERS
AS A NEW DATA SOURCE FOR PRICE STATISTICS:
POSSIBILITIES FOR CONSUMER PRICE INDEX COMPILATION**

ABSTRAKT

Digitalizácia maloobchodného predaja významne mení spôsob nakupovania spotrebiteľov a zároveň vytvára nové možnosti na modernizáciu produkcie oficiálnych štatistík. Zatiaľ čo tradičný zber cien je založený na zisťovaní cien vybraných reprezentantov, nové zdroje údajov umožňujú pracovať s rozsiahlymi databázami cien a transakcií získavaných priamo z elektronických obchodov. Jedným z perspektívnych zdrojov sú transakčné údaje internetových obchodov, ktoré okrem cien obsahujú aj informácie o počte predaných položiek a hodnote predaja jednotlivých produktov. Takéto údaje poskytujú presnejšie zachytenie skutočného spotrebiteľského správania a vytvárajú predpoklady na zostavovanie ekonomicky presnejších indexov spotrebiteľských cien. Cieľom článku je predstaviť možnosti využitia transakčných údajov z internetových obchodov pri zostavovaní indexov spotrebiteľských cien v podmienkach Štatistického úradu Slovenskej republiky. Príspevok opisuje spôsob automatizovaného získavania a spracovania údajov, predstavuje metodický rámec výpočtu bilaterálnych a multilaterálnych indexov spotrebiteľských cien a analyzuje rozdiely medzi váženými a neváženými prístupmi. Empirická analýza vychádza z experimentálnych transakčných údajov poskytovaných platformou Heureka pre vybrané produktové skupiny slovenského internetového trhu. Výsledky ukazujú, že využitie transakčných údajov z internetových obchodov umožňuje konštruovať vážené indexy spotrebiteľských cien založené na reálnom objeme predaja, pričom multilaterálne metódy významne znižujú volatilitu časových radov. Súčasne sa potvrdzuje význam filtrovania extrémnych cenových zmien pred samotným výpočtom indexov. Článok predstavuje ďalší krok vo výskume implementácie alternatívnych zdrojov údajov do cenovej štatistiky a poukazuje na potenciál transakčných údajov z internetových obchodov ako perspektívneho zdroja pre budúcu produkciu oficiálnych indexov spotrebiteľských cien.

ABSTRACT

The digitalisation of retail trade is changing consumer purchasing behaviour while creating new opportunities for the modernisation of official statistics. While traditional price collection is based on selected representative products, new data sources enable statistical institutions to work with large databases of prices and transactions obtained directly from online shops. One of the most promising data sources is transaction data from online retailers, which, in addition to prices also contain information on the quantities of individual products sold and their sales value. Such data provide a more accurate representation of actual consumer purchasing behaviour and create the potential for compiling more economically accurate consumer price indices. The objective of this paper is to present the potential use of transaction data from online

retailers for the compilation of consumer price indices at the Statistical Office of the Slovak Republic. The paper describes the automated data collection process, presents the methodological framework for bilateral and multilateral consumer price indices, and analyses the differences between weighted and unweighted approaches. The empirical analysis is based on experimental transaction data obtained from the Heureka online platform for selected product categories on the Slovak market. The results indicate that transaction data enable the construction of economically more representative weighted price indices while multilateral methods substantially reduce the volatility of price index series. The findings also confirm the importance of filtering extreme price changes before index compilation. The paper contributes to the ongoing research into the implementation of alternative data sources in official price statistics and demonstrates the potential of transaction data from online retailers for future consumer price index production.

KĹÚČOVÉ SLOVÁ

cenová štatistika, index spotrebiteľských cien, transakčné údaje, internetový obchod, multilaterálne indexy, oficiálna štatistika

KEY WORDS

price statistics, consumer price index, transaction data, e-commerce, multilateral indices, official statistics

1. ÚVOD

Meranie inflácie patrí medzi základné úlohy národných štatistických inštitúcií. Index spotrebiteľských cien predstavuje jeden z najvýznamnejších makroekonomických ukazovateľov, ktorý slúži nielen na hodnotenie cenového vývoja v ekonomike, ale zároveň tvorí dôležitý podklad pre menovú politiku, valorizáciu sociálnych dávok, dôchodkov či obchodných kontraktov. Kvalita odhadu inflácie je preto priamo závislá od kvality údajov, ktoré vstupujú do výpočtu indexov spotrebiteľských cien.

Tradičný spôsob zisťovania cien je založený na pravidelnom zbere cien reprezentantov vo vybraných kamenných predajniach alebo prostredníctvom internetového zisťovania cien. Takýto prístup bol dlhé roky postačujúci, avšak dynamický rozvoj elektronického obchodu významne mení spôsob spotrebiteľského správania. Rastúci podiel online nákupov spôsobuje, že čoraz väčšia časť spotrebiteľských výdavkov vzniká mimo tradičných predajní. Súčasne vznikajú nové zdroje údajov, ktoré dokážu zachytiť cenový vývoj podstatne detailnejšie ako klasické štatistické zisťovania.

Jedným z najvýznamnejších smerov rozvoja cenovej štatistiky je využívanie alternatívnych zdrojov údajov. Medzi najznámejšie patria skenerové údaje získavané priamo od maloobchodných obchodných reťazcov, webscrapované údaje získavané automatizovaným zberom cien z internetových stránok a najnovšie aj transakčné údaje internetových obchodov. Zatiaľ čo webscraping poskytuje informácie o ponúkaných cenách produktov, transakčné údaje zachytávajú skutočne realizované nákupy spotrebiteľov. Okrem ceny jednotlivých produktov obsahujú aj informácie o počte predaných kusov a hodnote predaja, čím umožňujú zostavovať vážené indexy spotrebiteľských cien založené na reálnych spotrebiteľských výdavkoch.

Štatistický úrad Slovenskej republiky sa problematike implementácie nových zdrojov údajov venuje systematicky už niekoľko rokov. Po úspešnom výskume využitia webscrapovaných údajov nasledovala spolupráca s platformou Heureka, ktorá umožnila pravidelné získavanie agregovaných transakčných údajov zo slovenských internetových obchodov. Významnou výhodou tohto riešenia je možnosť úplnej automatizácie celého procesu od získania údajov až po výpočet experimentálnych indexov spotrebiteľských cien.

Na rozdiel od webscrapovaných údajov, ktoré reprezentujú ponukové ceny produktov, transakčné údaje z internetových obchodov umožňujú presnejšie zachytiť ekonomickú realitu trhu. Spotrebiteľ totiž nemusí nakupovať produkt s najnižšou ponúkanou cenou a časť produktov ponúkaných v internetových obchodoch nemusí byť vôbec predaná. Použitie transakčných údajov preto predstavuje prirodzený krok k ekonomicky presnejšiemu meraniu cenového vývoja, podobne ako implementácia skenerových údajov do cenovej štatistiky kamenných predajní.

Napriek uvedeným výhodám je pri interpretácii transakčných údajov z platformy Heureka potrebné zohľadniť aj niektoré špecifiká a obmedzenia tohto dátového zdroja. Údaje nezachytávajú všetky transakcie realizované na slovenskom internetovom trhu, keďže nie všetky internetové obchody pôsobiace na trhu s platformou Heureka spolupracujú a rozsah ich zapojenia môže byť rozdielny. Zároveň nie všetky nákupy produktov vyhľadávaných prostredníctvom platformy Heureka musia byť zachytené v dostupných transakčných údajoch, keďže používateľ môže uskutočniť nákup priamo zo stránky konkrétneho internetového obchodu. Pokrytie transakcií sa preto môže medzi jednotlivými produktmi a produktovými skupinami líšiť. Ďalším limitom údajov z platformy Heureka môže byť, že štruktúra produktov zastúpených na platforme nemusí v rovnakej miere reprezentovať jednotlivé časti spotrebného koša. Uvedené limitácie predstavujú priestor na ďalší výskum. Ten by sa mal zamerať predovšetkým na posúdenie reprezentatívnosti dostupných transakčných údajov a na kvantifikáciu podielu jednotlivých produktov a produktových skupín zachytených prostredníctvom platformy Heureka na celkovom relevantnom internetovom trhu. Takéto hodnotenie umožní presnejšie posúdiť možnosti širšieho využitia tohto zdroja údajov pri zostavovaní indexov spotrebiteľských cien v praxi.

Cieľom článku je predstaviť možnosti využitia transakčných údajov z internetových obchodov pri zostavovaní indexov spotrebiteľských cien. Pozornosť sa sústreďuje na opis automatizovaného procesu spracovania údajov, porovnanie bilaterálnych a multilaterálnych metód výpočtu indexov spotrebiteľských cien a analýzu rozdielov medzi váženými a neváženými prístupmi. Empirická časť vychádza z experimentálnych údajov získavaných Štatistickým úradom Slovenskej republiky priamo od spoločnosti Heureka a nadväzuje na doterajší výskum implementácie nových zdrojov údajov do produkcie oficiálnych cenových štatistík.

Výsledky predstavené v tomto článku vychádzajú z príspevku prezentovaného na 64. svetovom štatistickom kongrese (*ISI World Statistics Congress*), ktorý sa konal v roku 2025 v Haagu. Medzinárodná odborná diskusia počas konferencie poskytla cennú spätnú väzbu k navrhovanej metodike využitia transakčných údajov z internetových obchodov pri zostavovaní indexov spotrebiteľských cien. Predkladaný článok nadväzuje na prezentované výsledky, rozširuje ich o podrobnejší metodický

opis a zasadzuje ich do širšieho kontextu implementácie alternatívnych zdrojov údajov v oficiálnej cenovej štatistike.

2. PREHĽAD LITERATÚRY

V posledných dvoch desaťročiach prechádza cenová štatistika významnou transformáciou, ktorá je spôsobená rastúcou digitalizáciou maloobchodu a dostupnosťou nových zdrojov údajov. Tradičný spôsob zisťovania cien prostredníctvom manuálneho zberu alebo výberových štatistických zisťovaní je postupne dopĺňaný, prípadne nahrádzaný automatizovanými zdrojmi údajov, ako sú skenerové údaje, webscrapované údaje alebo transakčné údaje internetových obchodov. Hlavným cieľom tohto vývoja je zvýšiť ekonomickú reprezentatívnosť indexov spotrebiteľských cien pri súčasnom znížení administratívnej záťaže a nákladov na zber údajov.

Teoretické základy zostavovania indexov spotrebiteľských cien podrobne opisuje manuál vydaný medzinárodnými organizáciami ILO, IMF, OECD, UNECE, Eurostatom a Svetovou bankou (ILO et al., 2004). Publikácia predstavuje základné princípy tvorby spotrebiteľských košov, výberu reprezentantov a konštrukcie elementárnych indexov spotrebiteľských cien. Nadväzujú na ňu monografie Balka (2008) a von der Lippeho (2007), ktoré rozvíjajú axiomatický prístup k teórii indexových čísel a analyzujú matematické vlastnosti jednotlivých indexových vzorcov. Významným príspevkom k modernej teórii indexových čísel sú aj práce Diewerta (1976; 2022), ktoré rozpracovali ekonomické základy superlatívnych indexov a problematiku reťazového vychýľovania (tzv. *chain drift*) pri dynamických spotrebných košoch.

Významný posun vo výskume nastal po rozšírení skenerových údajov, ktoré obsahujú informácie o skutočne realizovaných predajoch jednotlivých produktov. Na rozdiel od tradičného zisťovania cien umožňujú skenerové údaje pracovať so všetkými predanými položkami vrátane predaného množstva a tržieb, čím vytvárajú predpoklady na zostavovanie ekonomicky reprezentatívnejších vážených indexov spotrebiteľských cien. Spracovanie takýchto údajov si však vyžaduje nové metodické prístupy pri čistení údajov, filtrovaní extrémnych hodnôt, párovaní produktov a výbere vhodnej indexovej formuly. Tieto otázky podrobne analyzujú napríklad Białek a Beręsewicz (2021) alebo de Haan et al. (2021).

S rozvojom internetového predaja sa pozornosť výskumu presunula aj na využitie cien získavaných automatizovaným zberom údajov z internetových stránok, označovaným ako webscraping. Webscrapované údaje predstavujú ponukové ceny produktov z internetových obchodov, ktoré možno získavať vo vysokej frekvencii a bez potreby manuálneho zberu údajov. Výhodou tohto prístupu je možnosť monitorovať prakticky všetky produkty dostupné na internetovom trhu. Nevýhodou však zostáva skutočnosť, že ide o ponukové ceny, ktoré nemusia odrážať reálne nákupné správanie spotrebiteľov. Navyše webscrapované údaje neobsahujú informácie o predanom množstve ani o hodnote predaja, čo znemožňuje konštrukciu vážených indexov spotrebiteľských cien. Význam webscrapingu pre modernizáciu cenovej štatistiky analyzovali napríklad Polidoro et al. (2015).

Rozsiahly výskum sa venuje aj samotnej metodike výpočtu indexov spotrebiteľských cien pri nových zdrojoch údajov. Tradičné bilaterálne indexy, ktoré porovnávajú dve časové obdobia, sú pri vysokej dynamike produktov náchylné na vznik reťazového vychýľovania. Z tohto dôvodu sa čoraz väčšia pozornosť venuje

multilaterálnym metódam, ktoré porovnávajú viacero časových období súčasne. Medzi najvýznamnejšie patrí metóda GEKS, ktorú na časové porovnania cien rozpracovali Ivancic et al. (2011). Súčasný metodický rámec multilaterálnych indexov pre harmonizovaný index spotrebiteľských cien podrobne sumarizuje metodická príručka Eurostatu (Eurostat, 2022), ktorá sa stala referenčným dokumentom pre národné štatistické úrady pri implementácii skenerových údajov.

Problematike implementácie nových zdrojov údajov sa systematicky venuje aj Štatistický úrad Slovenskej republiky. H. Glaser-Opitzová (2019) predstavila metodické aspekty využitia skenerových údajov v cenovej štatistike a následne spolu v článku s P. Mazurekovou analyzovali vplyv dĺžky referenčného obdobia pri spracovaní skenerových údajov (Glaser-Opitzová & Mazureková, 2023). Na tieto práce nadviazali štúdie zamerané na webscrapované údaje, ktoré analyzovali agregáciu denných cien, multilaterálne indexy spotrebiteľských cien a využitie hedonických regresných modelov pri výpočte indexov spotrebiteľských cien (Knížat, 2023a; Knížat, 2023b; Knížat et al., 2024). Výsledky uvedených výskumov vytvorili metodický základ pre ďalší krok modernizácie cenovej štatistiky, ktorým je využitie transakčných údajov internetových obchodov.

Predkladaný článok nadväzuje na uvedené práce a rozširuje doterajší výskum implementácie alternatívnych zdrojov údajov do cenovej štatistiky. Na rozdiel od predchádzajúcich štúdií založených na webscrapovaných údajoch sa empirická analýza sústreďuje na transakčné údaje internetových obchodov, ktoré umožňujú porovnať správanie vážených a nevážených bilaterálnych aj multilaterálnych indexov spotrebiteľských cien a posúdiť ich potenciál pre budúcu produkciu oficiálnych cenových štatistík.

3. METODIKA

Táto kapitola podrobne opisuje metodiku použitú pri spracovaní transakčných údajov z online predajov a výpočte experimentálnych indexov spotrebiteľských cien. Najskôr predstavujeme zdroj údajov a ich základné charakteristiky, následne proces ich automatizovaného spracovania a výpočtu bilaterálnych a multilaterálnych indexových metód. Záverečná časť kapitoly je venovaná postupom na identifikáciu a odstránenie extrémnych cenových zmien, ktoré prispievajú k zvýšeniu kvality výsledných indexov.

3.1. TRANSAKČNÉ ÚDAJE Z ONLINE PREDAJOV

Modernizácia cenovej štatistiky je úzko spojená s implementáciou nových zdrojov údajov, ktoré umožňujú presnejšie zachytiť spotrebiteľské správanie na trhu. V podmienkach elektronického obchodu predstavujú perspektívny zdroj údajov transakčné údaje internetových obchodov, ktoré obsahujú informácie o skutočne realizovaných nákupoch spotrebiteľov.

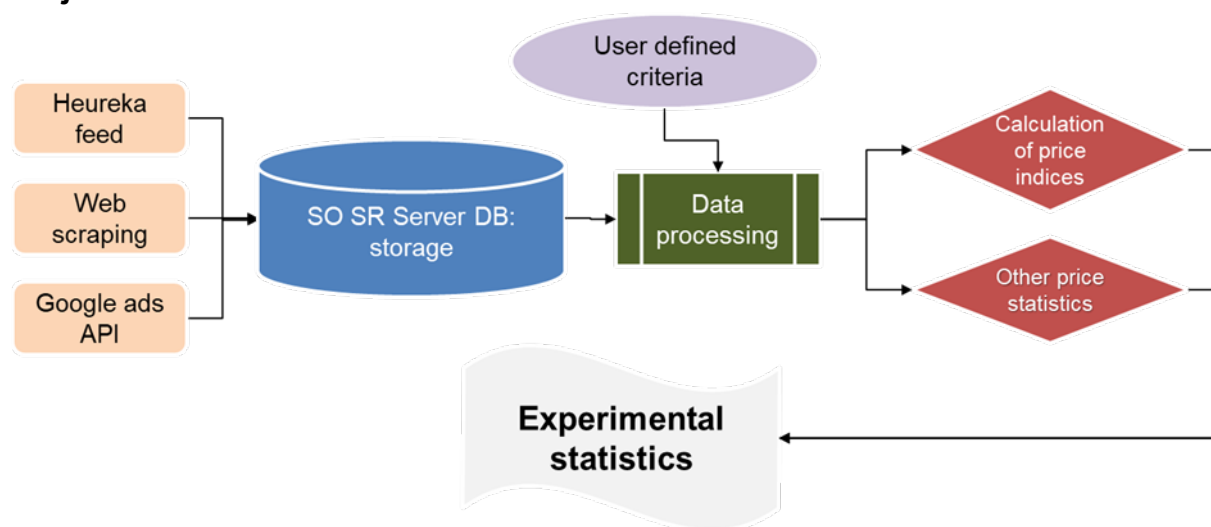
Na rozdiel od webscrapovaných údajov, ktoré predstavujú ponukové ceny produktov zverejnené na internetových stránkach, transakčné údaje obsahujú cenu, za ktorú bol produkt skutočne predaný, počet predaných kusov a hodnotu realizovaných tržieb za sledované obdobie. Tieto informácie umožňujú zostaviť vážené indexy spotrebiteľských cien, ktorých váhy sú odvodené od skutočných spotrebiteľských výdavkov.

Empirická analýza vychádza z experimentálnych údajov získavaných prostredníctvom platformy Heureka, ktorá združuje internetové obchody pôsobiace na slovenskom trhu. Platforma agreguje transakčné údaje jednotlivých internetových obchodov a poskytuje ich Štatistickému úradu Slovenskej republiky vo forme pravidelného denného dátového toku. Každý záznam obsahuje identifikáciu produktu, jeho charakteristiky, počet predaných kusov a hodnotu predaja za príslušný deň.

3.2. AUTOMATIZOVANÉ SPRACOVANIE ÚDAJOV

Implementácia online transakčných údajov z internetových obchodov do cenových štatistík si vyžaduje vytvorenie automatizovaného procesu spracovania údajov. Na obrázku 1 je znázornená štruktúra spracovania údajov použitá v experimentálnom prostredí Štatistického úradu Slovenskej republiky.

Obrázok 1: Automatizovaný proces spracovania transakčných a webscrapovaných údajov



Zdroj: nákras autora

Obrázok 1 demonštruje, že celý proces pozostáva z niekoľkých na seba nadväzujúcich krokov.

V prvom kroku platforma Heureka pravidelne poskytuje agregované transakčné údaje z internetových obchodov. Súčasťou systému sú aj údaje získané webscrapingom alebo prostredníctvom rozhraní API internetových platforiem.

V druhom kroku sa údaje automaticky ukladajú do databázového prostredia Štatistického úradu Slovenskej republiky. Centralizované úložisko zabezpečuje archiváciu jednotlivých dátových dávok, ich kontrolu a následné spracovanie.

Nasleduje etapa spracovania údajov, ktorá zahŕňa kontrolu kvality údajov, vytvorenie analytických dátových súborov a prípravu vstupov na výpočet indexov spotrebiteľských cien. V tejto fáze je možné aplikovať rôzne filtre na odstránenie extrémnych cenových zmien alebo iných neštandardných pozorovaní.

Poslednou etapou je výpočet experimentálnych indexov spotrebiteľských cien a ich využitie pri ďalších produktoch cenovej štatistiky alebo experimentálnych štatistikách.

Výhodou navrhnutého riešenia je vysoký stupeň automatizácie celého procesu. Po nastavení dátových tokov je možné jednotlivé kroky realizovať bez manuálnych zásahov, čím sa znižujú náklady na produkciu štatistických údajov a zároveň sa zvyšuje ich aktuálnosť.

3.3. BILATERÁLNE INDEXY SPOTREBITEĽSKÝCH CIEN

Na porovnanie vývoja cien sme použili reťazené bilaterálne indexy spotrebiteľských cien, ktoré porovnávajú ceny medzi dvoma časovými obdobiami. V článku analyzujeme nevážený Jevonsov index a vážený Törnqvistov index.

Jevonsov index patrí medzi najčastejšie používané elementárne indexy spotrebiteľských cien. Je definovaný ako geometrický priemer individuálnych cenových relatív všetkých produktov zaradených do spotrebného koša (Balk, 2008)

$$J_{0,t} = \prod_{i=1}^N \left(\frac{p_i^t}{p_i^0} \right)^{\frac{1}{N}}, \quad (1)$$

kde p_i^0 predstavuje cenu produktu i v bázičkom období 0, p_i^t cenu produktu v aktuálnom období t a $i = 1, \dots, N$ počet produktov zahrnutých do výpočtu.

Výhodou Jevonsovho indexu je jeho jednoduchá interpretácia a splnenie viacerých axiomatických vlastností indexových čísel (Balk, 2008). Keďže však všetkým produktom priraduje rovnakú váhu, nezohľadňuje rozdielny význam jednotlivých produktov z pohľadu spotrebiteľských výdavkov.

Na odstránenie tejto nevýhody možno využiť Törnqvistov index, ktorý využíva váhy odvodené z podielov výdavkov jednotlivých produktov v porovnávaných obdobiach (Balk, 2008)

$$T_{0,t} = \prod_{i \in S; t=1} \left(\frac{p_i^t}{p_i^0} \right)^{\frac{1}{2}(s_i^0 + s_i^t)}, \quad (2)$$

kde $s_i^0 = \frac{p_i^0 q_i^0}{\sum_{j=1}^N p_j^0 q_j^0}$ a $s_i^t = \frac{p_i^t q_i^t}{\sum_{j=1}^N p_j^t q_j^t}$ predstavujú podiel výdavkov produktu i na celkových výdavkoch v bázičkom 0 a aktuálnom období t , pričom q_j^0 a q_j^t predstavujú počet predaných kusov produktu i v bázičkom 0 a aktuálnom období t v danom poradí.

Použitie Törnqvistovho indexu je možné práve vďaka tomu, že online transakčné údaje obsahujú informácie o objeme predaja jednotlivých produktov. Na rozdiel od webscrapovaných údajov tak možno zostaviť ekonomicky reprezentatívnejší index spotrebiteľských cien.

3.4. MULTILATERÁLNE INDEXY SPOTREBITEĽSKÝCH CIEN

Pri vysokom počte nových a zanikajúcich produktov môžu bilaterálne indexy po ich reťazení vykazovať postupné systematické vychýľovanie (tzv. *chain drift*). Z tohto dôvodu sa na výpočet indexov spotrebiteľských cien čoraz častejšie využívajú multilaterálne metódy.

V tejto štúdii sme použili multilaterálnu metódu GEKS (navrhnutá: Gini, 1931; Éltető & Köves 1964; Szulc, 1964). Všeobecnú formu metódy GEKS môžeme zapísať ako (Eurostat, 2022)

$$GEKS_{0,t} = \prod_{l=0}^T (P_{0,t} \times P_{l,t})^{\frac{1}{T+1}}, \quad (3)$$

kde $P_{l,t}$ predstavuje bilaterálny index medzi referenčným obdobím l a obdobím t , pričom $T + 1$ označuje počet období zahrnutých do výpočtu. Ak sa ako bilaterálny index použije Jevonsov index definovaný v rovnici (1), vzniká nevážený multilaterálny GEKS-Jevonsov index, zatiaľ čo pri použití Törnqvistovho indexu definovaného v rovnici (2) vzniká vážený multilaterálny GEKS-Törnqvistov index.

V empirickej analýze sme použili 25-mesačné pohyblivé časové okno $T = 25$ ktoré zodpovedá odporúčaniam Eurostatu (Eurostat, 2022) pri zostavovaní multilaterálnych indexov spotrebiteľských cien.

Keďže pri každom pridaní nového mesiaca dochádza pri použití multilaterálnych metód k revízii historických hodnôt indexov, bolo potrebné použiť jednu zo spájacích metód, a to tzv. *half splice on published indices* (HASP). Tá zabezpečuje nadviazanie nového časového okna na už publikované indexy bez spätnej revízie historických údajov. Pre 25-mesačné okno ($2T + 1 = 25$ teda $T = 12$) sa HASP zapisuje ako (Eurostat, 2022)

$$P_{0,t}^{pub} = P_{0,t-12}^{pub} \times P_{t-12,t}^{[t-24,t]}, \quad t > 24, \quad (4)$$

kde $P_{0,t}^{pub}$ je publikovaný index medzi obdobiami 0 a t , $P_{0,t-12}^{pub}$ je už publikovaný index s pred 12 mesiacov a $P_{t-12,t}^{[t-24,t]}$ je multilaterálny index vypočítaný v aktuálnom 25-mesačnom okne $[t - 24, t]$. Takto zostavený index zachováva kontinuitu časového radu a zároveň využíva výhody multilaterálnych indexových metód.

3.5. FILTROVANIE EXTRÉMNYCH HODNÔT

Transakčné údaje z internetových obchodov predstavujú rozsiahly a heterogénny zdroj informácií, ktorý môže obsahovať veľmi veľa cenových pozorovaní za jednotlivé referenčné obdobia. Hoci ide o údaje založené na reálne uskutočnených nákupoch, ich využitie pri zostavovaní indexov spotrebiteľských cien si vyžaduje dôslednú kontrolu kvality. V dátach sa môžu vyskytovať extrémne cenové zmeny spôsobené chybami pri spracovaní údajov, zmenami balenia alebo množstva výrobku, jednorazovými marketingovými akciami alebo špecifickými obchodnými stratégiami predajcov. Ak by takéto pozorovania neboli identifikované a odstránené, mohli by významne ovplyvniť výsledné indexy spotrebiteľských cien, najmä pri použití vážených indexových metód. V tejto štúdii sme použili nasledujúce filtre (Eurostat, 2022; United Nations Task Team on Scanner Data, 2024b):

- **Extrémne zmeny cien** – prvý filter bol zameraný na identifikáciu extrémnych cenových zmien (outliers). Za extrémne sme považovali pozorovania, pri

ktorých sa ceny medzi dvoma po sebe nasledujúcimi obdobiami nachádzajú mimo intervalu 0,3 až 3.

$$0,3 \leq \frac{p_i^t}{p_i^0} \leq 3 \quad (5)$$

- **Dumpingové ceny** – druhý filter bol zameraný na identifikáciu dumpingových cien. Za dumpingové sme označili prípady, pri ktorých boli ceny menšie alebo sa rovnali 0,8 a súčasne relatívna zmena obratu nepresiahla hodnotu 0,2. Cieľom tohto filtra bolo eliminovať cenové zmeny, ktoré pravdepodobne neodrážajú bežný trhový vývoj, ale sú dôsledkom krátkodobých marketingových akcií alebo iných špecifických obchodných stratégií.

$$\frac{p_i^t}{p_i^{t-1}} \leq 0,8 \quad \wedge \quad \frac{S_i^t}{S_i^{t-1}} \leq 0,2, \quad (6)$$

kde S_i^t označuje obrat v období t .

UN-CEBD dokumentuje pri filtri na extrémne zmeny cien napríklad hranice 0,25 – 3 v Poľsku, 0,33 – 3 v Spojenom kráľovstve a 0,25 – 4 v Nórsku, zatiaľ čo pri dumpingovom cenovom filtri Nórsko používa kombináciu cenového relatíva 0,9 a relatíva predaja 0,2 (United Nations Task Team on Scanner Data, 2024a). V našom prípade sme pri spracovaní online transakčných údajov použili filtre analogické s filtrami aplikovanými pri spracovaní skenerových údajov v Štatistickom úrade Slovenskej republiky, ktoré sú porovnateľné s prahovými hodnotami dokumentovanými v medzinárodnej praxi.

4. VÝSLEDKY

V predchádzajúcej kapitole sme predstavili metodické postupy spracovania online transakčných údajov internetových obchodov a výpočtu experimentálnych indexov spotrebiteľských cien. V tejto časti tieto metodické prístupy aplikujeme na reálne transakčné údaje s cieľom analyzovať správanie jednotlivých indexových metód v praktických podmienkach. Pozornosť venujeme porovnaniu bilaterálnych a multilaterálnych indexov, hodnoteniu rozdielov medzi váženými a neváženými prístupmi a posúdeniu vplyvu filtrovania extrémnych cenových zmien na výsledné indexy spotrebiteľských cien.

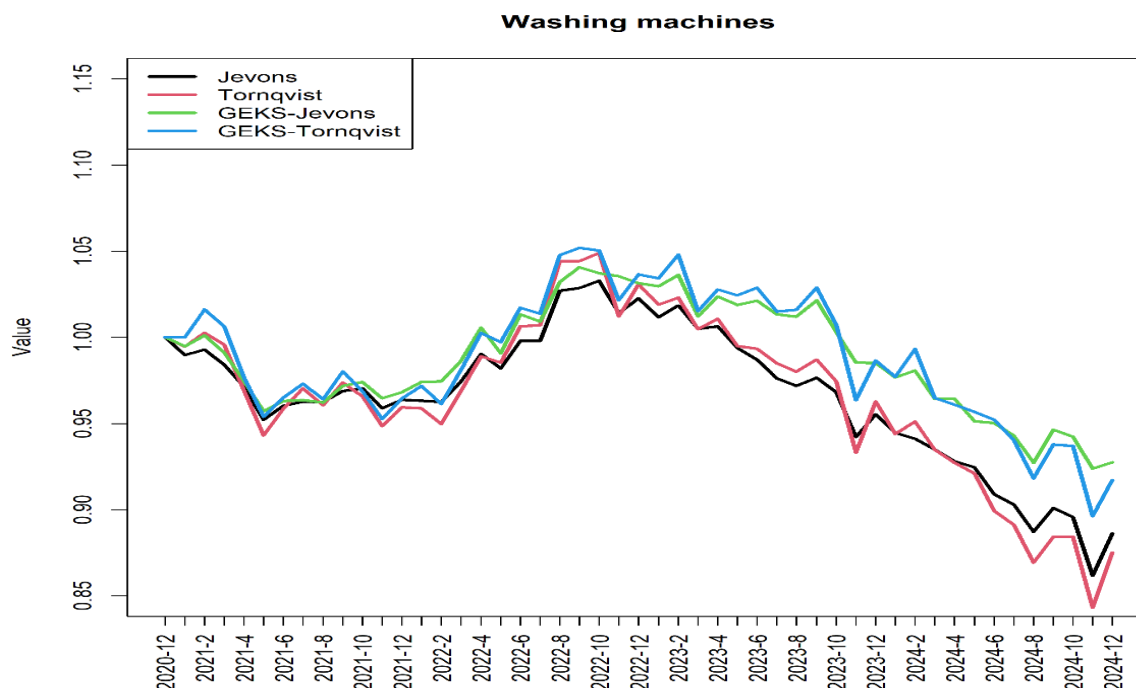
Experimentálna analýza bola realizovaná samostatne pre tri produktové skupiny – automatické práčky, umývačky riadu a sušičky bielizne. Tieto skupiny boli vybrané z dôvodu dostatočného rozsahu transakčných údajov a vysokej miery zastúpenia v elektronických obchodoch na platforme Heureka, čo umožňuje spoľahlivé porovnanie jednotlivých indexových metód.

Obrázky 2 až 4 zobrazujú experimentálne indexy spotrebiteľských cien vypočítané z online transakčných údajov internetových obchodov bez aplikácie filtrov uvedených v podkapitole 2.2. Pre každú produktovú skupinu sme porovnali štyri indexové metódy:

- bilaterálny nevážený Jevonsov index,
- bilaterálny vážený Törnqvistov index,

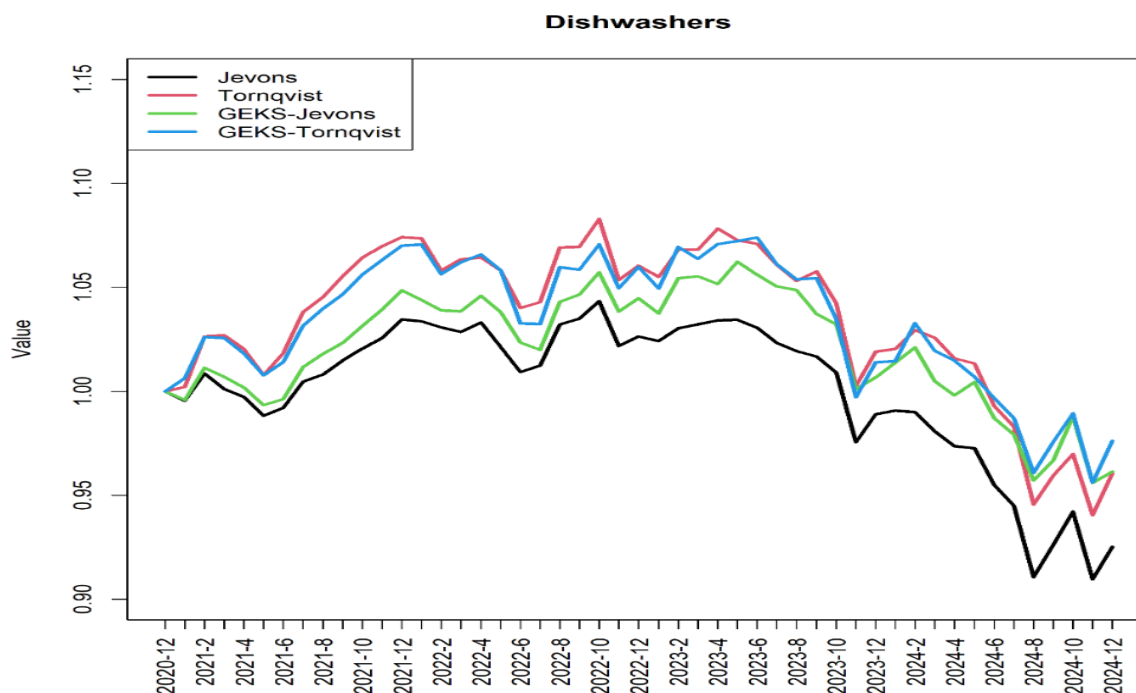
- multilaterálny nevážený GEKS-Jevonsov index,
- multilaterálny vážený GEKS-Törnqvistov index.

Obrázok 2: Automatické práčky (Washing machines) – bez použitia filtrov



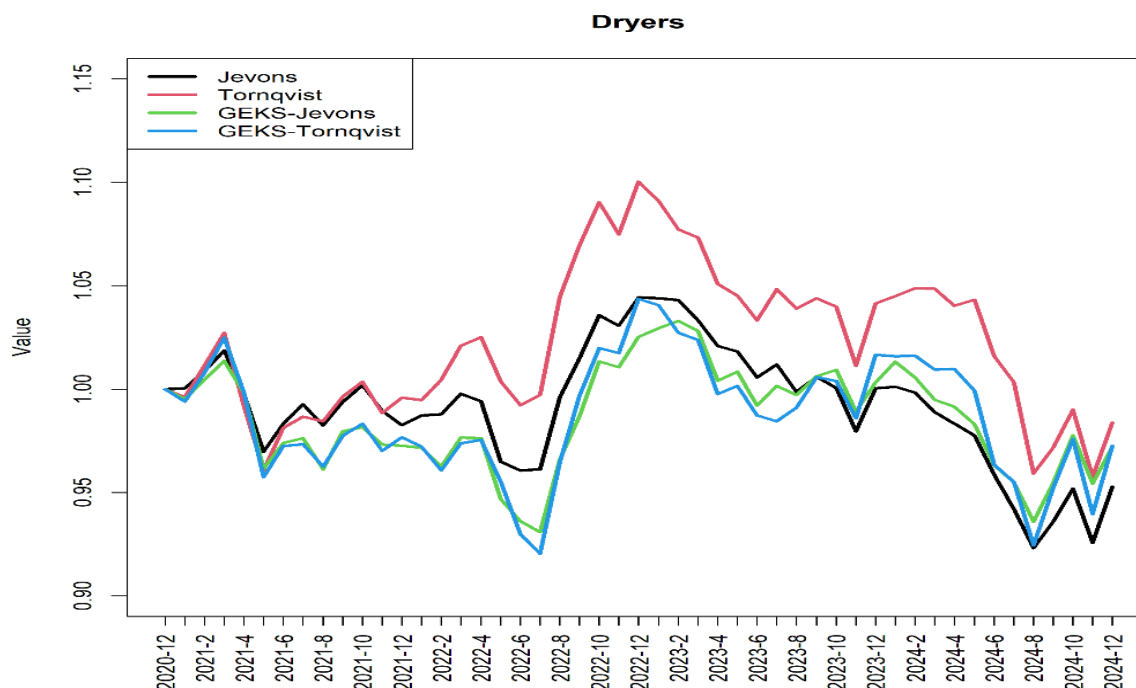
Zdroj údajov: Štatistický úrad Slovenskej republiky/Heureka, výpočet autora

Obrázok 3: Umývačky riadu (Dishwashers) – bez použitia filtrov



Zdroj údajov: Štatistický úrad Slovenskej republiky/Heureka, výpočet autora

Obrázok 4: Sušičky bielizne (Dryers) – bez použitia filtrov



Zdroj údajov: Štatistický úrad Slovenskej republiky/Heureka, výpočet autora

Na obrázku 2 možno pozorovať, že všetky štyri indexy vykazujú veľmi podobný priebeh. Po počiatočnej stabilizácii cien v roku 2021 nasleduje rast cien počas rokov 2022 a začiatkom roka 2023, po ktorom dochádza k postupnému poklesu až do konca analyzovaného obdobia. Bilaterálne indexy (Jevonsov a Törnqvistov) dosahujú vo všeobecnosti nižšie hodnoty ako multilaterálne metódy GEKS. Rozdiel medzi váženým a neváženým prístupom je pomerne malý, zatiaľ čo multilaterálne indexy vykazujú počas väčšiny analyzovaného obdobia stabilnejší priebeh a zmiernujú krátkodobé výkyvy spôsobené zmenami produktového portfólia.

Zároveň na obrázku 2 možno pozorovať, že najmä od roku 2023 sa multilaterálne indexy GEKS-Jevonsov a GEKS-Törnqvistov nachádzajú na mierne vyššej úrovni ako príslušné bilaterálne indexy. Jedným z vysvetlení môže byť dynamika produktového sortimentu, napríklad vstup a výstup jednotlivých produktov, keďže multilaterálne metódy využívajú informácie z viacerých období súčasne a sú konštruované tak, aby lepšie pracovali s meniacim sa produktovým univerzom. Ďalej môže pozorovaný rozdiel súvisieť aj s tým, že v tejto časti časového radu sa pri multilaterálnych indexoch začína uplatňovať spájanie (splicing) nadväzujúcich časových okien. Pozorované rozdiely tak zároveň poukazujú na potrebu ďalšieho skúmania vplyvu obmeny produktov a spôsobu spájania multilaterálnych indexov.

Výsledky pre umývačky riadu, znázornené na obrázku 3, poukazujú predovšetkým na rozdiel medzi váženým a neváženým prístupom. Jevonsov index dosahuje počas analyzovaného obdobia najnižšie hodnoty, zatiaľ čo Törnqvistov index sa pohybuje na najvyššej úrovni počas väčšiny analyzovaného obdobia. Výraznejší rozdiel medzi týmito dvoma bilaterálnymi metódami môže súvisieť s použitím váh v Törnqvistovom indexe. Ak sa cenový vývoj produktov s vyšším podielom na realizovaných tržbách odlišuje od cenového vývoja menej predávaných produktov, vážený index môže

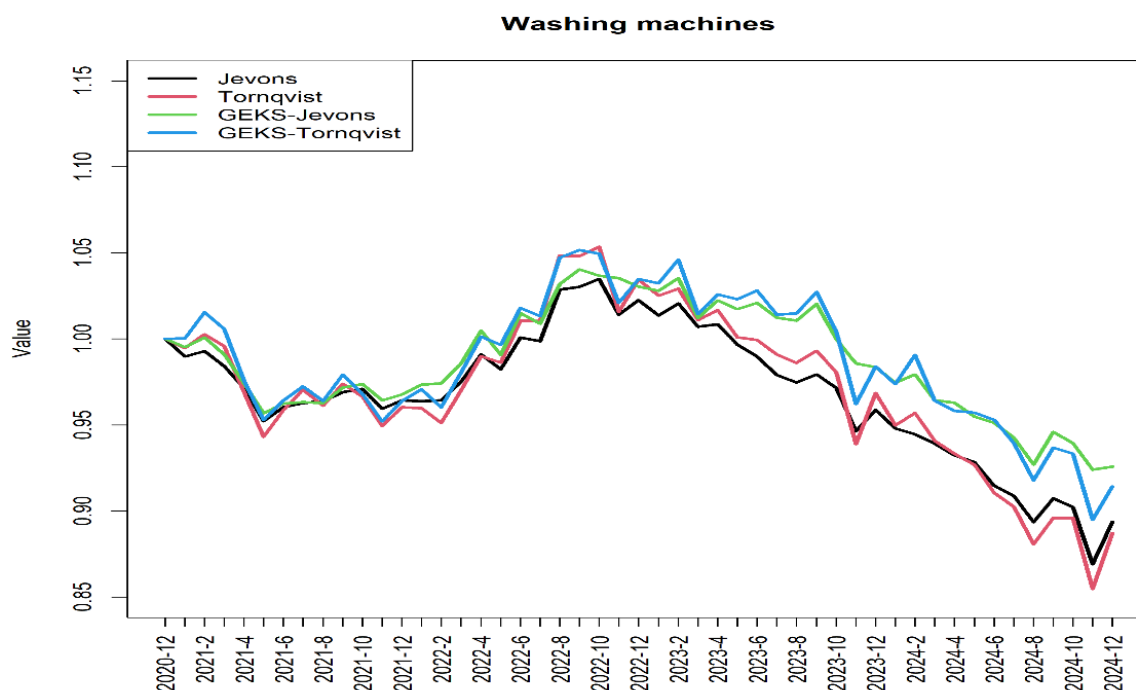
vykazovať odlišný priebeh od neváženého Jevonsovho indexu, ktorý prisudzuje jednotlivým cenovým relatívam rovnakú váhu. V prípade umývačiek riadu tak výsledky naznačujú, že štruktúra realizovaných výdavkov môže zohrávať významnejšiu úlohu pri výslednej hodnote indexu. Pri multilaterálnych indexoch sú rozdiely menej jednoznačné, čo naznačuje, že pozorovanú diferenciáciu výsledkov pri tejto produktovej skupine nemožno pripísať iba voľbe medzi bilaterálnym a multilaterálnym prístupom.

Pri sušičkách bielizne (obrázok 3) možno pozorovať špecifické správanie Törnqvistovho indexu, ktorý sa najmä v druhej polovici roka 2022 a počas roka 2023 výraznejšie odlišuje od ostatných troch indexových metód a dosahuje vyššie hodnoty. Naproti tomu Jevonsov, GEKS-Jevonsov a GEKS-Törnqvistov index vykazujú navzájom podstatne podobnejší priebeh. Pozorovaný rozdiel teda nemožno interpretovať ako všeobecnú odlišnosť medzi bilaterálnymi a multilaterálnymi metódami, ale predovšetkým ako špecifické správanie váženého bilaterálneho Törnqvistovho indexu. Jedným z možných vysvetlení je štruktúra realizovaných výdavkov, keďže Törnqvistov index zohľadňuje výdavkové váhy jednotlivých produktov. Výraznejšie cenové zmeny produktov s vysokým podielom na tržbách tak môžu mať väčší vplyv na jeho výslednú hodnotu. Skutočnosť, že GEKS-Törnqvistov index napriek využitiu váženého Törnqvistovho indexu ako bilaterálneho stavebného prvku zostáva bližšie k ostatným indexom, zároveň poukazuje na vplyv multilaterálnej konštrukcie GEKS, ktorá využíva cenové porovnania vo viacerých obdobiach.

Celkovo výsledky experimentu potvrdzujú, že voľba indexovej metódy ovplyvňuje výsledný priebeh indexu spotrebiteľských cien. Rozdiely medzi váženými a neváženými indexmi závisia od štruktúry spotrebiteľských výdavkov v jednotlivých produktových skupinách, zatiaľ čo multilaterálne metódy vo všetkých prípadoch vykazujú stabilnejší priebeh ako ich bilaterálne ekvivalenty.

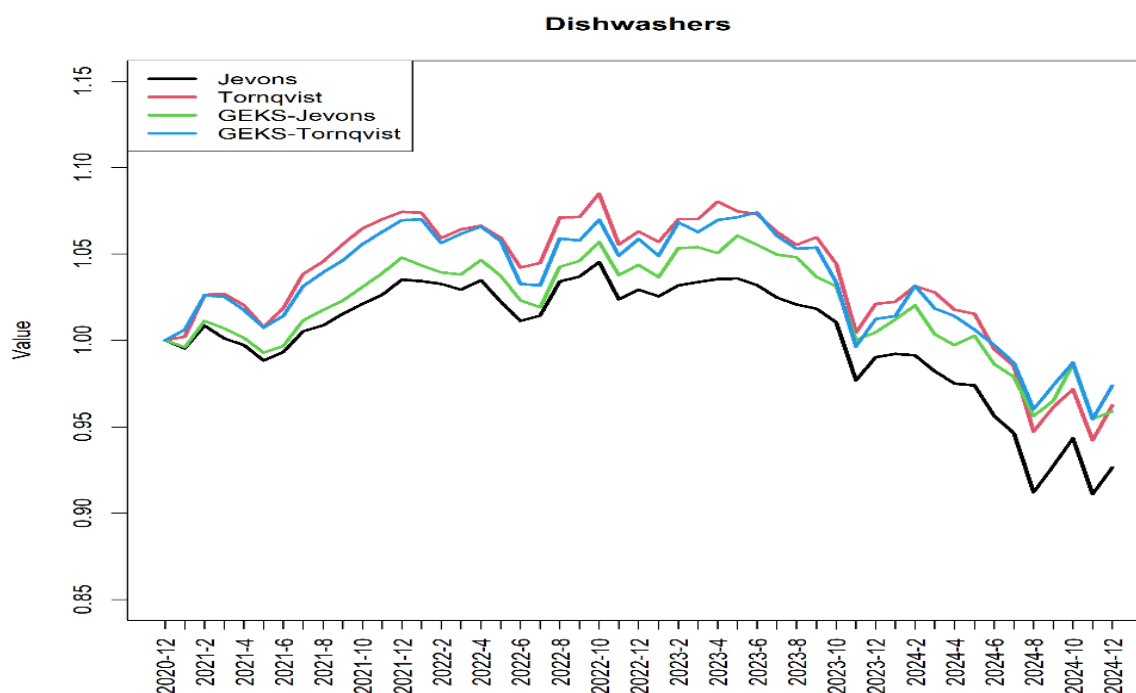
Podobne obrázky 5 až 7 zobrazujú experimentálne indexy spotrebiteľských cien vypočítané z online transakčných údajov internetových obchodov, tentoraz s aplikáciou filtrov uvedených v podkapitole 3.5.

Obrázok 5: Automatické práčky (Washing machines) – s použitím filtrov



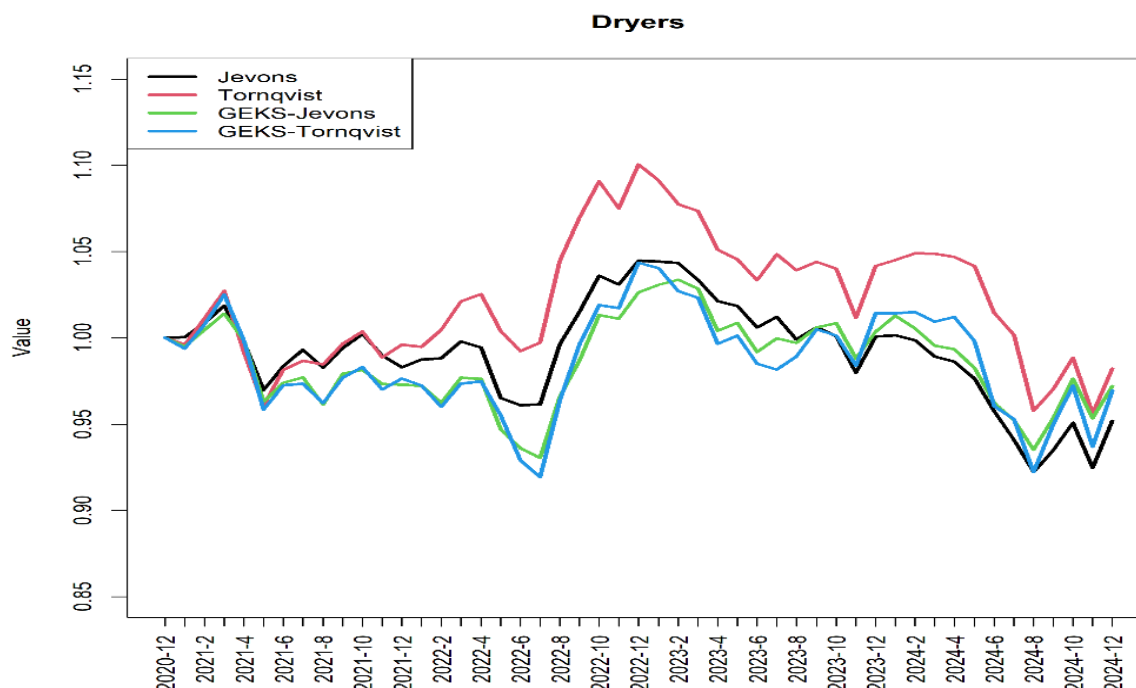
Zdroj údajov: Štatistický úrad Slovenskej republiky/Heureka, výpočet autora

Obrázok 6: Umývačky riadu (Dishwashers) – s použitím filtrov



Zdroj údajov: Štatistický úrad Slovenskej republiky/Heureka, výpočet autora

Obrázok 7: Sušičky bielizne (Dryers) – s použitím filtrov



Zdroj údajov: Štatistický úrad Slovenskej republiky/Heureka, výpočet autora

Pri porovnaní s výsledkami bez aplikácie filtrov možno konštatovať, že základný trend cenového vývoja zostáva vo všetkých produktových skupinách prakticky nezmenený. Aplikácia filtrov však vo viacerých prípadoch redukuje krátkodobé cenové výkyvy spôsobené extrémnymi pozorovaniami, čím prispieva k vyššej stabilite výsledných časových radov. Filtrovanie teda nemení ekonomickú interpretáciu jednotlivých indexových metód, ale zvyšuje robustnosť výsledných odhadov.

Na obrázku 5 možno pozorovať, že po aplikácii filtrov zostáva priebeh všetkých štyroch indexov veľmi podobný výsledkom získaným bez filtrovania. Základný trend rastu cien počas rokov 2022 až 2023 a ich následného poklesu v roku 2024 je zachovaný. V porovnaní s obrázkom 2 sú však krátkodobé výkyvy jednotlivých indexov miernejšie a rozdiely medzi nimi sa v niektorých obdobiach znižujú. Multilaterálne indexy GEKS-Jevons a GEKS-Törnqvist si naďalej zachovávajú najstabilnejší priebeh, zatiaľ čo bilaterálne indexy vykazujú o niečo vyššiu citlivosť na mesačné cenové zmeny.

Výsledky pre umývačky riadu po aplikácii filtrov sú znázornené na obrázku 6. V porovnaní s obrázkom 3 zostáva zachovaný celkový charakter časového radu aj vzájomné usporiadanie jednotlivých indexových metód. Súčasne však možno pozorovať mierne zníženie krátkodobej volatility najmä pri multilaterálnych indexoch. Jevonsov index naďalej dosahuje najnižšie hodnoty počas väčšiny analyzovaného obdobia, zatiaľ čo Törnqvistov a GEKS-Törnqvistov index zostávajú na vyššej úrovni. Výsledky potvrdzujú, že filtrovanie odstraňuje predovšetkým lokálne extrémny, pričom rozdiely vyplývajúce z použitia výdavkových váh zostávajú zachované.

Najvýraznejší vplyv filtrov možno pozorovať pri sušičkách bielizne (obrázok 7). V porovnaní s výsledkami bez filtrovania (obrázok 4) dochádza k zmierneniu extrémnych odchýlok Törnqvistovho indexu. Napriek tomu si zachováva vyššie hodnoty počas značnej časti analyzovaného obdobia, čo odráža vplyv výdavkových váh pri produktoch s vyšším objemom predaja. Multilaterálne indexy aj naďalej vykazujú najplynulejší priebeh a menej reagujú na krátkodobé cenové výkyvy než bilaterálne metódy.

Celkovo porovnanie obrázkov 2 až 4 s obrázkami 5 až 7 ukazuje, že aplikácia filtrov nemení dlhodobý trend cenového vývoja, ale prispieva k eliminácii extrémnych pozorovaní a znižuje volatilitu výsledných indexov. Tento efekt je najvýraznejší pri produktových skupinách s vyššou variabilitou cien, zatiaľ čo pri stabilnejších skupinách predstavuje filtrovanie skôr jemnú korekciu výsledných časových radov. Experiment tak potvrdzuje, že kombinácia automatizovaného filtrovania, výdavkových váh a multilaterálnych indexových metód predstavuje vhodný metodický prístup k spracovaniu rozsiahlych online transakčných údajov pri zostavovaní experimentálnych indexov spotrebiteľských cien.

Z porovnania jednotlivých indexových metód vyplýva, že zahrnutie informácie o realizovaných tržbách môže mať na výsledný cenový index nezanedbateľný vplyv, pričom jeho rozsah sa medzi analyzovanými produktovými skupinami líši. Vážený Törnqvistov index a vážený GEKS-Törnqvistov index zohľadňujú ekonomický význam jednotlivých produktov podľa ich podielu na realizovaných výdavkoch, zatiaľ čo nevážený Jevonsov index a nevážený GEKS-Jevonsov index prisudzujú cenovým relatívam jednotlivých produktov rovnaký význam. Výsledky preto poukazujú na dôležitú pridanú hodnotu transakčných údajov, ktoré umožňujú pri výpočte indexov využiť popri cenách aj informácie o skutočne realizovaných nákupoch. Na základe uskutočneného experimentu však nemožno jednoznačne konštatovať všeobecnú preferenciu vážených indexov pred neváženými, keďže rozdiely medzi nimi závisia od konkrétnej produktovej skupiny a štruktúry realizovaných výdavkov. Výsledky skôr podporujú ďalšie skúmanie vážených metód ako perspektívneho prístupu pri využívaní transakčných údajov v cenovej štatistike, pričom ich prípadná preferencia by mala vychádzať aj z ďalšieho hodnotenia reprezentatívnosti transakčných údajov a stability výsledkov na širšom súbore produktových skupín.

5. ZÁVER

Digitalizácia maloobchodu vytvára nové možnosti na modernizáciu cenovej štatistiky a postupnú transformáciu tradičných spôsobov zisťovania cien. Transakčné údaje z internetových obchodov predstavujú perspektívny zdroj údajov, ktorý na rozdiel od webscrapovaných údajov umožňuje zachytiť nielen ceny produktov, ale aj skutočné spotrebiteľské správanie prostredníctvom informácií o počte predaných kusov a hodnote realizovaných tržieb. Tieto údaje vytvárajú predpoklady na zostavovanie ekonomicky reprezentatívnejších indexov spotrebiteľských cien založených na reálnych spotrebiteľských výdavkoch.

Článok predstavil metodický rámec spracovania online transakčných údajov od ich automatizovaného získavania až po výpočet bilaterálnych a multilaterálnych indexov spotrebiteľských cien. Experimentálna analýza na údajoch platformy Heureka ukázala, že použitie transakčných údajov umožňuje aplikáciu vážených indexových metód, ktoré lepšie zohľadňujú význam jednotlivých produktov z pohľadu spotrebiteľských

výdavkov. Súčasne sa potvrdilo, že multilaterálne indexy vykazujú stabilnejší priebeh časových radov ako bilaterálne indexy, čím prispievajú k eliminácii problémov spojených s častou obmenou sortimentu v internetových obchodoch.

Výsledky experimentu zároveň potvrdili význam predspracovania údajov prostredníctvom filtrov extrémnych cenových zmien a dumpingových cien. Aplikácia filtrov nemení dlhodobý trend cenového vývoja, avšak prispieva k eliminácii extrémnych pozorovaní a k zvýšeniu robustnosti výsledných indexov spotrebiteľských cien. Tento prístup je dôležitý najmä pri spracovaní rozsiahlych transakčných databáz, kde sa prirodzene vyskytujú krátkodobé cenové výkyvy spôsobené marketingovými akciami alebo špecifickými obchodnými stratégiami predajcov.

Predstavená metodika sa v súčasnosti využíva na experimentálne zostavovanie indexov spotrebiteľských cien v podmienkach Štatistického úradu Slovenskej republiky a predstavuje ďalší krok v implementácii alternatívnych zdrojov údajov do produkcie oficiálnych štatistík. Získané poznatky zároveň vytvárajú metodický základ na rozšírenie experimentálnych analýz na ďalšie skupiny výrobkov a na ďalší výskum zameraný na porovnanie rôznych multilaterálnych indexových metód, optimalizáciu filtračných postupov a hodnotenie možností implementácie transakčných údajov do pravidelnej produkcie indexov spotrebiteľských cien.

So zámerom vylúčiť konflikt záujmov Peter Knížat ako autor článku neparticipoval na per rollam komunikácii a rokovaníach redakčnej rady časopisu Slovenská štatistika a demografia vo veciach týkajúcich sa tohto článku.

LITERATÚRA

- Balk, B. M. (2008). *Price and Quantity Index Numbers: Models for Measuring Aggregate Change and Difference*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511720758>
- Białek, J., & Beręsewicz, M. (2021). Scanner Data in Inflation Measurement: From Raw Data to Price Indices. *Statistical Journal of the IAOS*, 37(4), 1315 – 1336. <https://doi.org/10.3233/SJI-210816>
- de Haan, J., Hendriks, R., & Scholz, M. (2021). Price Measurement Using Scanner Data: Time-Product Dummy Versus Time Dummy Hedonic Indexes. *Review of Income and Wealth*, 67(2), 394 – 417. <https://doi.org/10.1111/roiw.12468>
- Diewert, W. E. (1976). Exact and Superlative Index Numbers. *Journal of Econometrics*, 4(2), 115 – 145. [https://doi.org/10.1016/0304-4076\(76\)90009-9](https://doi.org/10.1016/0304-4076(76)90009-9)
- Diewert, W. E. (2022). Scanner Data, Elementary Price Indexes and the Chain Drift Problem. In D. Chotikapanich, A. N. Rambaldi, & N. Rohde (Eds.), *Advances in Economic Measurement: A volume in honour of D. S. Prasada Rao* (pp. 445 – 606). Palgrave Macmillan. https://doi.org/10.1007/978-981-19-2023-3_11
- Éltető, Ö., & Köves, P. (1964). Egy nemzetközi összehasonlításoknál fellépő indexszámítási problémáról. *Statisztikai Szemle* 42(5), 507 – 518.
- Eurostat. (2022). *Guide on Multilateral Methods in the Harmonised Index of Consumer Prices* (2022 ed.). Publication Office of the European Union. <https://doi.org/10.2785/873932>

- Gini, C. (1931). On the Circular Test of Index Numbers, *Metron* 9(2), 3 – 24.
- Glaser-Opitzová, H. (2019). Nové zdroje údajov pre cenovú štatistiku a metódy ich spracovania. *Slovenská štatistika a demografia/Slovak Statistics and Demography*, 29(4), 49 – 66. <https://ssad.statistics.sk/SSaD/index.php/nove-zdroje-udajov-pre-cenovu-statistiku-a-metody-ich-spracovania/>
- Glaser-Opitzová, H., & Mazureková, P. (2023). Vplyv parciálneho časového pokrytia údajov zo skenerov na presnosť cenových indexov. *Slovenská štatistika a demografia/Slovak Statistics and Demography*, 33(3), 39 – 54. <https://ssad.statistics.sk/SSaD/index.php/vplyv-parcialneho-casoveho-pokrytia-udajov-zo-skenerov-na-presnost-cenovych-indexov/>
- ILO, IMF, OECD, UNECE, Eurostat, & The World Bank. (2004). *Consumer Price Index Manual: Theory and Practice*. International Labour Office.
- Ivancic, L., Diewert, W. E., & Fox, K. J. (2011). Scanner Data, Time Aggregation and the Construction of Price Indexes. *Journal of Econometrics*, 161(1), 24 – 35. <https://doi.org/10.1016/j.jeconom.2010.09.003>
- Knížat, P. (2023a). Hedonic Consumer Price Index: Diagnostics and Analysis of Variance. *Slovenská štatistika a demografia/Slovak Statistics and Demography*, 33(3), 21 – 38. <https://ssad.statistics.sk/SSaD/index.php/hedonic-consumer-price-index-diagnostics-and-analysis-of-variance/?lang=en>
- Knížat, P. (2023b). Web Scraped Data in Consumer Price Indices. *Statistical Journal of the IAOS*, 39(1), 203 – 212. <https://doi.org/10.3233/SJI-220115>
- Knížat, P., Glaser-Opitzová, H., Furková, A., & Vojtková, M. (2024). Multilateral Indices in Official Price Statistics and a New Additive Splicing Method. *Quality & Quantity*, 58(5), 4207 – 4222. <https://doi.org/10.1007/s11135-024-01848-3>
- Polidoro, F., Giannini, R., Lo Conte, R., Mosca, S., & Rossetti, F. (2015). Web scraping techniques to collect data on consumer electronics and airfares for Italian HICP compilation. *Statistical Journal of the IAOS*, 31(2), 165 – 176. <https://doi.org/10.3233/sji-150901>
- Szulc, B. (1964). Indeksy dla porównań wieloregionalnych. *Przegląd Statystyczny*, 3, 239 – 254.
- United Nations Task Team on Scanner Data. (2024a, May 8). *Current practice of data filtering*. UN Statistics Wiki. Retrieved September 7, 2026, from <https://millenniumindicators.un.org/wiki/spaces/GWGSD/pages/240910484/Current%2Bpractice%2Bof%2Bdata%2Bfiltering>
- United Nations Task Team on Scanner Data. (2024b, May 8). *Data filtering and missing prices*. UN Statistics Wiki. Retrieved September 10, 2026, from <https://millenniumindicators.un.org/wiki/spaces/GWGSD/pages/240910480/Data+filtering+and+missing+prices>
- von der Lippe, P. (2007). *Index Theory and Price Statistics*. Peter Lang.

RESUMÉ

Článok sa venuje možnostiam využitia transakčných údajov z internetových obchodov pri zostavovaní indexov spotrebiteľských cien v podmienkach Štatistického úradu Slovenskej republiky. Predstavuje automatizovaný proces získavania a spracovania údajov, metodiku výpočtu bilaterálnych a multilaterálnych indexov spotrebiteľských cien a postupy filtrovania extrémnych cenových zmien. Experimentálna analýza využíva údaje platformy Heureka o vybraných skupinách domácich spotrebičoch a porovnáva nevážené a vážené indexové metódy. Výsledky ukazujú, že transakčné údaje umožňujú zostavovať ekonomicky reprezentatívnejšie vážené indexy spotrebiteľských cien, pričom multilaterálne metódy poskytujú stabilnejší priebeh

časových radov ako tradičné bilaterálne indexy. Zároveň sa potvrdil pozitívny vplyv filtrovania extrémnych cenových zmien na robustnosť výsledných indexov. Príspevok poukazuje na význam transakčných údajov ako perspektívneho zdroja údajov na ďalšiu modernizáciu oficiálnej cenovej štatistiky.

RESUME

The paper examines the potential use of e-commerce transaction data for the compilation of consumer price indices at the Statistical Office of the Slovak Republic. It presents an automated framework for data acquisition and processing, the methodology for calculating bilateral and multilateral consumer price indices, and procedures for filtering extreme price changes. The empirical analysis is based on transaction data provided by the Heureka platform for selected household appliance product groups and compares weighted and unweighted index methods. The results show that transaction data enable the construction of economically more representative weighted consumer price indices, while multilateral methods provide more stable time series than traditional bilateral indices. The findings also confirm the effect of filtering extreme price changes on the robustness of the resulting price indices. The paper highlights the potential of transaction data from online retailers as a valuable source for the further modernisation of official price statistics.

PROFESIJNÝ ŽIVOTOPIS

Peter Knížat, PhD., MSc., pôsobí v Štatistickom úrade Slovenskej republiky (ŠÚ SR), kde je zodpovedný za výskum a vývoj inovácií oficiálnych a experimentálnych štatistík využitím nových dátových zdrojov. V ŠÚ SR tiež pôsobí ako školiťel odborných kurzov, ktoré sú zamerané na softvérové prostredie R, dátovú analýzu a ekonometriu, a je členom redakčnej rady vedeckého časopisu *Slovenská štatistika a demografia*. V rámci Európskeho štatistického systému (ESS) zastupuje ŠÚ SR v expertnej skupine ESS Innovation Network a je spoluškoliťelom odborného kurzu *Econometrics in macroeconomic statistics*. Súčasne pôsobí na Katedre ekonómie a financií Fakulty managementu Univerzity Komenského v Bratislave ako akademický pracovník, kde prispieva k výučbe predmetov zameraných na matematiku a štatistiku. Jeho výskum je zameraný na analýzu funkcionálnych údajov, teóriu cenových indexov, viackritériové rozhodovacie metódy, priestorové regresné a ekonometrické modely.

KONTAKT

peter.knizat@statistics.sk
knizat3@uniba.sk