

Uniwersytet Łódzki
Wydział Ekonomiczno-Socjologiczny

Pomiar dynamiki cen produktów skanowanych

Mgr Natalia Pawelec

Autoreferat rozprawy doktorskiej
przygotowanej pod kierunkiem
naukowym
prof. dr. hab. Jacka Białka

Łódź, 2026

Podziękowania kieruję do mojego Promotora, Profesora Jacka Białka za zaangażowanie, cierpliwość i merytoryczne wsparcie na każdym etapie powstawania tej pracy.

Spis treści

Spis treści	5
Streszczenie w języku polskim	6
Streszczenie w języku angielskim	6
Wprowadzenie.....	7
Przedmiot i zakres badań.....	10
Cele, problemy i hipotezy badawcze.....	18
Synteza wyników oraz wkład do literatury	24
Bibliografia.....	31
Artykuł 1 The use of scanner data for price indices calculation: the case of Poland....	33
Artykuł 2 Proposition of the quality adjusted GEKS-type price index.....	58
Artykuł 3 Estimating the elasticity of substitution when compiling the CES cost of living index on scanner data	74
Artykuł 4 The use of transitive Montgomery indicators for scanner data analysis.....	100
Artykuł 5 Zastosowanie indykatorów Benneta i Montgomery’ego w analizie danych skanowanych – porównanie własności i oszacowań efektów	114
Artykuł 6 Comparison of transitive Bennet and Montgomery indicators on the basis of scanner data sets	128
Załączniki – Oświadczenia współautorów	143

Streszczenie w języku polskim

Przedstawiony cykl publikacji dotyczy pomiaru inflacji z wykorzystaniem danych skanowanych oraz analizy własności indeksów i indykatorów cen i ilości w warunkach wysokiej dynamiki struktury koszyka inflacyjnego. Badania obejmują zarówno ocenę istniejących metod, jak i rozwój nowych rozwiązań metodycznych oraz ich weryfikację empiryczną. Uzyskane wyniki wskazują na ograniczenia klasycznych indeksów bilateralnych oraz na rekomendację metod dostosowanych do specyfiki danych wysokiej szczegółowości i częstotliwości (indeksów i indykatorów multilateralnych). W szczególności, w cyklu pracy podkreślono rolę decyzji metodologicznych po stronie urzędu statystycznego, takich jak wybór poziomu agregacji danych i procedury filtrowania danych, jak również wskazano czynniki determinujące wyniki pomiaru dynamiki cen. Wkład pracy polega na rozszerzeniu narzędzi analizy inflacji oraz dostarczeniu wniosków istotnych dla nowoczesnej statystyki cen.

Streszczenie w języku angielskim

The presented publication cycle concerns the measurement of inflation using scanner data and the analysis of the properties of price and quantity indices and indicators under conditions of high dynamics in the structure of the inflation basket. The research encompasses both an evaluation of existing methods and the development of new methodological solutions, along with their empirical verification. The obtained results indicate the limitations of classical bilateral indices and support the recommendation of methods tailored to the specific characteristics of high-granularity and high-frequency data (multilateral indices and indicators). In particular, the cycle of works emphasises the role of methodological decisions on the part of the statistical office, such as the choice of data aggregation level and data filtering procedures, while also identifying the factors that determine the results of price dynamics measurement. The contribution of the work lies in extending the toolkit for inflation analysis and in providing conclusions of relevance to modern price statistics.

Wprowadzenie

Inflacja należy do tych zjawisk ekonomicznych, których konsekwencje dotyczą niemal każdego uczestnika życia gospodarczego: pracowników, przedsiębiorców, osoby o stałych dochodach, oszczędzających, kredytobiorców czy też kredytodawców. Gdy inflacja staje się mniej przewidywalna, zarówno konsumenci, jak i przedsiębiorstwa stają wobec rosnącej niepewności co do przyszłych stóp procentowych, wynagrodzeń, podatków i zysków, co może prowadzić do opóźniania decyzji inwestycyjnych i ograniczenia aktywności gospodarczej (Federal Reserve Bank of Cleveland, 2026).

Na podstawie wskaźnika cen towarów i usług konsumpcyjnych (CPI) banki centralne podejmują decyzje o stopach procentowych, rządy przeprowadzają waloryzacje świadczeń społecznych i emerytur, a instytucje finansowe indeksują umowy kredytowe i kontrakty międzybankowe. Wartość inflacji wpływa zatem na najważniejsze decyzje makroekonomiczne i codzienne decyzje finansowe. Potwierdzeniem tej zależności jest fakt, że Europejski Bank Centralny uzasadniał wybór dodatniego celu inflacyjnego na poziomie zbliżonym do 2% występowaniem potencjalnego obciążenia pomiarowego wskaźnika HICP (Nickel i in., 2021). Oznacza to, że sama konstrukcja polityki pieniężnej Eurosystemu uwzględnia niedoskonałości związane z pomiarem inflacji. Niedoszacowanie lub przeszacowanie inflacji generuje zatem realne koszty ekonomiczne zarówno dla gospodarstw domowych, jak i dla całej gospodarki (Hałka A., Leszczyńska A. 2011).

Trafność pomiaru inflacji nabiera szczególnego znaczenia w warunkach dynamicznych zmian struktury rynku detalicznego. Postępująca digitalizacja handlu, rosnąca segmentacja oferty produktowej i zmienność preferencji konsumentów sprawiają, że tradycyjne metody pomiaru cen oparte na bezpośrednich notowaniach ankierskich i statycznym koszyku inflacyjnym coraz wyraźniej tracą na reprezentatywności. W konsekwencji rośnie ryzyko systematycznego obciążenia pomiaru, które bezpośrednio przekłada się na jakość decyzji podejmowanych przez banki centralne, rządy i inwestorów. W tych warunkach pytanie o sposób konstrukcji indeksu cenowego, wybór formuły indeksowej, źródeł danych i metod agregacji przestaje być kwestią wyłącznie techniczną i staje się zagadnieniem o fundamentalnym znaczeniu ekonomicznym. To właśnie doskonalenie metodologii pomiaru dynamiki cen,

ze szczególnym uwzględnieniem nowoczesnych źródeł danych i ich specyfiki, stanowi oś niniejszego cyklu publikacji.

Prezentowana rozprawa doktorska ma formę cyklu sześciu artykułów naukowych, które tworzą spójną całość podporządkowaną jednemu, nadrzędnemu celowi badawczemu: **ocenie i rozwojowi metod pomiaru dynamiki cen produktów skanowanych, ze szczególnym uwzględnieniem własności multilateralnych wersji indeksów i indyktorów**. Spoiwem łączącym wszystkie prace jest: wspólne źródło danych, jakim są dane skanowane pochodzące z sieci handlowych działających w Polsce¹, wspólna perspektywa metodologiczna (porównanie własności aksjomatycznych i empirycznych indeksów oraz indyktorów cen) oraz wspólna motywacja aplikacyjna (dostarczenie rekomendacji dla praktyki nowoczesnej statystyki cen).

Na prezentowany dorobek składa się sześć artykułów przygotowanych samodzielnie lub we współautorstwie, opublikowanych w recenzowanych i rozpoznawalnych czasopismach naukowych:

Artykuł 1: Pawelec, N. (2026). *The Use of Scanner Data for Price Indices Calculation: the Case of Poland*. Acta Universitatis Lodzensis. Folia Oeconomica, 1(374), 22-45. <https://doi.org/10.18778/0208-6018.374.02>

Artykuł 2: Białek, J., & Pawelec, N. (2025). *Proposition of the quality adjusted GEKS-type price index*. Statistika: Statistics and Economy Journal, 105(3), 291–305. <https://doi.org/10.54694/stat.2024.71>

Artykuł 3: Białek, J., Pawelec, N., & Roszkowska, S. (2024). *Estimating the elasticity of substitution when compiling the CES cost of living index on scanner data. Quality & Quantity*. <https://doi.org/10.1007/s11135-024-01924-8>

Artykuł 4: Białek, J., & Pawelec, N. (2024a). *The use of transitive Montgomery indicators for scanner data analysis*. Argumenta Oeconomica, 2(53), 5–26. <https://doi.org/10.15611/aoe.2024.2.01>

¹ Zbiory danych skanowanych wykorzystane w niniejszym cyklu publikacji pochodzą od sieci handlowych prowadzących działalność w Polsce. Tożsamość dostawców danych objęta jest zobowiązaniem do poufności, w związku z czym nazwy sieci handlowych nie są ujawniane, a dane zostały zanonimizowane na potrzeby przeprowadzonych analiz.

Artykuł 5: Pawelec, N. (2025). *Zastosowanie indyktorów Benneta i Montgomery'ego w analizie danych skanowanych – porównanie własności i oszacowań efektów*. *Wiadomości Statystyczne. The Polish Statistician*, 70(3), 11-23. <https://doi.org/10.59139/ws.2025.03.2>

Artykuł 6: Białek, J., & Pawelec, N. (2024b). *Comparison of transitive Bennet and Montgomery indicators on the basis of scanner data sets*. *European Research Studies Journal*, 27(4), 472–485. <https://doi.org/10.35808/ersj/3528>

Artykuły te tworzą logiczną i spójną całość, gdzie lista poruszanych zagadnień przebiega od diagnozy ograniczeń klasycznych indeksów bilateralnych (artykuł 1), przez propozycję nowego indeksu multilateralnego uwzględniającego korektę jakościową (artykuł 2) oraz analizę podejścia ekonomicznego opartego na modelu CES (artykuł 3), aż po konstrukcję i porównanie tranzytywnych indyktorów cen i ilości dedykowanych danym skanowanym (artykuły 4, 5 i 6²).

Każdy z artykułów podejmuje odmienny aspekt problematyki pomiaru inflacji z wykorzystaniem danych skanowanych, tworząc spójną całość badawczą ukierunkowaną na analizę własności indeksów i indyktorów cen stosowanych w procesach pomiaru inflacji. Jako całość, prace te pozwalają na pogłębioną ocenę możliwości wykorzystania danych skanowanych w analizie dynamiki cen oraz na identyfikację wyzwań metodologicznych związanych z ich zastosowaniem.

Wkład niniejszego cyklu prac do literatury przedmiotu jest trojaki: metodologiczny, empiryczny i aplikacyjny.

Wkład metodologiczny: cykl dostarcza dwóch oryginalnych propozycji metodycznych. Pierwszą jest nowy multilateralny indeks cen GEKS-GAQI, który jako pierwszy w swojej klasie łączy podejście GEKS z korektą jakościową produktów i jednocześnie spełnia restrykcyjny test identyczności. Drugą jest konstrukcja tranzytywnych multilateralnych indyktorów Montgomery'ego dedykowanych analizie

² Artykuł 6 ukazał się w *European Research Studies Journal*, który figuruje na listach czasopism budzących wątpliwości co do standardów recenzji (predatory journals). Artykuł powstał jako publikacja pokonferencyjna; Autorzy w momencie zgłoszenia nie dysponowali informacją o docelowym czasopiśmie. Autorzy są świadomi zaistniałej sytuacji niemniej ze względu na istotną wartość merytoryczną artykułu i jego rolę w budowaniu spójności cyklu pozostaje on włączony do niniejszej rozprawy.

danych skanowanych, co stanowi jedno z pierwszych tego rodzaju rozwiązanie w literaturze przedmiotu.

Wkład empiryczny: wszystkie analizy zostały przeprowadzone na rzeczywistych danych transakcyjnych z polskiego rynku detalicznego, obejmujących zróżnicowane grupy produktowe, co pozwala na weryfikację własności teoretycznych indeksów i indyktorów w realnych warunkach rynkowych z pełną rotacją produktów, zmiennością cen i efektami sezonowymi.

Wkład aplikacyjny (wdrożeniowy): uzyskane wyniki stanowią podstawę do dyskusji nad wyborem rozwiązań metodycznych dla urzędów statystycznych wdrażających dane skanowane do obliczeń CPI, w szczególności w zakresie wyboru formuły indeksowej, poziomu agregacji danych oraz progów stosowanych filtrów. Ma to bezpośrednie znaczenie praktyczne: ponad 40% krajów Unii Europejskiej wdrożyło lub wdraża dane skanowane w obliczeniach CPI, podczas gdy dostępne przewodniki metodyczne w tym Eurostat (2022) wciąż dyskutują o preferowanej formule indeksowej (Białek, Pawelec, 2025; Eurostat, 2022).

Przedmiot i zakres badań

Tradycyjne metody pozyskiwania informacji w ramach pomiaru inflacji oparte na bezpośrednich notowaniach ankierskich stają się niewystarczające w obliczu skali i zmienności współczesnego rynku. Przeobrażenia te, obejmujące segmentację handlu detalicznego, zmiany struktury demograficznej oraz wzorce zachowań nabywczych, mogą skutkować spadkiem reprezentatywności prób badawczych. W konsekwencji ogranicza to zdolność badacza do precyzyjnego opisywania procesów rynkowych oraz trafnego prognozowania zjawisk inflacyjnych. Z tego względu kluczowym postulatem współczesnej statystyki cen jest adaptacja metodologii do nowych uwarunkowań, determinowanych przez dostęp do alternatywnych źródeł danych oraz innowacyjnych technologii.

W tym kontekście decydującą rolę odgrywają technologie, narzędzia i zbiory *Big Data*, generowane w tempie wykładniczym przez nowoczesne systemy transakcyjne. Efektywna analiza ogromnych zbiorów danych w czasie rzeczywistym wykracza poza możliwości tradycyjnych narzędzi i technik statystycznych. Mimo że zbiory te zawierają kluczowe informacje z punktu widzenia różnych grup interesariuszy, ich implementacja,

choć niezbędna, wiąże się z licznymi wyzwaniami metodologicznymi i technologicznymi. Wymagają one precyzyjnego podejścia do zarządzania danymi masowymi w celu wygenerowania wiarygodnych wskaźników cenowych, stanowiących fundament dla nowoczesnych analiz inflacyjnych. Implementacja alternatywnych źródeł danych w pomiarze inflacji wymusza stosowanie innowacyjnych formuł indeksów cen, zdolnych do poprawnej agregacji informacji o wysokiej częstotliwości pozyskiwania i odpornych na olbrzymią rotację produktów przy jednoczesnym zachowaniu pożądanых własności aksjomatycznych.

Niniejsza rozprawa stanowi próbę odpowiedzi na te potrzeby, koncentrując się na rozwinięciu, porównaniu oraz ocenie własności nowoczesnych metod pomiaru inflacji (lub ogólniej: dynamiki cen), w tym multilateralnych indeksów cen typu GEKS, metod opartych na modelu stałej elastyczności substytucji (CES) oraz wybranych indykatorów cen i ilości, integrując tym samym podejścia aksjomatyczne oraz ekonomiczne w konstrukcji miar dynamiki cen. Omawiany cykl prac ogniskuje uwagę czytelnika na nowoczesnym źródle danych w pomiarze inflacji, jakim są *dane skanowane*. Cykl tych prac obejmuje swoim zasięgiem następujące obszary tematyczne:

- 1. Wyzwania, ograniczenia i korzyści związane ze stosowaniem indeksów cen w pomiarze inflacji bazującym na danych skanowanych.**

W ramach tego obszaru prace z cyklu dokonują:

- porównania bilateralnych i multilateralnych indeksów cen na bazie danych skanowanych;
- analizy wrażliwości indeksów cen na rotację produktów skanowanych oraz na zmienność ich cen i poziom konsumpcji.

- 2. Aksjomatyczna teoria indeksów cen.**

W ramach tego obszaru prace ujęte w zgłoszonym cyklu obejmują:

- propozycję nowego multilateralnego indeksu cen stanowiącego modyfikację indeksu GEKS, który uwzględnia korektę jakościową (dopasowanie jakościowe);

- analizę aksjomatycznych własności zaproponowanego indeksu GEKS-GAQI i jego porównanie z funkcjonującymi w literaturze przedmiotu indeksami multilateralnymi.

3. Podejście ekonomiczne w konstrukcji indeksu cenowego.

W ramach tego obszaru jedna z prac z cyklu dokonuje:

- estymacji elastyczności substytucji dla różnych zbiorów danych skanowanych, różnych poziomów agregacji tych danych oraz przy zastosowaniu różnych metod estymacji;
- porównania wpływu metody estymacji elastyczności substytucji na indeks CES (indeks Lloyda-Moultona).

4. Metodologia indykatorów cen i ilości.

W przypadku tego obszaru, prace z cyklu obejmują:

- dyskusję nad tranzytywnymi wersjami „klasycznych” indykatorów Benneta i Montgomery’ego;
- analizę własności tranzytywnych indykatorów cen i ilości oraz ich dekompozycję na efekt cenowy i ilościowy z wykorzystaniem danych skanowanych.

W dalszej części omówiono **główne obszary tematyczne** cyklu publikacji w sposób bardziej szczegółowy. Obszary te stanowią konceptualną ramę cyklu publikacji oraz punkt odniesienia dla sformułowania celów pracy i dalszych analiz.

1. Wyzwania, ograniczenia i korzyści związane ze stosowaniem indeksów cen w pomiarze inflacji bazującym na danych skanowanych

Dane skanowane stanowią dane transakcyjne pozyskiwane przede wszystkim od sieci handlowych, które są rejestrowane w punktach sprzedaży przez terminale handlowe, obejmując szczegółowe informacje o cenach, ilościach oraz cechach produktów, umożliwiając tym samym analizę zmian cen na poziomie pojedynczych transakcji (Eurostat 2017). Identyfikacja produktu odbywa się tu za pomocą kodu

kreskowego GTIN³ lub jego europejskiej wersji EAN⁴ (lub innych kodów kreskowych takich jak SKU⁵, PLU⁶, UPC⁷). Dane skanowane wykorzystywane przez krajowe urzędy statystyczne nie są specjalne tworzone i zbierane na potrzeby obliczeń wskaźników cen konsumpcyjnych, lecz stanowią one źródło danych wtórnych. Zbiory mogą być pozyskiwane od szerokiego zakresu detalistów tj.: aptek, supermarketów, sklepów z elektroniką i wielu innych. Cecha ta stanowi istotną przewagę nad tradycyjnym pomiarem cen, ponieważ umożliwia pozyskiwanie informacji o rzeczywistych cenach transakcyjnych w dłuższym okresie, a nie jedynie w jednym dniu miesiąca, jak ma to miejsce w przypadku klasycznego poboru cen. Dodatkowo, detaliczność danych skanowanych pozwala na wykluczenie produktów, które nie zostały sprzedane (lub których sprzedaż jest relatywnie niska) lub ograniczenie wpływu ekstremalnych zmian cen powstałych na skutek promocji, rabatów czy ofert specjalnych. To wszystko odbywa się przy znacznym ograniczeniu kosztów badania, pracochłonności i czasochłonności procesu pozyskiwania danych.

Wskazane zalety danych skanowanych wynikają z ich szczegółowości oraz transakcyjnego charakteru. Jednocześnie te same cechy wiążą się również z wyzwaniami metodologicznymi. Ocena dynamiki cen oparta na danych skanowanych w istotnym stopniu zależy od zastosowanej formuły indeksowej. W tym kontekście szczególnego znaczenia nabiera porównanie indeksów bilateralnych i multilateralnych, ponieważ każda z tych rodzin indeksów w odmienny sposób wykorzystuje informacje zawarte w zbiorach danych oraz różni się własnościami aksjomatycznymi i statystycznymi. W przypadku tradycyjnego zbierania danych (np. poprzez ankieterów w terenie i wsparcie Badania Budżetów Gospodarstw Domowych) wybór formuły indeksu cenowego jest w dużej mierze determinowany dostępnością informacji i opiera się na indeksach bilateralnych, w których ceny ustalonego (domyślnie niezmiennego) koszyka dóbr i usług w jednym lub kilku okresach są bezpośrednio porównywane z analogicznymi cenami w wybranym okresie bazowym. W zależności od formuły indeksu, wykorzystywane ilości nabywanych dóbr lub wydatki mogą pochodzić z okresu bazowego (np. indeks Laspeyresa), okresu bieżącego (np. indeks Paaschego),

³ Global Trade Item Number

⁴ European Article Number

⁵ Stock Keeping Unit

⁶ Price Look-Up

⁷ Universal Product Code

zarówno z okresu bazowego, jak i bieżącego (np. indeks Fishera, Walsha, Törnqvista), lub z dowolnego innego okresu wcześniejszego (np. indeks Lowe'a, czy Younga) (ILO, 2020).

W przypadku statycznego zbioru dóbr zastosowanie bilateralnej formuły indeksu stanowi prosty sposób pomiaru zagregowanej i wypadkowej zmiany cen koszyka inflacyjnego. Dane skanowane charakteryzują się jednak znaczną rotacją produktów. Zjawisko to obejmuje występowanie zarówno dóbr znikających jak i kompletnie nowych w analizowanym okresie i jest nazywane w literaturze przedmiotu jako *product churn*. Rotacja sprzedawanych produktów związana jest zarówno z silną sezonowością asortymentu jak i zmieniającymi się preferencjami konsumentów. Oznacza to, że mamy do czynienia z dynamicznym zbiorem produktów, który wymaga szczególnej uwagi statystyka, a zastosowanie wyłącznie indeksów bilateralnych mogłoby prowadzić do zbyt dużego obciążenia pomiaru dynamiki cen. Wynika to z faktu, iż indeksy bilateralne biorą pod uwagę tylko okres bieżący oraz bazowy, czyli nie są w stanie uchwycić wspomnianej wcześniej dużej rotacji produktów skanowanych i wszystkich tych zmian w sprzedawanym asortymencie, jakie zachodzą pomiędzy tymi okresami.

W przypadku dynamicznych zbiorów produktów skanowanych, gdzie struktura koszyka ulega zmianie niemal w każdym miesiącu, bezpośrednie porównywanie dwóch okresów oddalonych w czasie może przyczynić się do ignorowania istotnych procesów substytucji zachodzących w miesiącach pośrednich. W przypadku danych skanowanych, nawet na najniższym poziomie agregacji, czyli na poziomie kodu kreskowego, dostępne są informacje o cenie produktu oraz wielkości jego sprzedaży. Umożliwia to konstrukcję ważonych indeksów cenowych opartych na dowolnym okresie z analizowanego przedziału czasowego, a tym samym zastosowanie nie tylko bilateralnych formuł indeksowych, lecz także indeksów łańcuchowych i multilateralnych, które lepiej odpowiadają specyfice danych o dużej zmienności i rotacji produktów.

Wszystko to prowadzi do wniosku, że wybór odpowiedniej formuły indeksu w analizie danych skanowanych pozostaje kwestią otwartą i wciąż nierozstrzygniętą w literaturze, co wynika z kompromisu pomiędzy własnościami teoretycznymi indeksów, a ich zachowaniem empirycznym w warunkach wysokiej rotacji produktów i

zmienności cen i poziomu konsumpcji (Jan de Haan i in., 2016; Eurostat, 2022; Marie Leclair et in., 2020).

2. Aksjomatyczna teoria indeksów cen

Wobec braku jednoznacznych rozstrzygnięć dotyczących wyboru optymalnej formuły indeksu, w literaturze przedmiotu rozwinięto różne podejścia metodologiczne do konstrukcji indeksów cenowych. Jednym z podstawowych nurtów w teorii indeksów cen jest podejście aksjomatyczne (*axiomatic approach*), które koncentruje się na analizie własności formalnych indeksów oraz ich zgodności z określonym zestawem. W tym ujęciu wybór indeksu nie wynika z modelu ekonomicznego, lecz z oceny stopnia spełniania pożądaných własności teoretycznych (aksjomatów, tzw. *testów*), co stanowi szczególnie istotny punkt odniesienia w analizie danych skanowanych (W. Erwin Diewert, 1976; Ralph Turvey, 1989; Eurostat, 2022). Wymaga się, aby indeks posiadał określone własności lub spełniał określony zestaw testów (tzw. *systemy minimalnych wymagań*). Wybór indeksu może być dokonany na podstawie liczby spełnionych przez niego testów. Nie wszystkie testy są jednakowo istotne, a niespełnienie jednego lub dwóch kluczowych testów może być wystarczającym powodem, aby odrzucić daną formułę indeksu z rozważań (ILO 2020). Zestaw aksjomatów (testów) jakich wymaga się od poprawnie skonstruowanego indeksu multilateralnego można znaleźć w pracach Eurostat (2018), czy ABS (2016), natomiast aksjomatyka dla indeksów bilateralnych jest drobiazgowo omówiona w pracy von der Lippe (2007).

W kontekście danych skanowanych szczególne znaczenie przypisuje się własności tranzytywności (*transitivity*) oraz testowi identyczności (*identity test*). *Tranzytywność* oznacza spójność porównań międzyokresowych w całym analizowanym oknie czasowym, tj. możliwość przedstawienia zmiany cen pomiędzy dwoma okresami jako iloczynu zmian w podokresach, co w praktyce eliminuje problem łańcuchowego dryfu (*chain drift*). Należy zaznaczyć, iż indeksy multilateralne (działające na oknie czasowym) są z definicji tranzytywne. Z kolei *test identyczności* wymaga, aby indeks przyjmował wartość równą jeden w sytuacji, gdy ceny produktów w okresie bieżącym powracają do poziomu z okresu bazowego, co stanowi podstawowy warunek poprawnej interpretacji zmian cen. Jak wskazuje literatura spełnienie obu tych własności jednocześnie stanowi istotne wyzwanie w konstrukcji indeksów multilateralnych, gdyż wiele powszechnie stosowanych metod nie spełnia testu identyczności pomimo

zachowania tranzytywności (von der Lippe, 2007; Eurostat, 2018; Białek, 2023; Białek & Pawelec, 2025). Takimi indeksami są popularne indeksy GEKS, TPD czy indeks Geary-Khamisa.

3. Podejście ekonomiczne w konstrukcji indeksu cenowego

Podejście ekonomiczne do konstrukcji indeksów cenowych opiera się na koncepcji indeksu kosztów utrzymania (*Cost of Living Index, COLI*), który mierzy zmiany minimalnych wydatków niezbędnych do osiągnięcia danego poziomu użyteczności koszyka dóbr i usług przy zmieniających się cenach jego komponentów. W tym ujęciu zakłada się, że racjonalny konsument maksymalizuje użyteczność lub minimalizuje wydatki przy danym ograniczeniu budżetowym, a preferencje konsumenta są reprezentowane przez pewną funkcję użyteczności (Diewert, 1976; Diewert, 2002). Indeks cenowy wynika z relacji funkcji kosztów w porównywanych okresach i odzwierciedla zmiany kosztu utrzymania stałego poziomu dobrobytu, co odróżnia go od tradycyjnych indeksów typu Laspeyresa, które nie uwzględniają możliwości substytucji między dobrami (Triplett, 2001).

W praktyce bezpośrednia konstrukcja „prawdziwego” COLI jest utrudniona (choćby z powodu braku wiedzy o postaci funkcji użyteczności konsumentów), dlatego stosuje się jego przybliżenia. Za najlepsze przybliżenie COLI uznaje się w szczególności indeksy *superlatywne* (Diewert, 1976) oraz indeksy oparte na modelu CES (*Constant Elasticity of Substitution*). Podejście CES umożliwia jawne uwzględnienie efektu substytucji poprzez parametr elastyczności substytucji, który określa reakcję struktury konsumpcji na zmiany relatywnych cen. Wartość tego parametru ma kluczowe znaczenie dla wyników indeksu, a jego poprawna estymacja stanowi jedno z głównych wyzwań empirycznych (Feenstra, 1994).

Zastosowanie podejścia ekonomicznego jest zasadne w przypadku danych skanowanych, które dostarczają szczegółowych informacji o cenach i ilościach na poziomie pojedynczych produktów, umożliwiając estymację modeli popytu oraz parametrów takich jak elastyczność substytucji. Nie jest to jednak takie oczywiste, ponieważ koncepcja COLI jest koncepcją bilateralną, a w przypadku danych skanowanych urzędy statystyczne coraz częściej sięgają po indeksy multilateralne. Jednocześnie duża rotacja produktów oraz coraz częstsza zmiana rozmiaru produktów (*downsizing, upsizing*) stanowią istotne wyzwanie dla pomiaru zmian cen i wymagają

uwzględnienia efektów jakościowych oraz zmian w strukturze konsumpcji (Feenstra, 1994; de Haan i in., 2016). Z punktu widzenia podejścia ekonomicznego, dynamiczny charakter danych skanowanych może budzić wątpliwości co do zasadności założeń modelu CES (stała w czasie wartość w elastyczności substytucji).

4. Metodologia indykatorów cen i ilości

Teoria indykatorów⁸ opiera się na porównywaniu w dwóch okresach: bieżącym i bazowym, cen, ilości i wydatków na towary i usługi, ale przy założeniu, iż zmiana poziomu wydatków daje się dekomponować na efekt cenowy i ilościowy. Metodologia ta stanowi alternatywne podejście do klasycznej teorii indeksów. Indeksy cenowe opierają się na ilorazach obserwowanych zmiennych, podczas gdy indykatory bazują na różnicach wartości tych zmiennych (cen, ilości) eliminując tym samym techniczny problem dzielenia przez zero. Jest to kluczowa zaleta indykatorów zwłaszcza w przypadku analizy danych skanowanych, gdzie wysoka rotacja produktów i występowanie zerowych cen i/lub ilości generuje techniczny problem "tradycyjnym" indeksom (dzielenie przez zero jest niewykonalne).

Warto mieć na uwadze, że indykatory nie zastępują indeksów, lecz je uzupełniają, dostarczając dodatkowych informacji o mechanizmie obserwowanej dynamiki sprzedaży oraz o roli zmian struktury konsumpcji produktów i zmian cen produktów w kształtowaniu się tej dynamiki. Indykatory umożliwiają analizę wkładu komponentów cenowych i ilościowych w zmianę wartości całkowitej sprzedaży, co stanowi punkt wyjścia do dalszych rozważań nad ich własnościami formalnymi. W tym kontekście istotne znaczenie zyskuje problem tranzytywności (*transitivity*), który jest szczególnie pożądaną własnością w analizie porównań przestrzennych oraz dynamicznych szeregów czasowych. W przypadku indykatorów, które operują na bezwzględnych różnicach (a nie ilorazach), tranzytywność przyjmuje postać addytywną (Białek, Pawelec 2024b). Tranzytywność gwarantuje, że wyniki pomiaru dynamiki cen i ilości na oknie czasowym nie są zależne od wyboru okresu bazowego, a co najważniejsze eliminuje ona problem łańcuchowego dryfu (*chain drift*). Klasyczne bilateralne indykatory Benneta oraz Montgomery'ego nie spełniają warunku tranzytywności (Białek, Pawelec 2024b), podobnie jak klasyczne indeksy Laspeyresa i

⁸ Indykator (indicator) nie jest synonimem wskaźnika/indeksu (index). W literaturze występuje wyraźne rozróżnienie tych pojęć (Diewert 2005).

Paaschego również nie są tranzytywne. Uwaga ta prowadzi do rozwoju tranzytywnych wersji klasycznych indykatów Benneta i Montgomery'ego, które pozwalają zachować addytywną dekompozycję zmian przy jednoczesnym zapewnieniu spójności pomiaru w całym oknie czasowym (Fox, 2006; Białek & Pawelec, 2024a).

Cele, problemy i hipotezy badawcze

W publikacjach przedłożonych w ramach rozprawy doktorskiej dokonano analizy kluczowych problemów związanych z wykorzystaniem danych skanowanych w statystyce cen, w szczególności dotyczących wpływu dużej rotacji produktów o wysokiej zmienności cen oraz ilości na własności indeksów cen i indykatów, wyboru pomiędzy podejściem bilateralnym i multilateralnym, a także roli poziomej agregacji danych i procedur filtrowania danych w kształtowaniu wyników pomiaru inflacji. W pracach tych przedstawiono zarówno analizę własności stosowanych metod indeksowych, jak i propozycje pewnych rozwiązań metodologicznych dostosowanych do specyfiki danych skanowanych (np. rekomendację dla estymacji elastyczności substytucji stosowanej w indeksach typu CES, czy propozycję indeksu multilateralnego typu GEKS stosującego dopasowanie jakościowe).

Analizy przedstawionych problemów badawczych dokonano z wykorzystaniem danych skanowanych pochodzących z sieci handlowych funkcjonujących w Polsce, wykorzystując odpowiednie metody przygotowania tych danych (czyszczenie, filtrowanie, dopasowanie w czasie, klasyfikowanie do odpowiednich kategorii) i uwzględniając powszechnie stosowane w tym zakresie formuły indeksów cen. Każdy ze zbiorów zawiera informacje o cenach produktów, ilościach ich sprzedaży, dacie zakupu (transakcji), lokalizacji punktu sprzedaży sieci (outletu), etykiecie produktu oraz identyfikatorach produktu (kody stosowane przez sieć i kody kreskowe, jak EAN, GTIN, SKU i inne). Zbiory te dotyczą różnych grup produktów (rozważono grupy z kategorii spożywczych i niespożywczych), co umożliwia analizę zróżnicowanych wzorców zmian cen oraz porównanie wartości indeksów cen w odmiennych warunkach rynkowych.

Poszczególne artykuły cyklu koncentrują się na różnych obszarach metodologicznych (patrz obszary 1-4 na str. 7), ale łączy je jeden, główny cel badawczy, jakim jest **ocena i rozwój metod pomiaru dynamiki cen produktów skanowanych**,

ze szczególnym uwzględnieniem własności multilateralnych wersji indeksów i indyktorów. W jego ramach wyróżniono pięć celów szczegółowych.

Cel 1: Ocena możliwości wykorzystania danych skanowanych w pomiarze inflacji oraz identyfikacja ograniczeń bilateralnych indeksów cen w warunkach wysokiej dynamiki struktury koszyka inflacyjnego.

Punktem wyjścia prowadzonych analiz była rosnąca dostępność elektronicznych danych transakcyjnych (danych skanowanych) pochodzących z systemów sprzedażowych sieci handlowych oraz ich potencjalne znaczenie dla rozwoju statystyki cen. Dane skanowane zyskują na znaczeniu jako lepsza alternatywa dla tradycyjnego pomiaru inflacji otwierając nowe możliwości analizy procesów cenowych. (Marie Leclair i in. (2020)). W związku z tym przeprowadzono ocenę możliwości wykorzystania danych skanowanych w pomiarze inflacji, a następnie dokonano identyfikacji ograniczeń bilateralnych indeksów cen w warunkach wysokiej dynamiki struktury koszyka inflacyjnego (Pawelec, 2026). Analiza miała na celu weryfikację czy indeksy multilateralne są lepiej przystosowane do szacowania inflacji na danych skanowanych niż indeksy bilateralne. Z tak sformułowanym celem związane są dwie hipotezy badawcze:

Hipoteza badawcza 1A: Indeksy multilateralne są lepiej przystosowane do szacowania inflacji na danych skanowanych, ponieważ są mniej wrażliwe na dużą rotację produktów skanowanych oraz wysoką zmienność ich cen i sprzedawanych ilości aniżeli indeksy bilateralne.

Hipoteza badawcza 1B: Skala wrażliwości obu grup indeksów zależy od poziomu agregacji danych skanowanych i rodzaju zastosowanego filtru danych. Im niższy jest poziom agregacji danych, tym większe pojawiają się różnice we wskazaniach indeksów cen. Na wartości indeksów cen największy wpływ ma filtr niskich sprzedaży.

Cel 2: Konstrukcja multilateralnego indeksu typu GEKS wykorzystującego dopasowanie jakościowe produktów, który w odróżnieniu od klasycznych indeksów multilateralnych spełnia test identyczności (*identity*).

Kolejnym z pytań postawionych w ramach cyklu publikacji jest pytanie badawcze dotyczące wyboru indeksu multilateralnego, który z jednej strony powinien spełniać odpowiednie wymogi formalne (spełniać tzw. *testy* dla indeksów

multilateralnych), a z drugiej strony powinien generować jak najmniej obciążone szacunki inflacji (Eurostat 2022). Wyzwanie to jest szczególnie istotne, gdyż wybór formuły indeksu dedykowanego danym skanowanym nie jest jednoznaczny. Część krajów stosuje tzw. podejście dynamiczne (ang. *dynamic approach*) z łańcuchowym indeksem Jevonsa jako naczelną formułą obliczeniową, część krajów z kolei eksperymentuje z indeksami multilateralnymi (z reguły decydując się na użycie indeksów typu GEKS, indeksu Geary-Khamisa czy indeksu TPD lub TDH). Oczywiście dostępne są pewne wskazówki czy sugestie, jakie formuły w ogóle powinny być brane pod uwagę (Eurostat, 2022), ale nadal w przewodnikach po tych metodach, czy manualach do CPI (2020) i HICP (2018) brakuje wyraźnego wskazania ostatecznej formuły indeksu cen dedykowanego danym skanowanym. Rosnąca dostępność i wykorzystanie danych skanowanych w statystyce cen przyczynia się bezpośrednio do rozwoju nowych metod konstrukcji indeksów cen, które będą dostosowane do specyfiki tych danych. Z uwagi na pewną niszę w metodologii indeksów multilateralnych, w jednej z prac cyklu skoncentrowano się na konstrukcji nowego multilateralnego indeksu typu GEKS, który z jednej strony stosuje dopasowanie jakościowe produktów (jest zatem bliski idei indeksu Geary-Khamisa), ale z drugiej strony, w odróżnieniu od dotychczas stosowanych indeksów multilateralnych, spełnia test identyczności (ang. *identity test*). Zaproponowany indeks (GEKS-GAQI) dużo czerpie także z idei indeksów klasy GEKS, ponieważ stosuje geometryczne uśrednienie indeksów częściowych pozyskiwanych na zadanym oknie czasowym (np. trzynastomiesięcznym). Reasumując, analiza miała na celu sprawdzenie hipotezy:

Hipoteza badawcza 2: Możliwa jest konstrukcja multilateralnego indeksu GEKS, który: (a) stanowiłby pomost pomiędzy metodyką indeksu GEKS (stanowiącego geometryczne uśrednienie indeksów bilateralnych), a metodyką indeksu Geary-Khamisa (stosującego dopasowanie jakościowe produktów i współczynniki korygujące); (b) spełniałby restrykcyjny test identyczności (*identity test*).

Cel 3: Analiza modelowego podejścia do indeksów kosztów utrzymania opartego na funkcji CES (*Constant Elasticity of Substitution*), ze szczególnym uwzględnieniem możliwości estymacji elastyczności substytucji na podstawie danych skanowanych.

Indeksy typu GEKS, jak i inne klasy indeksów, można oceniać przez pryzmat spełnianych przez nie aksjomatów (tzw. testów) i jest to najpopularniejsze podejście w metodologii indeksów cen (ang. *axiomatic approach*). Równolegle w literaturze przedmiotu rozwijane są inne podejścia służące selekcji pożądanych formuł indeksów cen (np. podejście ekonomiczne (*economic approach*), stochastyczne (*stochastic approach*) czy łańcuchowe (*chain index approach*). W szczególności, podejście ekonomiczne oparte jest na teorii kosztów utrzymania i wykorzystuje model stałej elastyczności substytucji CES (*Constant Elasticity of Substitution model*) do konstrukcji tzw. indeksów superlatywnych (ang. *superlative indices*). Przeprowadzenie badań służących ostatecznej rekomendacji indeksu cenowego (lub indykatora cenowego – w zależności ode celu badania) może mieć wpływ nie tylko na gospodarkę, ale może mieć znaczenie dla inwestorów czy osób zarządzającymi bankami i instytucjami finansowymi. W związku z tym kolejnym celem była analiza modelowego podejścia do indeksów kosztów utrzymania opartego na funkcji CES, ze szczególnym uwzględnieniem możliwości estymacji elastyczności substytucji na podstawie danych skanowanych. W ramach badania podjęto próbę oceny, w jakim stopniu zastosowane rozwiązania metodologiczne (tzw. metody algebraiczne i ekonometryczne) pozwalają na wiarygodne wyznaczenie parametru elastyczności substytucji oraz jak wyniki te zmieniają się pod wpływem doboru filtru danych i poziomu agregacji danych. Podjęty w cyklu prac wysiłek miał na celu wskazanie czy dane skanowane mogą stanowić dobre źródło danych do szacowania indeksów typu CES (indeks Lloyda-Moultona) i aproksymacji indeksów kosztu utrzymania (COLI – Cost of Living Index) w warunkach dużej zmienności cen i struktury sprzedaży. W ramach badań z tego zakresu sformułowano dwie hipotezy badawcze:

Hipoteza badawcza 3A: Algebraiczne metody estymacji elastyczności substytucji (występującej w indeksie CES) są znacznie mniej wrażliwe na wybór filtru danych skanowanych i poziom ich agregacji aniżeli metody oparte na regresji panelowej;

Hipoteza badawcza 3B: Im wyższy jest poziom agregacji danych, tym niższa jest wartość elastyczności substytucji

Cel 4: Konstrukcja multilateralnego indykatora Montgomery’ego dedykowanego analizie danych skanowanych i zbadanie jego aksjomatycznych własności.

Dotychczasowe rozważania koncentrowały się na indeksach cenowych, zarówno w ujęciu aksjomatycznym (np. GEKS) jak i ekonomicznym (podejście oparte na funkcji CES). Cennym uzupełnieniem analiz dotyczących indeksów cen jest – w mniemaniu autorki niniejszej pracy - dodatkowa analiza indyktorów cen i ilości. Z praktycznego punktu widzenia cenną własnością jest tu fakt, iż suma indykatora cenowego i ilościowego wyznacza różnicę pomiędzy całkowitą wartością sprzedaży w okresie bieżącym i całkowitą wartością sprzedaży w okresie bazowym. A zatem, dzięki indyktorom, przychód ze sprzedaży można zdekomponować na efekt cenowy i ilościowy, co w przypadku indeksów cen wiąże się tylko z nielicznymi formułami (np. indeks Fishera) i oznacza spełnienie testu odwróconych czynników (ang. *factor reversal test*). Należy zaznaczyć, iż teoria indyktorów cen i ilości jest znacznie mniej rozbudowana i rozpoznana, aniżeli teoria indeksów cen i ilości. Literaturową niszą w tym obszarze jest z pewnością brak tranzytywnych indyktorów cen i ilości operujących na przedziale czasowym (choć są dostępne tranzytywne wersje indyktorów do porównań międzyregionalnych operujących na ustalonym zbiorze obiektów – państw, miast, regionów, etc.). Tranzytywne indykatory skonstruowane dla zadanego przedziału czasowego w założeniu wyeliminowałyby zjawisko tzw. łańcuchowego dryfu (ang. *chain drift*), jakie obserwujemy w analizie danych skanowanych. Wyeliminowanie obciążenia pomiaru wynikającego z efektu łańcuchowego dryfu z technicznego punktu widzenia po prostu oznacza, iż w grupie produktów sezonowych, gdy ceny i ilości sprzedaży powracają po sezonie do swoich wyjściowych poziomów, indykatory cen przyjmują zerowe wartości, a indeksy cen mają wartość jeden. I tak powinno być, jeśli wymagamy nieobciążonego pomiaru.

Mając powyższe na uwadze, podjęto próbę konstrukcji multilateralnych indyktorów cen i ilości dedykowanych analizie danych skanowanych. Motywacją do takich analiz był brak prac związanych z wykorzystaniem indyktorów (zarówno bilateralnych, jak i multilateralnych) do analizy danych skanowanych. W szczególności skoncentrowano się na indyktorze Montgomery’ego, któremu w literaturze poświęca się relatywnie mniej uwagi w porównaniu z indykatorem Benneta (Białek i Pawelec 2024a). Badanie empiryczne w ramach tego celu zostało ukierunkowane przez dwie hipotezy:

Hipoteza badawcza 4A: Możliwa jest konstrukcja multilateralnego indykatora Montgomery’ego, który działając na oknie czasowym spełni trzy ważne aksjomatyczne

wymogi: test zmiany wartości (*value change test*), test odwróconego czasu (*time reversal test*) oraz wielowymiarowy test identyczności (*multi-period identity test*).

Hipoteza badawcza 4B: Powyższa konstrukcja jest możliwa nawet wtedy, gdy w zbiorze danych występują niedopasowane produkty i tym samym braki danych i/lub zerowe ceny.

Cel 5: Porównanie własności multilateralnych indykatorów Benneta i Montgomery’ego z wykorzystaniem danych skanowanych ze szczególnym uwzględnieniem wrażliwości tych indykatorów na rodzaj zastosowanego filtru danych oraz poziom agregacji danych.

Piątym i ostatnim poruszonym problemem badawczym było porównanie własności multilateralnych indykatorów Benneta i Montgomery’ego z wykorzystaniem danych skanowanych ze szczególnym uwzględnieniem wrażliwości tych indykatorów na rodzaj zastosowanego filtru danych oraz poziom agregacji danych (Białek i Pawelec 2024b, Pawelec 2025). Przeprowadzona analiza stanowi naturalne rozwinięcie przeprowadzonych wcześniej przez autorkę badań, będąc analogiem do analizy porównawczej indeksów bilateralnych i multilateralnych, w których wykazano istotną wrażliwość na specyfikacje danych skanowanych. Włączenie do analiz indykatora Benneta oraz Montgomery’ego pozwoliło na rozszerzenie perspektywy badawczej o podejście oparte na dekompozycji zmian wartości sprzedaży produktów skanowanych. W efekcie otrzymano szczegółową analizę mechanizmów kształtujących dynamikę cen i ilości w analizowanych grupach produktów, otrzymując porównanie wyników dekompozycji zmian wartości sprzedaży uzyskiwanych różnymi metodami. Pozwoliło to na ocenę przydatności indykatorów cen i ilości w praktyce pomiaru dynamiki cen (i przy okazji ilości) opartego na danych skanowanych. Realizacji tego celu towarzyszyły następujące hipotezy badawcze:

Hipoteza badawcza 5A: Różnice pomiędzy tranzytywnymi wersjami indykatorów Benneta i Montgomery’ego są tym mniejsze, im wyższy jest poziom agregacji danych.

Hipoteza badawcza 5B: Zastosowanie filtrów danych bardzo wyraźnie wpływa na cenowe i ilościowe indykatory Benneta i Montgomery’ego i ich implementacja może prowadzić nawet do zmiany znaku tych indykatorów.

Tak sformułowane cele i hipotezy wyznaczają oś badawczą niniejszego cyklu publikacji. Wyniki przeprowadzonych analiz empirycznych, ocena stopnia realizacji poszczególnych celów oraz weryfikacja hipotez badawczych zostały przedstawione w rozdziale *Synteza wyników oraz wkład do literatury*.

Synteza wyników oraz wkład do literatury

Przedstawiony cykl publikacji koncentruje się na problematyce wykorzystania danych skanowanych w pomiarze inflacji oraz analizie własności indeksów cen i indyktorów cen i ilości. Przeprowadzonym badaniom przyświecało pięć powiązanych ze sobą celów badawczych, dotyczących zarówno oceny istniejących metod indeksowych i indyktorowych, jak i rozwoju nowych rozwiązań metodycznych. Poniżej przedstawiono syntetyczne omówienie celów badań, uzyskanych wyników oraz wkładu do literatury.

Pierwszym z przedstawionych powyżej celów badań (**Cel 1**) była ocena możliwości wykorzystania danych skanowanych w pomiarze inflacji oraz identyfikacja ograniczeń bilateralnych indeksów cen w warunkach wysokiej dynamiki struktury koszyka inflacyjnego, przy uwzględnieniu różnic pomiędzy podejściem bilateralnym i multilateralnym oraz wpływu poziomu agregacji i procedur filtrowania danych. Realizacja tak zdefiniowanego celu oraz empiryczna weryfikacja przyjętych założeń została przeprowadzona w artykule *The Use of Scanner Data for Price Indices Calculation: the Case of Poland* (artykuł nr 1), w którym wykorzystano dane skanowane do analizy własności indeksów cen.

Uzyskane wyniki jednoznacznie wskazują, że indeksy bilateralne – w tym indeksy łańcuchowe – wykazują istotne ograniczenia w warunkach wysokiej rotacji produktów oraz dużej zmienności cen i wolumenów sprzedaży, co prowadzi do niestabilności oraz ograniczonej porównywalności uzyskiwanych miar inflacji. Jednocześnie wykazano, że indeksy multilateralne są lepiej dostosowane do specyfiki danych skanowanych, ponieważ w większym stopniu uwzględniają zmiany struktury koszyka i ograniczają wpływ krótkookresowych fluktuacji na poziomie pojedynczych produktów (**potwierdzenie hipotezy 1A**). Ponadto potwierdzono, że wyniki pomiaru inflacji są istotnie uzależnione od poziomu agregacji danych oraz zastosowanych procedur filtrowania – wraz ze spadkiem poziomu agregacji rosną rozbieżności

między indeksami, a szczególnie silny wpływ na wartości indeksów wywiera filtr niskich sprzedaży (**potwierdzenie hipotezy 1B**).

Wkład do literatury polega na empirycznym wykazaniu ograniczeń klasycznych indeksów bilateralnych w warunkach wykorzystania danych skanowanych oraz na uzasadnieniu zasadności stosowania indeksów multilateralnych jako bardziej adekwatnego narzędzia pomiaru inflacji w warunkach wysokiej dynamiki strukturalnej rynku. Dodatkowo wskazano i usystematyzowano znaczenie poziomu agregacji danych oraz procedur filtrowania jako kluczowych czynników wpływających na własności i interpretację uzyskiwanych miar inflacji.

W ramach drugiego celu badawczego (**Cel 2**), polegającego na konstrukcji multilateralnego indeksu typu GEKS wykorzystującego dopasowanie jakościowe produktów oraz spełniającego test identyczności, podjęto próbę integracji podejścia opartego na geometrycznym uśrednianiu indeksów bilateralnych z metodą uwzględniającą korekty jakościowej. Realizacja tego celu została przedstawiona w artykule *Proposition of the Quality Adjusted Geks-Type Price Index* (artykuł numer 2), w którym zaproponowano nowy indeks multilateralny GEKS-GAQI, stanowiący rozwinięcie klasycznej konstrukcji GEKS poprzez wprowadzenie elementów charakterystycznych dla podejścia Geary'ego-Khamisa.

Uzyskane wyniki potwierdziły, że możliwa jest konstrukcja indeksu łączącego zalety obu podejść. W szczególności wykazano, że zaproponowany indeks spełnia restrykcyjny test identyczności, co stanowi istotne rozszerzenie własności aksjomatycznych względem standardowych indeksów multilateralnych. Jednocześnie indeks GEKS-GAQI umożliwia uwzględnienie dopasowania jakościowego produktów, dzięki czemu jest odporny na zmiany jednostki sprzedaży oraz inne źródła nieporównywalności obserwacji charakterystyczne dla danych skanowanych. (**potwierdzenie hipotezy 2**). Oznacza to, że konstrukcja indeksu pozwala na bardziej spójne i teoretycznie uzasadnione ujęcie zmian cen w warunkach wysokiej heterogeniczności produktów.

Wkład do literatury polega na zaproponowaniu nowej formuły indeksu multilateralnego, który stanowi pomost pomiędzy metodologią GEKS a podejściem Geary'ego-Khamisa, łącząc ich kluczowe własności w jednej spójnej konstrukcji. W szczególności wykazano, że możliwe jest jednoczesne spełnienie silnych postulatów

aksjomatycznych (w tym testu identyczności) oraz uwzględnienie korekt jakościowych, co dotychczas stanowiło istotne ograniczenie istniejących rozwiązań. Tym samym zaproponowany indeks rozszerza możliwości analityczne w zakresie pomiaru inflacji na danych skanowanych i stanowi podstawę do dalszych badań nad konstrukcją indeksów cenowych dostosowanych do nowoczesnych źródeł danych.

W ramach trzeciego celu (**Cel 3**) badawczego, polegającego na analizie modelowego podejścia do indeksów kosztów utrzymania opartego na funkcji CES (*Constant Elasticity of Substitution*), ze szczególnym uwzględnieniem możliwości estymacji elastyczności substytucji na podstawie danych skanowanych, podjęto badanie wpływu zarówno poziomu agregacji danych, jak i wyboru metody estymacji na uzyskiwane wartości parametrów modelu regresji panelowej. Realizacja tego celu została przedstawiona w artykule *Estimating the elasticity of substitution when compiling the CES cost of living index on scanner data* (artykuł numer 3) poświęconym zastosowaniu funkcji CES w kontekście danych skanowanych oraz konstrukcji indeksów kosztów utrzymania.

Uzyskane wyniki potwierdziły przyjęte hipotezy badawcze. Metody algebraiczne estymacji elastyczności substytucji charakteryzują się większą stabilnością i znacznie mniejszą wrażliwością na wybór filtru danych oraz poziom agregacji niż podejścia oparte na regresji panelowej w szczególności metoda M-LM⁹ dostarcza stabilnych oszacowań niezależnie od zastosowanego filtru danych (**potwierdzenie hipotezy 3A**). Jednocześnie poziom agregacji wyraźnie wpływa na wyniki estymacji: przejście od danych zdezagregowanych (poziom kodu kreskowego GTIN¹⁰) do wyższych poziomów agregacji (poziom COICOP 6¹¹) prowadziło każdorazowo do kilkukrotnego spadku oszacowanej elastyczności substytucji, przy czym na poziomie GTIN różnice między metodami estymacji były bardziej widoczne i przekładały się na zauważalne różnice wartości indeksów CES (**potwierdzenie hipotezy 3B**). Oznacza to, że zarówno dobór metody estymacji, jak i sposób agregacji danych mają kluczowe znaczenie dla wiarygodności i porównywalności uzyskiwanych parametrów.

Wkład do literatury polega na uzupełnieniu nurtu badań nad indeksami kosztów utrzymania poprzez systematyczne wykazanie, że estymacja elastyczności substytucji

⁹ Method based on the Lloyd-Moulton index

¹⁰ Global Trade Item Number

¹¹ Classification of Individual Consumption According to Purpose

w modelu CES jest silnie zależna od decyzji metodycznych, w tym w szczególności od poziomu agregacji danych oraz wyboru metody estymacji. W szczególności wskazano przewagę metod algebraicznych nad podejściem ekonometrycznym w zakresie stabilności oszacowań, co stanowi istotne rozszerzenie dotychczasowych badań empirycznych opartych głównie na regresji panelowej i ma bezpośrednie implikacje dla praktyki konstrukcji indeksów kosztów utrzymania typu CES.

W ramach czwartego celu badawczego (**Cel 4**), polegającego na konstrukcji multilateralnego indykatora Montgomery’ego dedykowanego analizie danych skanowanych oraz zbadaniu jego własności aksjomatycznych, podjęto próbę rozszerzenia klasycznego podejścia indykatorowego o ujęcie tranzytywne działające w wielookresowym horyzoncie czasowym. Cel ten obejmował zarówno opracowanie nowej konstrukcji indykatora, jak i ocenę możliwości jej zastosowania w warunkach niepełnej obserwowalności danych, typowej dla danych skanowanych. Realizacja tego celu została przedstawiona w artykule *The use of transitive Montgomery indicators for scanner data analysis* (artykuł numer 4) poświęconym tranzytywnej wersji indykatora Montgomery’ego.

Uzyskane wyniki potwierdziły przyjęte hipotezy badawcze. W szczególności wykazano, że możliwa jest konstrukcja multilateralnego indykatora Montgomery’ego, który działając na oknie czasowym spełnia kluczowe wymogi aksjomatyczne, w tym test zmiany wartości (*value change test*), test odwróconego czasu (*time reversal test*) oraz wielowymiarowy test identyczności (*multi-period identity test*) (**potwierdzenie hipotezy 4A**). Jednocześnie wykazano, że zaproponowana konstrukcja zachowuje swoje własności również w warunkach występowania niedopasowanych produktów, braków danych oraz obserwacji o zerowych cenach, co czyni ją adekwatną do analizy rzeczywistych danych skanowanych (**potwierdzenie hipotezy 4B**). Dodatkowo pokazano, że zastosowanie podejścia multilateralnego eliminuje problem łańcuchowego dryfu, charakterystyczny dla klasycznych ujęć dynamicznych.

Wkład do literatury polega tu na rozszerzeniu dotychczasowego nurtu badań nad indykatorami cen i ilości poprzez zaproponowanie ich tranzytywnych, multilateralnych wersji oraz ich zastosowanie do danych skanowanych (pierwsze tego typu zastosowanie w literaturze przedmiotu). W szczególności wykazano komplementarność podejścia indeksowego i indykatorowego, wskazując, że multilateralne indykatory umożliwiają

precyzyjną dekompozycję zmian wartości sprzedaży na komponent cenowy i ilościowy, przy jednoczesnym ograniczeniu obciążeń pomiarowych wynikających z łańcuchowego podejścia do indeksów. Tym samym zaproponowane rozwiązanie stanowi istotne uzupełnienie istniejących metod analizy inflacji opartych na danych skanowanych.

W ramach piątego celu badawczego (**Cel 5**), polegającego na porównaniu własności multilateralnych indykatorów Benneta i Montgomery'ego z wykorzystaniem danych skanowanych, ze szczególnym uwzględnieniem ich wrażliwości na rodzaj zastosowanego filtru danych oraz poziom agregacji danych, przeprowadzono analizę empiryczną przedstawioną w dwóch artykułach: *Zastosowanie indykatorów Benneta i Montgomery'ego w analizie danych skanowanych – porównanie własności i oszacowań efektów* (artykuł numer 5) oraz *Comparison of transitive Bennet and Montgomery indicators on the basis of scanner data sets* (artykuł numer 6). Badania te stanowią komplementarne ujęcie problemu, obejmujące zarówno analizę własności teoretycznych indykatorów, jak i ich empiryczną weryfikację przydatności na danych skanowanych.

Uzyskane wyniki potwierdziły przyjęte hipotezy badawcze. W szczególności wykazano, że różnice pomiędzy tranzytywnymi wersjami indykatorów Benneta i Montgomery'ego maleją wraz ze wzrostem poziomu agregacji danych, co wskazuje na ich zbieżność w analizach prowadzonych na wyższych poziomach agregacji (**potwierdzenie hipotezy 5A**). Jednocześnie potwierdzono, że zastosowanie filtrów danych ma istotny wpływ na wartości zarówno cenowych, jak i ilościowych indykatorów, przy czym w skrajnych przypadkach może prowadzić do zmiany ich znaku (**potwierdzenie hipotezy 5B**). Wskazuje to na wysoką wrażliwość wyników na decyzje metodyczne związane z przetwarzaniem danych.

Wkład do literatury polega na kompleksowym i empirycznie ugruntowanym porównaniu własności multilateralnych indykatorów Benneta i Montgomery'ego w warunkach dużej rotacji produktów skanowanych, ze szczególnym uwzględnieniem roli filtracji i agregacji danych. Uzyskane wyniki pozwalają na sformułowanie wniosków o charakterze metodycznym, istotnych dla praktyki nowoczesnej statystyki cen, w tym w szczególności w zakresie doboru filtrów danych oraz ustalania progów dla ich zastosowania.

Wyniki uzyskane w ramach przeprowadzonych badań w rozprawie doktorskiej wskazują również na przekrojową, kluczową rolę filtracji danych oraz poziomu

agregacji jako czynników determinujących wskazania zarówno indeksów, jak i indyktorów cen oraz ilości. Uzyskane wnioski mają istotne implikacje dla nowoczesnej statystyki cen, w szczególności w zakresie doboru procedur filtrowania danych oraz ustalania progów ich zastosowania.

Najważniejsze efekty przeprowadzonych badań i analiz są następujące:

- W przeprowadzonej analizie wykazano ograniczenia klasycznych indeksów cenowych (bilateralnych, łańcuchowych) w warunkach dużej rotacji produktów oraz wysokiej zmienności cen i ilości.
- Zaproponowano nowy multilateralny indeks cen (GEKS-GAQI), który ma cenne własności aksjomatyczne (m.in. wyróżnia go spełniony test identyczności) i jednocześnie dokonuje korekty jakościowej produktów (czyli jest niewrażliwy na zmiany jednostki sprzedaży w analizowanym zbiorze danych i działa w przypadku heterogenicznych zbiorów produktów).
- Uzupełniono nurt badań nad indeksami kosztów utrzymania poprzez wskazanie, że estymacja elastyczności substytucji w modelu CES jest silnie zależna od poziomu agregacji danych oraz wyboru metody estymacji, co ma istotne implikacje dla praktyki konstrukcji indeksów kosztów utrzymania typu CES. W szczególności wykazano przewagę metod algebraicznych nad metodami regresji panelowej w zakresie stabilności oszacowań elastyczności substytucji, co stanowi istotne uzupełnienie dotychczasowych badań empirycznych w tym zakresie opartych głównie na podejściu ekonometrycznym.
- Literaturę przedmiotu dotyczącą indyktorów cen i ilości wzbogacono o wszechstronną analizę tranzytywnych wersji indyktorów Benneta i Montgomery'ego z wykorzystaniem danych skanowanych, co - według najlepszej wiedzy autorki cyklu prac – jest pionierskie i suplementarne w stosunku do tranzytywnych indeksów cen. W szczególności wykazano komplementarność podejścia indeksowego i indyktorowego w analizie danych skanowanych wskazując, że multilateralne indykatory umożliwiają precyzyjną dekompozycję sumarycznej wartości sprzedaży produktów na efekt cenowy i ilościowy przy jednoczesnej eliminacji obciążenia pomiaru z tytułu łańcuchowego dryfu.
- Zidentyfikowano kluczową rolę filtracji danych oraz poziomu agregacji danych na jakim przeprowadzana jest analiza jako czynników determinujących

wskazania zarówno indeksów, jak i indykatorów cen (także indykatorów ilości). Wnioski z tej części badań mogą mieć ważne implikacje dla nowoczesnej statystyki cen (choćby w postaci rekomendacji wobec wyboru filtru danych i jego progów odcięcia).

Przedstawiony cykl publikacji stanowi spójny i samodzielny wkład w rozwój metodologii pomiaru inflacji, odpowiadając na realne potrzeby zarówno teorii indeksów cen, jak i praktyki statystycznej.

Jednocześnie przeprowadzone badania ujawniają obszary wymagające dalszej eksploracji. Warto zaznaczyć, że nie wszystkie dostępne formuły indeksów multilateralnych zostały dotychczas poddane pełnej weryfikacji empirycznej w warunkach charakterystycznych dla danych skanowanych, w szczególności w kontekście zjawisk takich jak *downsizing* (zmniejszenie gramatury produktu przy utrzymaniu ceny), czy *upsizing* (zwiększenie gramatury, często połączone ze zmianą ceny). Zjawiska te stanowią formę ukrytego wzrostu cen (to tzw. *ukryta inflacja*), trudnego do uchwycenia przez standardowe formuły indeksowe nieuwzględniające korekt jakościowych i zasługują na odrębne opracowanie metodologiczne. Weryfikacja własności indeksów i indykatorów w warunkach intensywnych zmian gramatury produktów stanowi zatem jeden z kierunków, które warto podjąć w kolejnych pracach badawczych. Kolejnym, potencjalnym kierunkiem dalszych badań może być ocena wpływu długości okna czasowego na jakim pracują formuły multilateralne na uzyskiwane wyniki oszacowań dynamiki cen skanowanych. Obserwowany rozwój alternatywnych źródeł danych, w tym danych pozyskiwanych z Internetu (*web-scraped data*), stwarza zatem nowe możliwości dla dalszego rozwijania metodologii zaproponowanej w niniejszym cyklu.

Bibliografia

Australian Bureau of Statistics. (2016). *Making greater use of transactions data to compile the Consumer Price Index* (Information Paper 6401.0.60.003). Canberra.

Białek, J. (2023). *Quality adjusted GEKS-type indices for price comparisons based on scanner data*. *Statistics in Transition New Series*, 24, 151–169.

Białek, J., Kłopotek, M., i Panek, T. (red.). (2024). *Nowoczesne technologie i nowe źródła danych w pomiarze inflacji*. Warszawa: Główny Urząd Statystyczny.

Białek, J., & Pawelec, N. (2024a). *The use of transitive Montgomery indicators for scanner data analysis*. *Argumenta Oeconomica*, 2(53), 5–26.
<https://doi.org/10.15611/aoe.2024.2.01>

Białek, J., & Pawelec, N. (2024b). *Comparison of transitive Bennet and Montgomery indicators on the basis of scanner data sets*. *European Research Studies Journal*, 27(4), 472–485. <https://doi.org/10.35808/ersj/3528>

Białek, J., & Pawelec, N. (2025). *Proposition of the quality adjusted GEKS-type price index*. *Statistika: Statistics and Economy Journal*, 105(3), 291–305.
<https://doi.org/10.54694/stat.2024.71>

Białek, J., Pawelec, N., & Roszkowska, S. (2024). *Estimating the elasticity of substitution when compiling the CES cost of living index on scanner data*. *Quality & Quantity*. <https://doi.org/10.1007/s11135-024-01924-8>

de Haan, J., Willenborg, L., & Chessa, A. G. (2016). *A review of price index methods for scanner data*. Statistics Netherlands (CBS).

Diewert, W. E. (1976). *Exact and superlative index numbers*. *Journal of Econometrics*, 4(2), 115–145.

Eurostat. (2018). *Practical guide for processing supermarket scanner data: Harmonised Index of Consumer Prices*. Urząd Publikacji Unii Europejskiej.

Eurostat. (2022). *Guide on multilateral methods in the Harmonised Index of Consumer Prices*. Urząd Publikacji Unii Europejskiej. <https://ec.europa.eu/eurostat/>

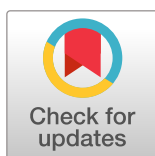
- Federal Reserve Bank of Cleveland (2026). *Why Should You Care about Inflation? Center for Inflation Research*. <https://www.clevelandfed.org/center-for-inflation-research/inflation-explained-your-guide-to-inflation-basics/why-should-you-care-about-inflation> [dostęp 27.05.2026]
- Fox, K. J. (2006). A method for transitive and additive multilateral comparisons: A transitive Bennet indicator. *Journal of Economics*, 87(1), 73–87. <https://doi.org/10.1007/s00712-005-0160-8>
- Hałka, A., Leszczyńska, A. (2011). *Wady i zalety wskaźnika cen towarów i usług konsumpcyjnych – szacunki obciążenia dla Polski*. *Gospodarka Narodowa*, nr 9(241), s. 51–75 <https://doi.org/10.33119/GN/101086>
- International Labour Organization. (2020). *Consumer price index manual: Concepts and methods*. International Labour Office. <https://doi.org/10.30875/24ad1c9e-en>
- Leclair, M., Léonard, I., Rateau, G., Sillard, P., Varlet, G., & Vernédal, P. (2020). *Scanner data: Advances in methodology and new challenges for computing consumer price indices*. OECD Publishing. <https://doi.org/10.1787/e7a8f5d2-en>
- Nickel, C. I in. (2021). *Inflation measurement and its assessment in the ECB's monetary policy strategy review*. ECB Occasional Paper Series, No. 265. European Central Bank. [dostęp 27.05.2026]
- Pawelec, N. (2026). *The Use of Scanner Data for Price Indices Calculation: the Case of Poland*. *Acta Universitatis Lodzensis. Folia Oeconomica*, 1(374), 22-45.
- Pawelec, N. (2025). *Zastosowanie indykatorów Benneta i Montgomery'ego w analizie danych skanowanych – porównanie własności i oszacowań efektów*. *Wiadomości Statystyczne. The Polish Statistician*, 70(3), 11-23.
- Turvey, R. (1989). *Consumer price indices: An ILO manual*. International Labour Office.
- von der Lippe, P. (2007). *Index theory and price statistics*. Peter Lang.

Artykuł 1

Artykuł 1 The use of scanner data for price indices calculation: the case of Poland

opublikowany w

Acta Universitatis Lodzensis. Folia Oeconomica nr 374(1)/2026



Natalia Pawelec  <https://orcid.org/0009-0008-5289-6204>

University of Lodz, Faculty of Economics and Sociology, Lodz, Poland, natalia.pawelec@uni.lodz.pl

The Use of Scanner Data for Price Indices Calculation: the Case of Poland

Abstract:

In the context of the growing use of scanner data in inflation measurement, a key methodological challenge remains the choice of the appropriate price index formula and data filtering method. This article aims to investigate the impact of these choices on the results of price dynamics measurement. The paper analyses transactional data collected from point-of-sale terminals in Polish supermarkets for selected product groups: coffee and milk. Various price index formulas – including unweighted bilateral (Dutot, Carli, Jevons), weighted bilateral (Törnqvist, Fisher), and multilateral (Geary-Khamis, CCDI) – were compared, and the influence of different scanner data filtering methods was tested. The results show that both the choice of the price index formula and the applied scanner data filtering method have a measurable and significant impact on the assessment of price dynamics. These differences can lead to diverse interpretations of inflation, which underscores the need for conscious and justified methodological choices when using scanner data in price statistics.

Keywords: scanner data, multilateral indices, bilateral indices

JEL: C43

Funding information: University of Lodz, Faculty of Economics and Sociology, Lodz, Poland.

Declaration regarding the use of GAI tools: Not used.

Conflicts of interests: None.

Ethical considerations: The Author assures of no violations of publication ethics and takes full responsibility for the content of the publication.

Received: 2025-02-25. Revised: 2025-11-10. Accepted: 2026-02-25



© by the Author, licensee University of Lodz – Lodz University Press, Lodz, Poland.
This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license CC-BY (<https://creativecommons.org/licenses/by/4.0/>)



This journal adheres to the COPE's Core Practices
<https://publicationethics.org/core-practices>

1. Introduction

The Consumer Price Index (CPI) should reflect actual inflation as accurately as possible for various reasons. One of them is the fact that it is used to index nominal values in the economy, influencing social policy and pricing decisions by business owners. One of the methodological challenges of consumer price research is the necessity to diversify data sources and automate data processing. The increasing availability of electronic transaction datasets for the Consumer Price Index (CPI) presents a significant opportunity for national statistical institutes (NSIs) to elevate the precision of their index numbers. The data facilitate the deployment of more advanced analytical techniques that can account for the fluid nature of consumption habits far more effectively than conventional fixed-basket approaches (Chessa, 2021). One solution is acquiring and processing scanner data. We define scanner data as ‘transaction data obtained from retail chains containing data on turnover, quantities per item code based on transactions for a given period and from which unit value prices can be derived at item code level’ (Eurostat, 2017).

Scanner data are generated by point-of-sales terminals in shops which log each transaction. Scanner data can be sourced from numerous types of retailers, including supermarkets, pharmacies, hardware stores, electronics shops, clothing stores, and many others. Presently, scanner data are mainly used to substitute price collection in supermarkets, especially for items such as food, beverages, and the personal and household care products available there. As previously mentioned, this guide will focus exclusively on these product categories. Scanner data have several advantages over traditional price collection. They are relatively inexpensive, with their collection and processing being automated. They include information on the actual expenditure for all item codes sold. Integrating scanner data into inflation measurement enhances the data collection process, reduces associated costs, and provides a more accurate representation of shifts in consumer behaviour.

The main objective of this study was to investigate how the choice of a price index formula impacts the final inflation indicator values. Specifically, the aim was to examine the differences that may occur between the values of bilateral and multilateral indices. The analysis was conducted using scanner data from the Polish retail market, covering the period from December 2019 to December 2020. The study focuses on comparing several types of price indices: unweighted bilateral indices (Dutot, Carli, Jevons), weighted bilateral indices (Törnqvist, Fisher), and multilateral indices (Geary-Khamis, CCDI). According to Eurostat (2020b), multilateral indices are the most strongly recommended approach for scanner data analysis, as they better account for product turnover and mitigate chain drift.

The results clearly indicate that the selection of a specific price index formula significantly influences the values of the inflation indicator, which leads to varied interpretations of price dynamics. The differences in the calculation results are presented and discussed in detail, allowing for an assessment of the methodological consequences within the Polish context.

2. Measuring Inflation

2.1. Indices for Measuring Inflation

The CPI is a measure of inflation that is closely observed by consumers, businesses, and policymakers as price movements in the economy reflect changes in the value of money over time. In most countries, inflation measurement is conducted by statistical offices, although central banks (e.g., the National Bank of Poland) are also interested in this measurement. In Poland, Statistics Poland makes calculations based on representative goods and services, including typical products or services most commonly purchased by households (Statistics Poland, 2019).

In official statistics, two price indices are used to measure inflation. The first is the CPI, which represents the inflation rate in a given country. It approximates changes in household consumption costs to maintain a constant utility level (Hałka, Leszczyńska, 2011) and serves as a proxy for the Cost of Living Index (COLI). The CPI measures changes in the prices of goods and services purchased by households for consumption and informs monetary policy, social protection benefits, financial instruments, and the indexing of commercial contracts, wages, pensions, and retirement benefits (Eurostat, 2024).

The second indicator is the Harmonised Index of Consumer Prices (HICP), created by the European Central Bank to ensure a high-quality comparison of consumer price inflation between countries. The HICP aims to deliver a reliable and consistent metric for comparing consumer price inflation. It employs different formulas to aggregate prices and price indices depending on whether products are weighted or not. The HICP is a cost of goods (and services) index, i.e., it measures the changing costs of a fixed basket of products at different sets of prices over time. Prices used to measure the HICP should reflect those actually paid by households and must include costs related to reservations and deliveries often associated with online purchases (Eurostat, 2024).

The CPI is measured monthly. To accurately reflect the current consumption patterns of households, the data must be sourced from various sources. According to Statistics Poland: 'Price data are sourced from price lists, regulations, and decisions regarding uniform prices applicable nationwide or in parts thereof, issued by government and local administration bodies, and economic entities (e.g., prices of selected pharmaceutical products, notarial services, or banking services). Another method of data collection is searching for appropriate data on websites and virtual domains of stores. This method is used to gather data on, among other things, airline ticket fees' (Statistics Poland, 2019).

Statistics Poland collects over 230,000 prices monthly, with the most significant number (72,000) being single prices for the 'food and non-alcoholic beverages' category. Statistics Poland also gathers information on the characteristics of selected products, such as raw material composition, model, and weight. To measure inflation, it is necessary to have information on price notations for identical goods and services in two compared periods

and on the expenditure structure of households (Statistics Poland, 2019). The number of price notations collected is increasing yearly due to the growing diversity of goods and services offered, as well as new ways to acquire such information, such as using scanner data.

The weights necessary to calculate the CPI are obtained by grouping expenditures on goods and services collected from household budget surveys. The 'inflation basket' structure considers typical goods and services most commonly purchased by households.

Data used to construct the HICP weights also come from various sources. National accounts data (at different levels of aggregation) are the main source because regulations require this. Their undeniable advantage is that they are structured and organised according to the European Classification of Individual Consumption by Purpose. Other data sources include household surveys, market research, retail trade data, and administrative sources. The traditional method of price recording by interviewers has a well-established methodology and a long history. However, technological advancements, particularly in information technology, enable the use of alternative sources of product price data, such as scanner data (which are primarily obtained from supermarket chains) and scraped data (which are automatically collected from websites). The scope of price notations collected from alternative sources has been increasing yearly (Statistics Poland, 2024).

3. Scanner Data as a Modern Data Source in Inflation Measurement

3.1. Advantages, Disadvantages, and Methodological Challenges

The advent of barcode scanner technology in the 1970s gave rise to scanner data, a new form of information quickly recognised for its value in understanding consumer choices and demands. The data have since fuelled extensive research across various economic behaviours, particularly in industrial organisation, marketing, and increasingly in public, health, and macro/monetary economics (Dubois, Griffith, O'Connell, 2022). Scanner data serve as a modern source of transaction data, providing highly detailed and accurate information due to the immense number of observations obtained by scanning consumer goods barcodes. The most common barcodes are GTIN (Global Trade Item Number) and EAN (European Article Number), the European version of GTIN (Eurostat, 2017).

Scanner data are a modern source of information about prices and consumption levels. Specifically, they provide data on prices and quantities of purchased products at the lowest level of data aggregation, namely at the individual barcode level. One of the many advantages of using scanner data is that they contain information at the barcode or GTIN level, enabling products to be 'recognised' and classified into appropriate COICOP groups.

A clear benefit of scanner data is that they provide additional product information and information related to the seller's code, point-of-sale identifier, product label, unit of sale (e.g., 'pcs,' 'kg'), sales value, number of units sold, flag (containing information, e.g., about discounted products), and VAT information. This is useful for aggregating elements into homogeneous groups. Scanner data are collected automatically and relatively cheaply (Białek,

2020a). Moreover, they provide information about actual expenditures for all sold items, excluding unsold products. These data typically cover 2–4 weeks of each month and are sometimes transferred with daily or weekly frequency. All these advantages suggest that using scanner data offers hope for improving the quality of the HICP (Eurostat, 2017).

However, one problem that arises from using scanner data is the overly detailed level of information about the product contained in the GTIN code. The GTIN code comprises 13 or 14 digits, including: 1 digit to indicate the packaging level, a 3-digit GS1 national organisation code (country code), a 4–7-digit GS1 coding unit number, a 2–5-digit product code, and a check digit. This makes it challenging to identify homogeneous product groups using this code because different codes may represent the same product with only packaging variations. In addition to GTIN and EAN barcodes, other barcodes, such as PLU (Price Look-Up) and SKU (Stock Keeping Unit), also exist. The PLU code is shorter, and the SKU is more general than the GTIN, and in practice, further complicating product identification.

The choice of the time frame can be problematic when aggregating scanner data because many products change periodically, sometimes every quarter, month, week, or even day. Prices are often higher on Saturdays, and sales value, transaction numbers, and prices peak in the afternoon or evening (Białek, 2020a).

The advantages and disadvantages of using scanner data are also linked to methodological challenges that must be addressed during scanner data processing. One challenge is the need to process and prepare the acquired data, which requires an appropriate IT structure and qualified personnel. Scanner data can be sourced from points of sale, such as supermarkets, pharmacies, and travel agencies. However, the most valuable suppliers are supermarkets with a large number of barcodes (typically between 10,000 and 25,000 barcodes, mainly for food and beverages). Another source of scanner data is market research companies and electronic trading platforms (e.g., Allegro or OLX, which operate in the Polish market). This results in greater dependence on the supplier than in the past when it was only necessary to obtain permission to visit the sales point. Nowadays, maintaining good relationships with retailers and having legal agreements with retailers to define the cooperation details is crucial. Another significant methodological challenge is selecting the appropriate price index formula for elementary homogeneous product groups.

4. Choice of Price Index Formula

The choice of formula is one aspect of the broad spectrum of decisions to be made when working with scanner data. When describing bilateral and multilateral methods (indices), it is important to acknowledge that using new data sources (including scanner data) is highly demanding for statistical offices. This is due to the vast volume of data and the high turnover of products associated with these sources (Statistics Poland, 2024).

In traditional data collection, elementary indices are used for the lowest level of data aggregation, i.e., where data on only the prices of goods and services are collected by interviewers in the field. Weighted indices are used at higher data aggregation levels, where the statistical

office has knowledge of the level of consumption obtained through the Household Budget Survey (Eurostat, 2024). Nevertheless, the exception to this rule is scanner data. Scanner data, even at the lowest level of data aggregation (i.e., at the barcode level), provide information on the value of product sales; thus, it is then possible to use both weighted and unweighted indices (Eurostat, 2017). Some EU countries use the **dynamic approach**, whereby a sample of products is taken from month to month by appropriate data filtering (e.g., a low sales filter is applied), and a chained Jevons index (unweighted formula) is calculated (Van Loon, Roels, 2018). However, there are also countries that implement weighted multilateral indices, e.g., Statistics Netherlands has implemented the Geary-Khamis index into regular production at the barcode level (Chessa, 2018).

The economic literature offers a wide range of price indices, each with its own specific applications as well as theoretical and empirical justifications. This analysis focuses on the indices selected by the author. A broader set of measures could be considered; however, to maintain clarity and transparency of the discussion, it was necessary to make a selection.

4.1. Bilateral Indices

Most National Statistical Offices use bilateral indices to measure the CPI, particularly in the traditional data collection framework, using the Jevons index at the lowest aggregation level and the Laspeyres index at the highest level (Białek, 2020b). These methods compare current prices and quantities of goods with corresponding prices and quantities from a base (fixed) period, making them transparent and easy to explain to users.

Unweighted Indices

The three most known elementary formulas are the Dutot index (Dutot, 1738), the Carli index (Carli, 1804) and the Jevons index (Jevons, 1865), all of which are based on unweighted average prices or price ratios (Eurostat, 2020b). Eurostat prohibits the use of the Carli index because it generates charges that overestimate inflation. The Dutot and Jevons indices, which are the two elementary aggregate formulas laid down in HICP regulations, do not suffer from this problem (Eurostat, 2024). According to EUROSTAT, the most recommended formula is the Jevons index (Eurostat, 2020a)

The Dutot index is defined as the ratio of the average prices of goods from the current and base months:

$$P_D^{0,t} = \frac{\frac{1}{N_{o,t}} \sum p_i^t}{\frac{1}{N_{o,t}} \sum p_i^0}, \quad (1)$$

where:

$N_{o,t}$ – the number of matched products in the compared months;

$\tau = 0$ – the base period;

$\tau = t$ – the current period;

p_i^τ – the price of the i -th product in the period τ .

The second formula discussed in the article is the Carli index, which represents the arithmetic mean of the partial indices:

$$P_c = \frac{1}{N_{o,t}} \sum_{i=1}^N \frac{p_i^t}{p_i^0}. \quad (2)$$

The last unweighted index discussed is the Jevons price index, which is calculated as the geometric mean of the partial price indices:

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

where:

p_i^τ – the price of the i -th product at the time $\tau \in \{0, t\}$;

$G_{o,t}$ – products observed simultaneously in the months o (base) and t (current);

$N_{o,t} = \text{card } G_{o,t}$.

Weighted Indices

The second group of indices are weighted indices, including superlative indices firstly proposed by Diewert (1976), such as the Törnqvist index (Törnqvist, 1936):

$$P_T^{0,t} = \prod_{i \in G_{0,t}} \left(\frac{p_i^t}{p_i^0} \right)^{\frac{s_i^0 + s_i^t}{2}}, \quad (4)$$

where:

s_i^0 and s_i^t represent the expenditure shares of the i -th in the base and current periods, respectively.

The Fisher index (Fisher, 1922) is the geometric mean of the Laspeyres (1871) ($P_{La}^{0,t}$) and Paasche ($P_{Pa}^{0,t}$) indices (Paasche, 1874):

$$P_F^{0,t} = \sqrt{P_{La}^{0,t} P_{Pa}^{0,t}}, \quad (5)$$

where:

$P_{La}^{0,t}$ – the Laspeyres index,

$P_{Pa}^{0,t}$ – the Paasche index.

4.2. Multilateral Indices

Multilateral indices are particularly useful for scanner data and were initially introduced to compare price levels between countries or regions (Gini, 1931; Éltető, Köves, 1964; Chessa, 2021). They provide transitive price comparisons, which is a desirable property because the results are independent of the choice of the base country. As such, multilateral indices are considered a solution to the problems encountered with bilateral indices (Eurostat, 2020b). The price index measures the aggregate price change in the current period compared to the base period. Multilateral indices include all products available in at least two periods of a time window, usually 13 months, ensuring that the seasonality of products is considered. Multilateral indices weigh each product according to its significance in each period and help avoid issues related to chain drift found in bilateral indices (Eurostat, 2020b). Numerous studies emphasise that the use of multilateral indices in scanner data analysis reveals their advantages over simple bilateral formulas (see, e.g., Vartia, 2018; Diewert, 2022).

Multilateral indices are transitive, meaning their result for any two periods within the time window does not depend on the choice of the base period. This transitivity is crucial for eliminating chain drift. 'In case of promotional sales with reduced prices, the quantities purchased often increase substantially. But when the prices return to their original level, the quantities purchased of storable goods may not return to their 'normal' level. This type of asymmetric behaviour can cause chain drift in superlative price indices, which is typically downward' (Eurostat, 2020b). Well-known multilateral methods include the GEKS index (Gini, 1931), the Geary-Khamis (GK) index (Geary, 1958), and the CCDI (Caves Christensen and Diewert) index (Caves, Christensen, Diewert, 1982; Eurostat, 2020b).

The standard formula of the GEKS index (Gini, 1931; Éltető, Köves, 1964), which was originally used for international comparisons between countries or regions, is based on the superlative Fisher index (Fisher, 1922). The GEKS index between the period 0 and period t is calculated as the unweighted geometric mean of $T + 1$ bilateral price indices $P^{t,t}$ and $P^{t,0}$, which are based on the same formula. Bilateral indices should pass the time reversal test, i.e., they should satisfy $P^{a,b} * P^{b,a} = 1$. The formula for the GEKS index can be expressed as:

$$P_{\text{GEKS}}^{0,t} = \prod_{\tau=0}^T \left(\frac{P_F^{\tau,t}}{P_F^{\tau,0}} \right)^{\frac{1}{T+1}}. \quad (6)$$

The Geary-Khamis index (Geary, 1958; Khamis, 1972) can be expressed as:

$$P_{\text{GK}}^{0,t} = \frac{\sum_{i \in G_t} p_i^t q_i^t / \sum_{i \in G_0} p_i^0 q_i^0}{\sum_{i \in G_t} v_i^{\text{GK}} q_i^t / \sum_{i \in G_0} v_i^{\text{GK}} q_i^0}, \quad (7)$$

where:

$$v_i^{\text{GK}} = \sum_{z=0}^T \varphi_{i,\text{GK}}^z \frac{p_i^z}{P_{\text{GK}}^{0,z}}, \quad (8)$$

$$\varphi_{z, GK}^z = \frac{q_i^z}{\sum_{\tau=0}^T q_i^{\tau}}. \quad (9)$$

The CCDI index (Caves, Christensen, Diewert, 1982) is essentially the GEKS index based on the Törnqvist formula, replacing the 'standard' Fisher index in GEKS. It is expressed as follows:

$$P_{CCDI}^{0,t} = \prod_{\tau=0}^T \left(\frac{P_T^{\tau,t}}{P_T^{\tau,0}} \right)^{\frac{1}{T+1}}, \quad (10)$$

where:

$$P_T^{\tau,t} = \prod_{i \in G_{\tau,t}} \left(\frac{p_i^t}{p_i^{\tau}} \right)^{\frac{s_j^{\tau} + s_j^t}{2}},$$

s_j^{τ} – the expenditure share of the i -th product at the time τ ;

s_j^t – the expenditure share of the i -th product at the time t .

Despite their complexity, multilateral indices are recommended for compiling price indices from transaction data. Due to their valuable axiomatic properties and the fact that they operate over the full-time window, multilateral indices are well-suited for index calculations when working with scanner data. Research is still ongoing to determine the best multilateral indices for compiling elementary groups using scanner data. Scanner data offer many possibilities for new research and development.

5. Selection of Samples: Static and Dynamic Approaches

Processing scanner data involves several key stages. One of them is to select the appropriate price index, which was discussed earlier. Another crucial step is to determine the sampling approach for products, which directly affects how the sales structure is represented in price calculations. In this context, two primary approaches are distinguished: the static approach and the dynamic approach (Van Loon, Roels, 2018).

5.1. The Static Approach

The static approach closely resembles traditional price collection methods and relies on a fixed product sample. Product codes are selected at the end of the year and remain in use for the following 12 months. This selection is carefully designed to ensure that the sample accurately represents the sales structure for the entire year while minimising the impact of seasonal sales fluctuations, such as those typical in December.

Maintaining this sample requires balancing representativeness and feasibility. If a product becomes less representative or is discontinued, it must be replaced with an equivalent item. While this method follows conventional sampling procedures, it benefits from full transaction data, allowing for a more informed selection of products and adjustments as necessary throughout the year.

The static approach is particularly useful when scanner data are used on a limited scale and need to be integrated with traditionally collected price data. In such cases, it may be practical to manually select product codes that best match the descriptions used in conventional price collection. However, this method also has drawbacks – it is labour-intensive and does not fully exploit the wealth of available scanner data.

5.2. The Dynamic Approach

The dynamic approach involves automatically updating the product sample on a monthly basis. This method selects product codes that appear in both the current and previous month and that account for a significant share of turnover in a given elementary aggregate. This ensures that important new products are included while less relevant ones are removed. Research shows that while a wide range of goods is available on the market, only a small portion of them generate the majority of turnover. The dynamic method takes this into account.

This method resembles a continuous replenishment system and chain-linked index calculation, making it well-suited for handling large volumes of scanner data. Its key advantage lies in its automation, which greatly simplifies data processing. However, frequent product relaunches or replacements must be handled separately to ensure appropriate quality adjustments.

The dynamic basket utilises a set of filters and an algorithm to select a matched sample. The dump filter allows for the removal of products that experience a simultaneous significant drop in price and sales value, indicating that the product will soon be withdrawn and will no longer be representative.

The low-sales filter is designed to eliminate product codes with very low sales or, conversely, to ensure that the selected products account for a sufficiently large share of total turnover (ranging from 50% to 80%). The low-sales algorithm is defined as:

$$\frac{s_{m-1} + s_m}{2} > \frac{1}{(n \times \lambda)},$$

where:

s_{m-1} and s_m – the turnover share in the month $m - 1$ and m ;

n – the number of matched items in the elementary aggregate;

$\lambda = 1.25$.

The value 1.25 for λ is empirically determined so that the selected item codes represent about 80% of turnover. Since the filter is based on turnover value, certain product categories, such as low-value or unbranded goods, are automatically excluded from the sample, as their low prices contribute to a low turnover.

6. Empirical Example

This chapter uses a sample set of scanner data to present the behaviour of various price indices, distinguishing between bilateral indices (weighted and unweighted) and multilateral indices. In the subsequent section, the dynamics of these indices are analysed using selected filters: the low sales filter and the dump prices filter.

The following empirical study uses daily scanner data from a retail chain in Poland for the period December 2018 to December 2019. The collections come from the PriceIndices package in R (Białek, 2021), which is publicly available and free of charge. The study analyses two datasets: coffee (which contains 42,561 observations (rows)) and milk (4,386 rows). The datasets contain six columns and include the following variables:

- time – transaction dates (year-month-day);
- prices – prices of sold products (PLN);
- quantities – quantities of sold products (milk in litres, coffee in kilograms);
- prodID – unique product codes;
- retID – unique identifiers for retail outlets;
- description – descriptions of sold products.

The milk dataset contains six different product descriptions that correspond to **subgroups** of the milk group, i.e., powdered milk, low-fat pasteurised milk, low-fat UHT milk, full-fat pasteurised milk, full-fat UHT milk, and goat milk. The coffee dataset includes three different homogeneous products: instant coffee, coffee beans, and ground coffee.

Below is an example of a scanner dataset used to calculate indices:

Table 1. Data frame with the first six milk harvest observations

	time	prices	quantities	prodID	retID	description
1	2018-12-01	8.78	9.0	14 215	2 210	powdered milk
2	2019-01-01	8.78	13.5	14 215	2 210	powdered milk
3	2019-02-01	8.78	0.5	14 215	1 311	powdered milk
4	2019-02-01	8.78	8.0	14 215	2 210	powdered milk
5	2019-03-01	8.78	0.5	14 215	1 311	powdered milk
6	2019-03-01	8.78	1.5	14 215	2 210	powdered milk

Source: PriceIndices package (Białek, 2021).

6.1. Unweighted and Weighted Bilateral Indices

Figures 1 and 2 compare selected elementary chain price indices calculated for the product groups. Figure 1 compares selected unweighted indices, while Figure 2 shows weighted bilateral indices. Tables 2 and 3 present the values of individual indices for each product category.

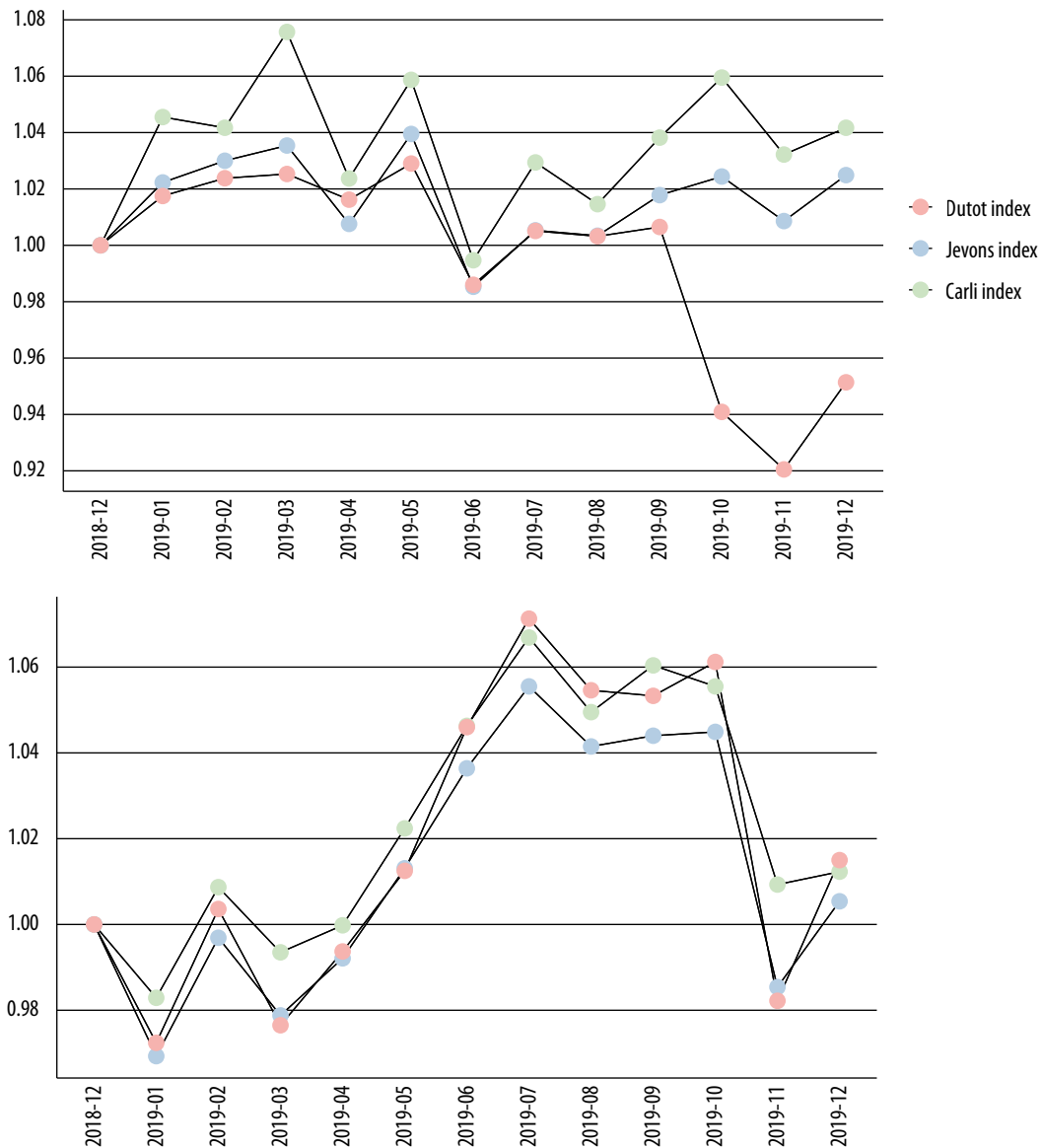


Figure 1. Unweighted bilateral indices

Source: own elaboration based on data from the PriceIndices package.

As shown in Figure 1, the values of individual unweighted indices differ slightly. For the coffee dataset, the indices generate similar values. The milk dataset shows more significant differences between the indices. Looking at the milk and coffee charts, it is not possible to unequivocally state which index generates the lowest or highest values as these relationships differ between datasets. Table 1 presents the values of the indices for each month discussed.

Table 2. Unweighted bilateral index values for milk products

Data	Dutot index	Jevons index	Carli index
2018-12	1.0000	1.0000	1.0000
2019-01	1.0175	1.0223	1.0455
2019-02	1.0238	1.0300	1.0417

Data	Dutot index	Jevons index	Carli index
2019-03	1.0253	1.0354	1.0757
2019-04	1.0162	1.0076	1.0237
2019-05	1.0290	1.0395	1.0587
2019-06	0.9860	0.9853	0.9947
2019-07	1.0051	1.0053	1.0294
2019-08	1.0032	1.0034	1.0146
2019-09	1.0065	1.0178	1.0382
2019-10	0.9409	1.0244	1.0595
2019-11	0.9205	1.0086	1.0322
2019-12	0.9514	1.0249	1.0417

Source: own elaboration based on data from the PriceIndices package in R.

Examining the values of the individual indices shows that the lowest values were recorded for the Dutot index and the highest for the Carli index. The most significant differences occurred in October, November, and December between the Dutot index and the Carli index, with discrepancies of -0.1186 pp, 0.1117 pp, and 0.0903 pp, respectively.

Table 3. Unweighted bilateral index values for coffee products

Data	Dutot index	Jevons index	Carli index
2018-12	1.0000	1.0000	1.0000
2019-01	0.9724	0.9693	0.9829
2019-02	1.0036	0.9969	1.0087
2019-03	0.9765	0.9788	0.9935
2019-04	0.9937	0.9921	0.9998
2019-05	1.0125	1.0131	1.0224
2019-06	1.0460	1.0364	1.0463
2019-07	1.0713	1.0555	1.0669
2019-08	1.0546	1.0415	1.0495
2019-09	1.0533	1.0440	1.0604
2019-10	1.0612	1.0449	1.0555
2019-11	0.9822	0.9854	1.0093
2019-12	1.0150	1.0054	1.0123

Source: own elaboration based on data from the PriceIndices package in R.

The unweighted bilateral indices calculated for coffee are characterised by smaller differences than for milk. In this case, the most significant differences in values between the indices occur in November, with a difference of 0.0270 pp. between the Dutot and Carli indices and -0.0239 pp. between the Dutot and Jevons indices.

In the following analysis, weighted bilateral indices for the products were examined. Figure 1 presents a graph comparing the indices, and Tables 3 and 4 show the values for each chart.

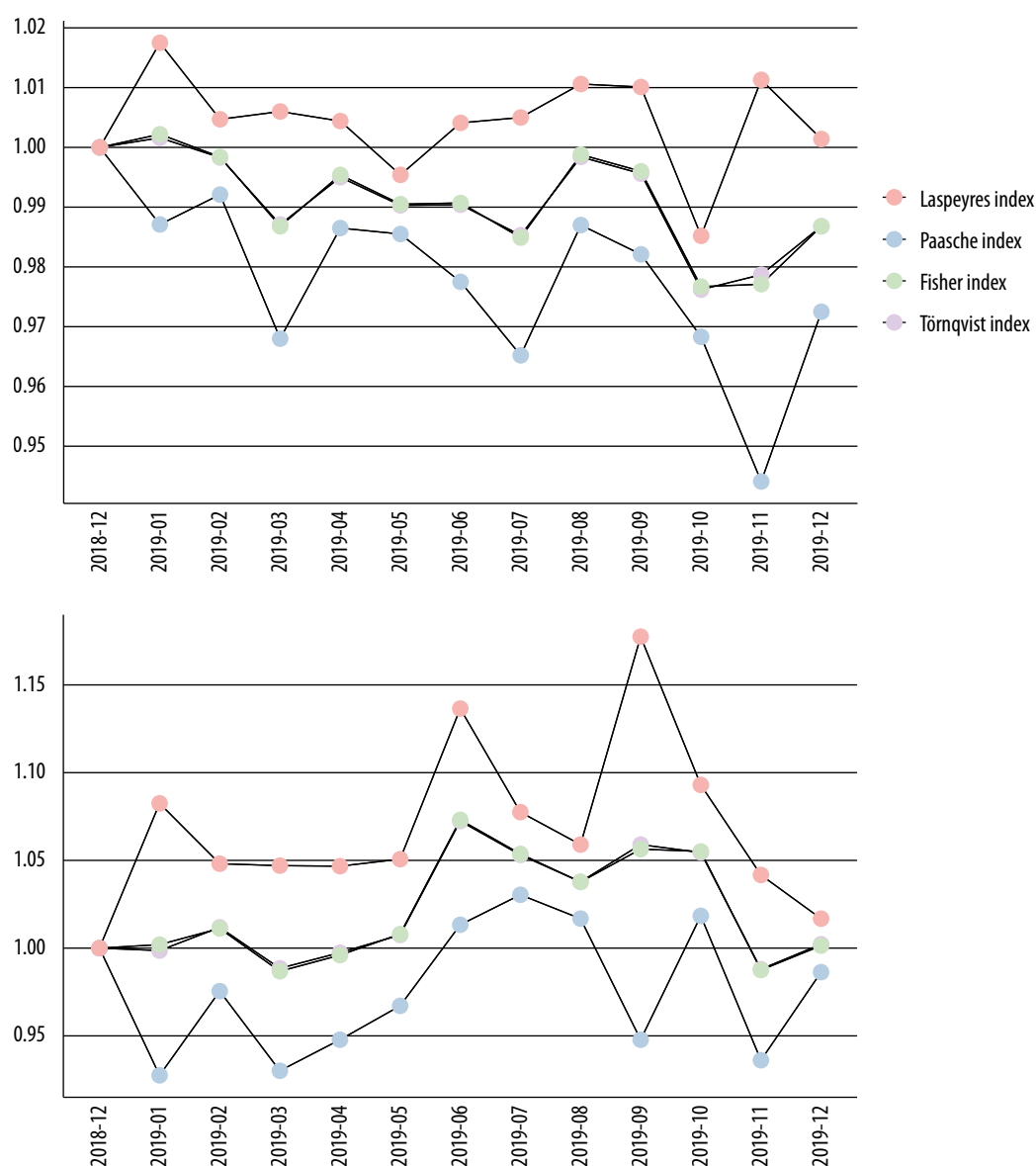


Figure 2. Weighted bilateral indices

Source: own elaboration based on data from the PriceIndices package in R.

For the weighted indices, a consistent pattern emerges across both datasets. The Fisher and Törnqvist indices have very close values, which is consistent with the analytically demonstrated mutual approximation of superlative indices (Białek, Beręsewicz, 2021). Conversely, the Laspeyres index generates the highest values, while the Paasche index generates the lowest. It is not clearly established whether a fixed-base approach or a chain version of these indices is better for the Fisher and Törnqvist indices.

Table 4. Weighted bilateral index values for milk products

Data	Laspeyres index	Paasche index	Fisher index	Törnqvist index
2018-12	1.0000	1.0000	1.0000	1.0000
2019-01	1.0175	0.9871	1.0022	1.0016
2019-02	1.0047	0.9921	0.9984	0.9983
2019-03	1.0060	0.9680	0.9868	0.9871
2019-04	1.0044	0.9865	0.9954	0.9950
2019-05	0.9954	0.9855	0.9905	0.9903
2019-06	1.0041	0.9775	0.9907	0.9904
2019-07	1.0050	0.9652	0.9849	0.9853
2019-08	1.0106	0.9870	0.9988	0.9984
2019-09	1.0101	0.9821	0.9960	0.9956
2019-10	0.9852	0.9683	0.9767	0.9762
2019-11	1.0113	0.9441	0.9771	0.9787
2019-12	1.0014	0.9725	0.9868	0.9868

Source: own elaboration based on data from the PriceIndices package in R.

For the weighted bilateral indices for milk, the most significant differences can be observed in November, with a difference of 0.0673 pp between the Laspeyres and Törnqvist indices. In March and July, significant differences were also noted for these indices, amounting to 0.0380 pp and 0.0398 pp, respectively.

Table 5. Weighted bilateral index values for coffee products

Data	Laspeyres index	Paasche index	Fisher index	Törnqvist index
2018-12	1.0000	1.0000	1.0000	1.0000
2019-01	1.0825	0.9275	1.0020	0.9986
2019-02	1.0481	0.9755	1.0112	1.0118
2019-03	1.0470	0.9301	0.9868	0.9885
2019-04	1.0467	0.9478	0.9960	0.9974
2019-05	1.0507	0.9671	1.0080	1.0075
2019-06	1.1365	1.0133	1.0731	1.0724
2019-07	1.0775	1.0304	1.0537	1.0531
2019-08	1.0590	1.0168	1.0377	1.0378
2019-09	1.1775	0.9478	1.0564	1.0590
2019-10	1.0930	1.0184	1.0551	1.0545
2019-11	1.0417	0.9361	0.9875	0.9880
2019-12	1.0168	0.9863	1.0014	1.0022

Source: own elaboration based on data from the PriceIndices package in R.

The average values for the Laspeyres index are 1.0168, for the Paasche index 0.9863, for the Fisher index 1.0014, and for the Törnqvist index 1.0021. The most significant differences were noted in January and September between the Laspeyres index and the Paasche index, amounting to 0.1550 pp and 0.2298 pp.

6.2. Comparison of Selected Bilateral and Multilateral Indices

Multilateral methods are increasingly used when working with scanner data because they address problems such as chain drift. Despite this, some countries still use the chain Jevons index to develop price indices for scanner data, although it is not the optimal choice. Some statistical institutions also use chain Törnqvist indices (e.g., the USA and Japan), which are weighted formulas.

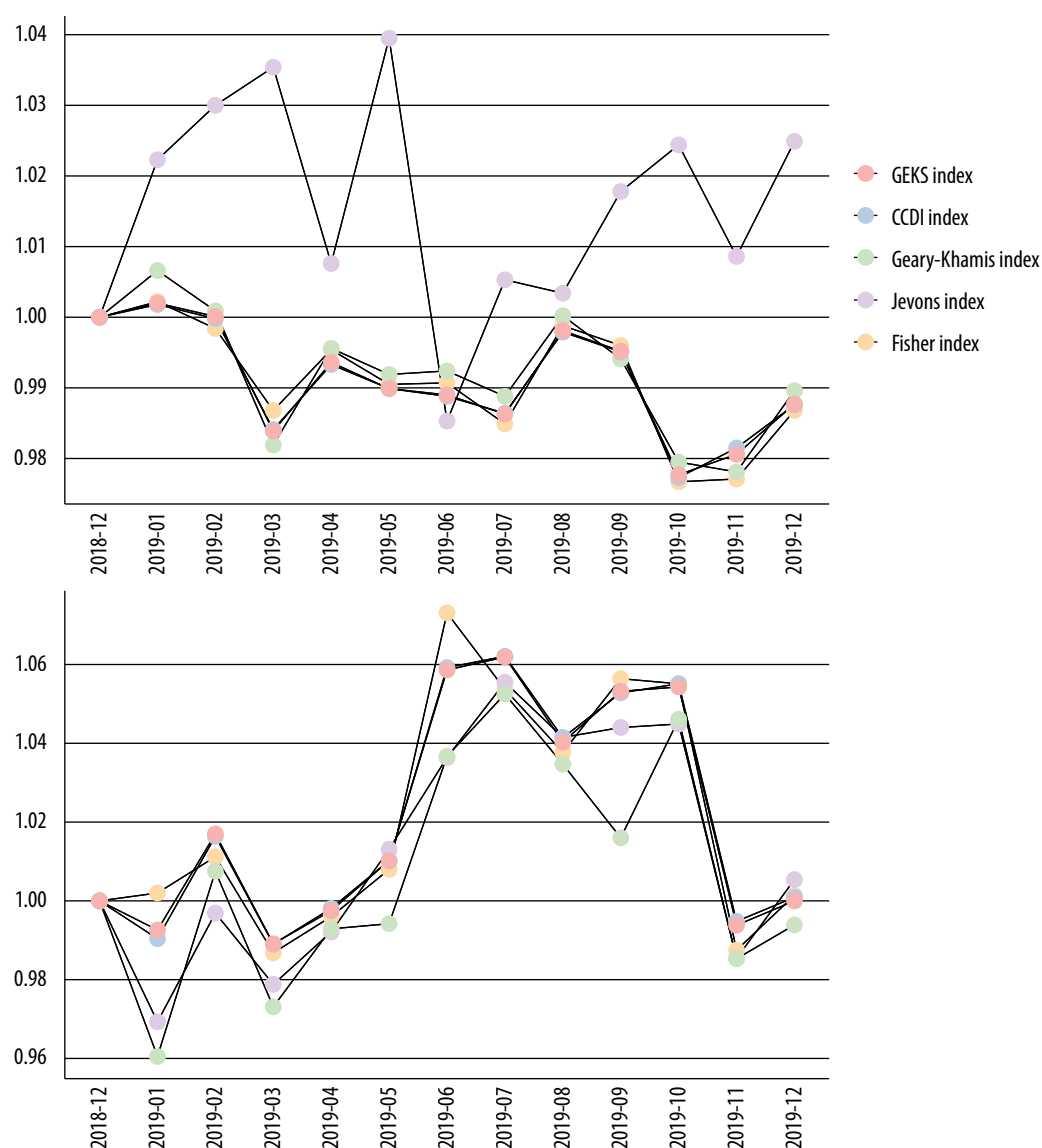


Figure 3. Comparison of bilateral and multilateral indices

Source: own elaboration based on data from the PriceIndices package.

As an alternative to bilateral indices, Figure 3 presents multilateral indices recommended for scanner data. Selecting the unweighted Jevons formula for the dataset results in significantly higher values than other indices. While its application to classically obtained prices at the lowest level of data aggregation is justified (as interviewers only note product prices), it is not the best choice for scanner data because it does not account for consumption volumes.

Table 6. Comparison of bilateral and multilateral indices for milk products

Data	GEKS index	CCDI index	Geary-Khamis index	Jevons index	Fisher index
2018-12	1.0000	1.0000	1.0000	1.0000	1.0000
2019-01	1.0020	1.0018	1.0066	1.0223	1.0022
2019-02	1.0001	0.9998	1.0009	1.0300	0.9984
2019-03	0.9839	0.9841	0.9819	1.0354	0.9868
2019-04	0.9936	0.9933	0.9956	1.0076	0.9954
2019-05	0.9899	0.9899	0.9919	1.0395	0.9905
2019-06	0.9890	0.9888	0.9924	0.9853	0.9907
2019-07	0.9863	0.9864	0.9888	1.0053	0.9849
2019-08	0.9981	0.9979	1.0002	1.0034	0.9988
2019-09	0.9952	0.9951	0.9941	1.0178	0.9960
2019-10	0.9777	0.9773	0.9795	1.0244	0.9767
2019-11	0.9806	0.9815	0.9781	1.0086	0.9771
2019-12	0.9877	0.9876	0.9896	1.0249	0.9868

Source: own elaboration based on data from the PriceIndices package in R.

When looking solely at the multilateral indices, the differences in values appear from the second or third decimal place. It is also not possible to clearly state which index consistently has the highest or lowest values. Adding bilateral indices to the analysis reveals that their values are higher compared to the multilateral indices, especially in the case of the Jevons index.

Table 7. Comparison of bilateral and multilateral indices for coffee products

Data	GEKS index	CCDI index	Geary-Khamis index	Jevons index	Fisher index
2018-12	1.0000	1.0000	1.0000	1.0000	1.0000
2019-01	0.9926	0.9904	0.9605	0.9693	1.0020
2019-02	1.0170	1.0164	1.0075	0.9969	1.0112
2019-03	0.9891	0.9890	0.9731	0.9788	0.9868
2019-04	0.9975	0.9981	0.9929	0.9921	0.9960
2019-05	1.0102	1.0102	0.9942	1.0131	1.0080
2019-06	1.0586	1.0592	1.0367	1.0364	1.0731

Data	GEKS index	CCDI index	Geary-Khamis index	Jevons index	Fisher index
2019-07	1.0619	1.0621	1.0525	1.0555	1.0537
2019-08	1.0403	1.0414	1.0347	1.0415	1.0377
2019-09	1.0532	1.0528	1.0160	1.0440	1.0564
2019-10	1.0543	1.0551	1.0461	1.0449	1.0551
2019-11	0.9938	0.9948	0.9853	0.9854	0.9875
2019-12	1.0000	1.0009	0.9939	1.0054	1.0014

Source: own elaboration based on data from the PriceIndices package in R.

In the case of the coffee dataset, no index stands out in terms of value, unlike the situation with the milk dataset. Here, the index values are more similar to each other. Among the multi-lateral indices, the Geary-Khamis index stands out slightly in comparison to the other two indices. When comparing the multilateral indices with the bilateral ones, the Fisher index stands out more, usually exhibiting the highest values.

6.3. Static and Dynamic Approaches

Figure 4 presents the comparison of selected bilateral indices (the Jevons index and the Carli index) and multilateral indices (the GEKS index and the CCDI index). They show how index values change after applying the low sales filter and the dump prices filter. The results are presented separately for the milk and coffee datasets.

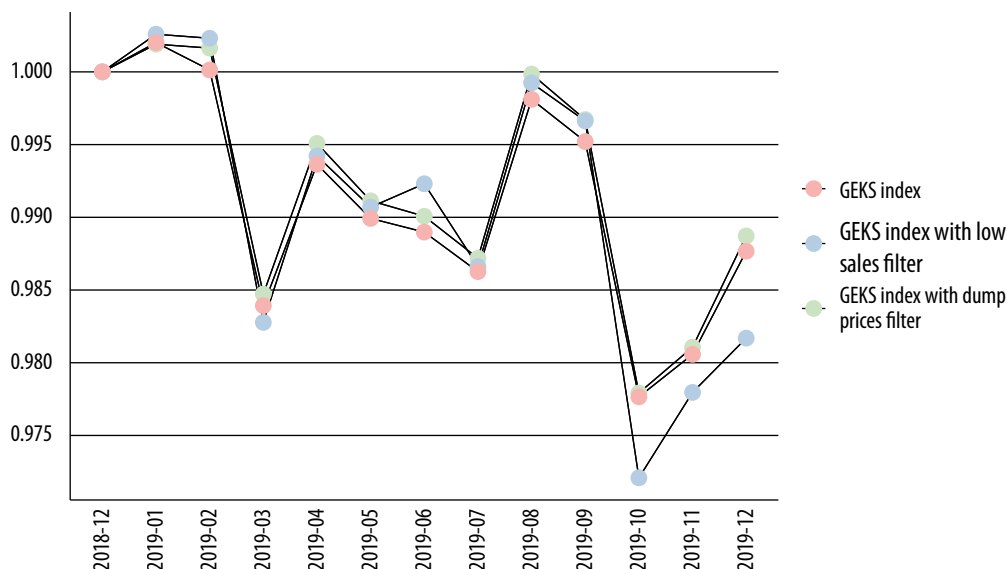


Figure 4. GEKS index values for the milk dataset depending on the applied data filter (low sales filter example)

Source: own elaboration based on data from the PriceIndices package in R.

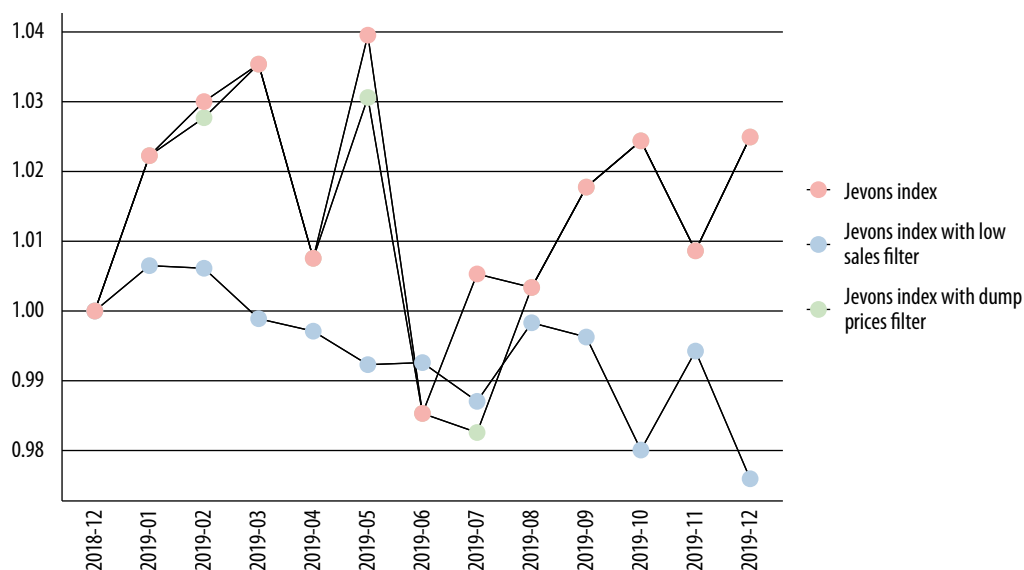


Figure 5. Jevons index values for the milk dataset depending on the applied data filter (low sales filter example)

Source: own elaboration based on data from the PriceIndices package in R.

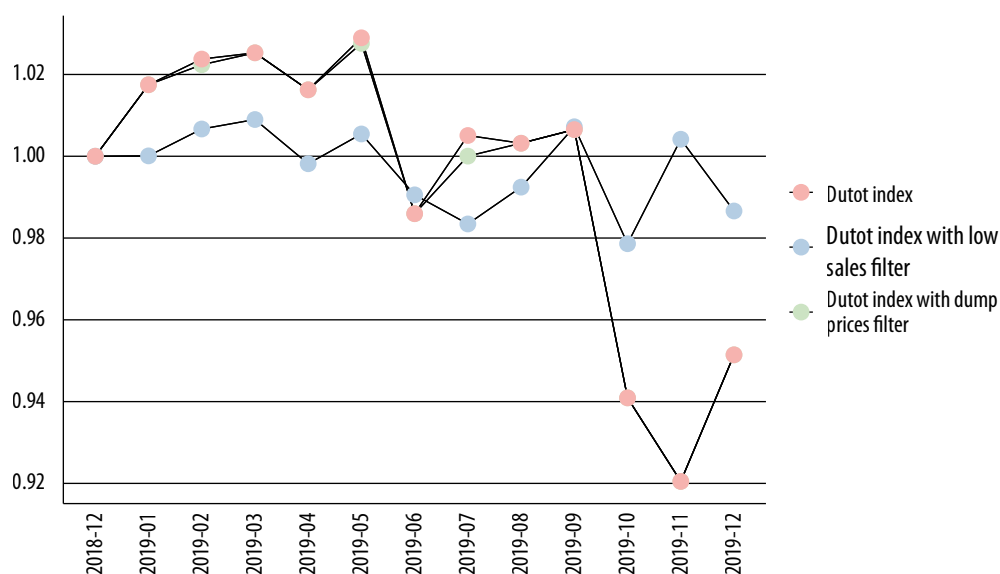


Figure 6. Dutot index values for the milk dataset depending on the applied data filter (low sales filter example)

Source: own elaboration based on data from the PriceIndices package in R.

For the milk dataset and across the three different indices (GEKS, Jevons, Dutot), the standard index and the index after applying the dump prices filter display similar trends. Conversely, the indices with the low sales filter applied tend to exhibit lower values. The multilateral GEKS index exhibits the least variability among the analysed differences. The differences between index variants are relatively small, with the low sales filter only significantly deviating from the other variants in a few instances. The Jevons index (Figure 5) shows greater

variability than GEKS, especially in certain periods (e.g., the first half of 2019). Clear differences between the indices can be seen, particularly for the version with the low sales filter which deviates from the index without filters. In the case of the Dutot index (Figure 6), the indexes without filters and with the dump prices filter record a decrease in value in 09–2019 while the index with the low sales filter over the entire observation remains relatively the same. In the other two cases of the above-mentioned indices, there is no such relationship.

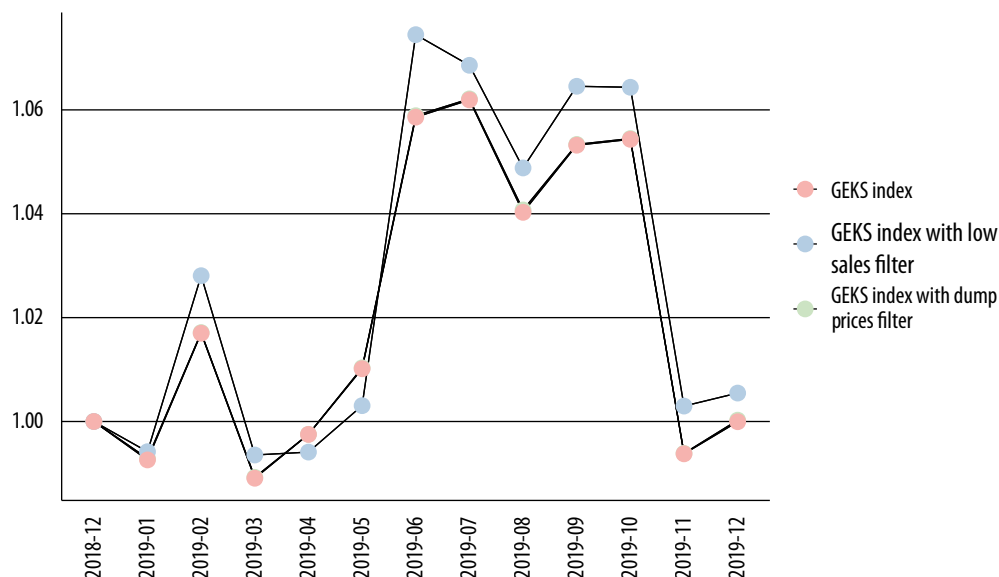


Figure 7. GEKS index values for the coffee dataset depending on the applied data filter (low sales filter example)

Source: own elaboration based on data from the PriceIndices package in R.

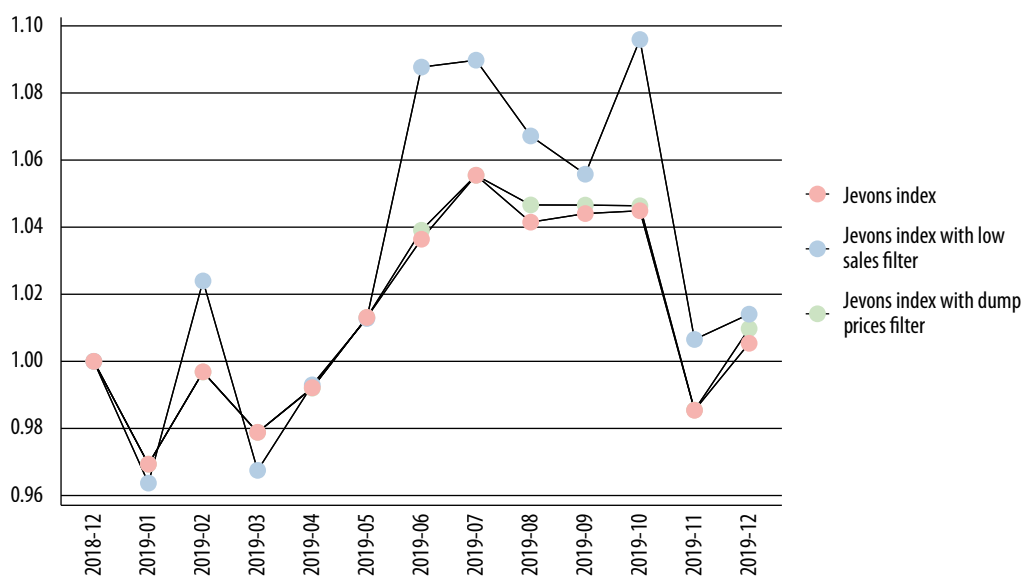


Figure 8. Jevons index values for the coffee dataset depending on the applied data filter (low sales filter example)

Source: own elaboration based on data from the PriceIndices package in R.



Figure 9. Dutot index values for the coffee dataset depending on the applied data filter (low sales filter example)

Source: own elaboration based on data from the PriceIndices package in R.

For the coffee dataset, three types of analysis are also presented: the indexes without applying filters, with the low sales filter and the dump prices filter for three indices (GEKS, Jevons, Dutot). In each case, the indices without filters and those after applying the dump prices filter follow a similar pattern, while the index with the low sales filter shows greater deviations. The Jevons index (Figure 8) and the Dutot index (Figure 9) exhibit high variability among the analysed methods. The index with the low sales filter shows stronger deviations, particularly in mid-2019, where it significantly exceeds the values of the other indices. The Dutot index (Figure 9) behaves similarly to the Jevons index, with moderate variability. The use of low sales filter also increases the index values. The multilateral GEKS index shows the smallest differences between the filtered and unfiltered variants. However, it is still evident that the low sales filter leads to higher index values, although this effect is less pronounced than in the case of the Jevons or Dutot indexes.

7. Conclusions

Scanner data hold enormous potential for refining and modernising inflation measurement. They provide information on consumption at the lowest level of aggregation, making the quality of scanner data incomparable to traditional datasets. However, it is important to recognise the many problems and challenges associated with using such data. As mentioned in Section 3, one of the main challenges is processing and preparing scanner data. The detailed level of information, the choice of time frame, and the large volume of data (typically ranging between 10,000 and 100,000 individual barcodes) require specialist knowledge in machine learning and statistics, as well as an appropriate IT environment for process automation.

This paper focused on comparing price indices and determining the scale of differences that arise from using individual indices. The choice of index formula is not straightforward, as the difference between indices for a single homogeneous product group can be significant. In the context of measuring inflation, such differences between the actual and measured price index are very dangerous and lead to serious financial consequences, such as excessive government expenditure indexing and taxes (Boskin et al., 1996). Other studies also show that at the elementary level, differences may result not only from the applied formula but also from the choice of data source (Białek, Panek, Zwierzchowski, 2022).

Multilateral indices, characterised by transitivity and the absence of chain drift, are a rational choice for scanner data. To formulate a recommendation for the most appropriate price index for scanner data, further research on index properties – in both theoretical and empirical terms – is needed, along with analysing the experiences of countries that already use this methodology.

The results obtained from scanner data analysis provide several important insights. Firstly, they confirm relationships that are already well-established in traditional price collection: superlative indices such as Fisher and Törnqvist tend to approximate each other, while the Laspeyres index systematically overestimates inflation and the Paasche index underestimates it. Secondly, however, the divergence between the Laspeyres and Paasche indexes is notably larger in scanner data than in traditional price collection. In conventional surveys, this difference in Poland, at the COICOP 2 level, usually does not exceed a few promilles. In contrast, when working with scanner data at a very low level of aggregation (barcode level), the differences between bilateral indices become significant, which clearly illustrates the substitution effect between goods. This finding highlights the methodological importance of carefully selecting the appropriate index formula in scanner data applications.

When applying the low sales and dump prices filters, it is evident that bilateral indices show greater variation in values compared to multilateral indices. In both datasets, the low sales filter deviated from the index without filters, while the dump prices filter displayed similar values.

The choice between the static and dynamic approaches depends on the characteristics of the data, their availability, and the analytical objectives. The static approach aligns more closely with traditional methodologies and is advantageous when integrating scanner data with conventional price collection. In contrast, the dynamic approach offers greater flexibility and makes better use of large datasets, particularly in rapidly changing markets.

References

- Białek J. (2020a), *Remarks on Price Index Methods for the CPI Measurement Using Scanner Data*, "Statistics and Economy Journal Statistika", vol. 100(1), pp. 55–70.
- Białek J. (2020b), *Wykorzystanie danych skanowanych do pomiaru inflacji – doświadczenia międzynarodowe i wyzwania metodologiczne*, "Wiadomości Statystyczne. The Polish Statistician", vol. 65(01), pp. 9–33.
- Białek J. (2021), *PriceIndices – a New R Package for Bilateral and Multilateral Price Index Calculations*, "Statistika – Statistics and Economy Journal", vol. 101(2), pp. 121–141.

- Białek J., Beręsewicz M. (2021), *Scanner data in inflation measurement: From raw data to price indices*, "Statistical Journal of the IAOS", vol. 37(4), pp. 1315–1336
- Białek J., Panek T., Zwierzchowski J. (2022), *Assessing the effect of new data sources on the consumer price index: a deterministic approach to uncertainty and sensitivity*, "Statistics in Transition New Series", vol. 23(3), pp. 1–25.
- Boskin M.J., Dulberger E.R., Gordon R.J., Griliches Z., Jorgenson D.W. (1996), *Toward a More Accurate Measure of the Cost of Living. Final Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index*, <https://www.ssa.gov/history/reports/boskinrpt.html> [accessed: 18.12.2025].
- Carli G.R. (1804), *Del valore e della proporzione de' metalli monetati*, "Scrittori Classici Italiani di Economia Politica", vol. 13, pp. 297–336.
- Caves D.W., Christensen L.R., Diewert W.E. (1982), *Multilateral comparisons of output, input, and productivity using superlative index numbers*, "Economic Journal", vol. 92(365), pp. 73–86.
- Chessa A. (2018), *Product definition and index calculation with MARS-QU: Applications to consumer electronics*, Statistics Netherlands.
- Chessa A. (2021), *A comparison of index extension methods for multilateral methods*, https://eventos.fgv.br/sites/eventos.fgv.br/files/arquivos/u161/index_extension_methods_compared_chessa_og19.pdf [accessed: 18.12.2025].
- Diewert W.E. (1976), *Exact and superlative index numbers*, "Journal of Econometrics", vol. 4(2), pp. 115–145.
- Diewert W.E. (2022), *Scanner data, elementary price indexes and the chain drift problem*, [in:] K.J. Fox, E.R. Lawrence (eds.), *Advances in Economic Measurement: A Volume in Honour of DS Prasada Rao*, Palgrave Macmillan, Singapore, pp. 445–606.
- Dubois P., Griffith R., O'Connell M. (2022), *The use of scanner data for economics research*, "Annual Review of Economics", vol. 14(1), pp. 723–745.
- Dutot C. (1738), *Reflexions Politiques sur les Finances et le Commerce*, Les freres Vaillant et Nicolas Prevost, The Hague.
- Éltető O., Köves P. (1964), *On a problem of index number computation relating to international comparison*, "Statistikai Szemle", vol. 42(10), pp. 507–518.
- Eurostat (2017), *Harmonised Index of Consumer Prices. Practical Guide for Processing Supermarket Scanner Data*, <https://circabc.europa.eu/sd/a/8e1333df-ca16-40fc-bc6a-1ce1be37247c/practical-guide-supermarket-scanner-data-september-2017.pdf> [accessed: 19.12.2025].
- Eurostat (2020a), *Consumer Price Index Manual – Concepts and Methods*, https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@dgreports/@stat/documents/publication/wcms_761444.pdf [accessed: 20.12.2025].
- Eurostat (2020b), *Guide on multilateral methods in the Harmonised Index on Consumer Prices (HICP) – 2022 edition*, <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-21-020> [accessed: 20.12.2025].
- Eurostat (2024), *Harmonised Index of Consumer Prices (HICP) Methodological manual – 2024 edition*, <https://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/w/ks-gq-24-003> [accessed: 21.12.2025].
- Fisher I. (1922), *The making of index numbers: a study of their varieties, tests, and reliability*, Houghton Mifflin, Boston–New York.
- Geary R.C. (1958), *A note on the comparison of exchange rates and purchasing power between countries*, "Journal of the Royal Statistical Society. Series A (General)", vol. 121(1), pp. 97–99.
- Gini C. (1931), *On the circular test of index numbers*, "Metron", vol. 9(9), pp. 3–24.

- Hałka A., Leszczyńska A. (2011), *Wady i zalety wskaźnika cen towarów i usług konsumpcyjnych – szacunki obciążenia dla Polski*, „Gospodarka Narodowa”, no. 9, pp. 51–75.
- Jevons W.S. (1865), *On the variation of prices and the value of the currency since 1782*, “Journal of the Statistical Society of London”, vol. 28(2) pp. 294–320.
- Khamis S.H. (1972), *A new system of index numbers for national and international purposes*, “Journal of the Royal Statistical Society: Series A (General)”, vol. 135(1), pp. 96–121.
- Laspeyres K. (1871), *ix. die berechnung einer mittleren waarenpreissteigerung*, “Jahrbücher für Nationalökonomie und Statistik”, vol. 16(1), pp. 296–318.
- Paasche H. (1874), *Über die Preisentwicklung der letzten Jahre nach den Hamburger Börsennotirungen*, “Jahrbücher für Nationalökonomie und Statistik”, vol. 12, pp. 168–178.
- Statistics Poland (2019), *Co warto wiedzieć o inflacji?*, Główny Urząd Statystyczny, Warszawa, https://stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultaktualnosci/5464/18/1/1/co_warto_wiedzi ec_o_inflacji.pdf [accessed: 21.12.2025].
- Statistics Poland (2024), *Co warto wiedzieć o inflacji? cz. II*, Główny Urząd Statystyczny, Warszawa, https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5464/18/2/1/co_warto_wiedzi ec_o_inflacji_cz.ii.pdf [accessed: 21.12.2025].
- Törnqvist L. (1936), *The bank of Finland's consumption price index*, “Bank of Finland Monthly Bulletin”, vol. 10, pp. 1–8.
- Van Loon K., Roels D. (2018), *Integrating big data in the Belgian CPI. Meeting of the Group of Experts on Consumer Price Indices*, UNECE, Geneva, <https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2018/Belgium.pdf> [accessed: 21.12.2025].
- Vartia Y., Suoperä A., Nieminen K., Montonen K. (2018), *Circular Error in Price Index Numbers Based on Scanner Data. Preliminary Interpretations*, https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2018/Finland_main.pdf [accessed: 29.12.2025].

Zastosowanie danych skanowanych do obliczania inflacji: na przykładzie Polski

Streszczenie:

W kontekście rosnącego wykorzystania danych skanowanych w pomiarze inflacji kluczowym wyzwaniem metodologicznym pozostaje wybór odpowiedniej formuły indeksu cenowego oraz metody filtrowania danych. Celem artykułu jest zbadanie wpływu tych wyborów na wyniki pomiaru dynamiki cen. W pracy przeanalizowano dane transakcyjne zebrane z terminali punktów sprzedaży w polskich supermarketach dla wybranych grup produktów: kawy oraz mleka. Porównano różne formuły indeksów cenowych – w tym bilateralne nieważone (Dutot, Carli, Jevons), bilateralne ważne (Törnqvist, Fisher) oraz multilateralne (Geary-Khamis, CCDI) – oraz przetestowano wpływ różnych metod filtrowania danych skanerowych. Wyniki pokazują, że zarówno wybór formuły indeksu cenowego, jak i zastosowana metoda filtrowania danych skanerowych mają mierzalny i istotny wpływ na ocenę dynamiki cen. Różnice te mogą prowadzić do odmiennych interpretacji inflacji, co podkreśla potrzebę świadomego i uzasadnionego wyboru metodologicznego przy wykorzystaniu danych skanerowych w statystyce cen.

Słowa kluczowe:

dane skanowane, indeksy multilateralne, indeksy bilateralne

Artykuł 2

Artykuł 2 Proposition of the quality adjusted GEKS-type price index

opublikowany w

Statistika: Statistics & Economy Journal, Vol 105, 2025

Wkład współautorów:

Natalia Pawelec: 80%

Jacek Białek: 20%

Wkład własny:

Wkład własny autorki obejmuje dominującą część pracy badawczej i twórczej związanej z powstaniem artykułu. Autorka ponosi odpowiedzialność za przeprowadzenie przeglądu literatury dotyczącego indeksów multilateralnych i jakościowo skorygowanych indeksów typu GEKS. Autorka samodzielnie przeprowadziła całość analizy empirycznej: przygotowanie i filtrowanie zbiorów danych skanowanych (kategorie mleko i kawa), implementację obliczeń w środowisku R z wykorzystaniem pakietu *PriceIndices*, obliczenie i porównanie wartości indeksów dla dwóch poziomów agregacji (GTIN i COICOP 6), a także opracowanie tabel i rysunków prezentujących wyniki.

Proposition of the Quality Adjusted Geks-Type Price Index

Jacek Białek¹ | *University of Lodz, Lodz, Poland*

Natalia Pawelec² | *University of Lodz, Lodz, Poland*

Received 21.10.2024, Accepted (reviewed) 14.11.2024, Published 12.9.2025

Abstract

Scanner data from retailers like supermarkets, electronics stores, and online shops provides detailed transaction information at the barcode level (e.g., GTIN, EAN), allowing for the use of various price index formulas, including weighted ones. Due to high product turnover and seasonality, multilateral index methods are ideal, as they use a whole-time window and are transitive, avoiding chain drift. However, most commonly used multilateral indices (e.g., GEKS, CCDI, GK, TPD) fail the identity test, which requires the index to return to one when prices revert to their original levels. This paper proposes a new multilateral index inspired by GEKS but incorporating quality adjustments like the Geary-Khamis method. The index satisfies the identity test and other key axioms, demonstrating its robustness. Comparisons with the SPQ index and quality-adjusted indices (GEKS-AQU, GEKS-AQI) confirm its effectiveness, making it a highly useful tool for scanner data analysis in both theory and practice.

Keywords

Scanner data, multilateral indices, the GEKS-index

DOI

<https://doi.org/10.54694/stat.2024.71>

JEL code

C43

INTRODUCTION

The term “scanner data” refers to transaction data that specify turnover and numbers of items sold using a Global Trade Article Number (GTIN) code, European Article Number (EAN) code, Stock Keeping Unit code (SKU) or other barcode. Scanner data have numerous advantages over traditional survey data; for example, such data sets are much bigger than traditional ones and contain complete transaction records, i.e. information about prices and quantities, along with additional information about products, including the grammage, unit, label with description, and VAT. In other words, scanner data contain full information at the most detailed item level. Since the information about prices and quantities at the lowest COICOP (Classification of Individual Consumption by Purpose) level is available in scanner

¹ Department of Statistical Methods, University of Lodz, Lodz, Poland. E-mail: jacek.bialek@uni.lodz.pl. Also Statistics Poland, Department of Trade and Services, Warsaw, Poland. ORCID: <<https://orcid.org/0000-0002-0952-5327>>.

² Department of Economic Mechanisms, University of Lodz, Rewolucji 1905r 41/43, Lodz, 90-214, Poland. E-mail: natalia.pawelec@uni.lodz.pl. ORCID: <<https://orcid.org/0009-0008-5289-6204>>.

data, statistical offices can use any price index formula at this level of aggregation, including weighted formulas. The use of scanner data when assessing inflation improves the data collection process, reduces costs and better reflects changes that occur in consumer behaviour. The form of a sample scanner dataset is presented in Table 1.

Table 1 Sample scanner data frame obtained from a retail chain in Poland

Time	Prices	Quantities	retID	EAN	Label	Grammage	Unit	prodID
2022-12-31	10.47	8.48	26-617	590674717126	long grain rice	0.4	kg	1
2022-12-31	12.47	5.87	40-772	590674717126	long grain rice	0.4	kg	1
2022-12-31	11.40	15.65	70-001	590674717126	long grain rice	0.4	kg	1
2022-12-31	13.20	16.95	85-791	590674717126	long grain rice	0.4	kg	1
2022-12-31	11.47	85.41	01-460	590674717126	long grain rice	0.4	kg	1
2022-12-31	11.97	7.82	05-820	590674717126	long grain rice	0.4	kg	1

Source: Białek et al. (2022: 71)

Nevertheless, the choice of price index when using scanner data to measure inflation is not straightforward. More than 40% of EU countries have implemented or are in the process of implementing scanner data in their Consumer Price Index (CPI) calculations, but there are still no clear recommendations on the choice of price index formula. While there are guides and guidelines available for choosing this index (Eurostat, 2020), and we understand the advantages and disadvantages of each approach (ABS, 2016), European Union (EU) countries still differ in the final index formulas they use. The topic of price index selection will be discussed in more detail in Section 1 but let us just mention here that this paper proposes another multilateral index, the design of which is related to the GEKS and Geary-Khamis methods. Indeed, the main purpose of the paper is to propose a new multilateral price index and to study its properties, which on the one hand has a structure similar to the GEKS index, on the other hand it requires quality adjusting, which makes it similar to the Geary-Khamis method.

The paper is structured as follows: Section 1 reviews the literature on selecting a price index for scanner data, examining available price index formulas and key selection criteria. Section 2 discusses the concept of GEKS-type indices and quality-adjusted GEKS-type indices (GEKS-AQU, GEKS-AQI). Subsection 2.3 proposes a new type of index, the GEKS-GAQI index. Section 3 presents the results of an empirical comparison of the proposed index with popular multilateral indices. The last section provides the main conclusions of the study.

1 LITERATURE REVIEW

In general, the procedure of scanner data processing consists of the following steps: data cleaning ► extracting information from the product description ► classification of products into COICOP groups ► product matching over time ► product filtering ► data standardization ► imputing missing prices (optional) ► price index calculations. It is not the purpose of the paper to discuss these individual stages, but the interested reader is referred, for example, to Białek and Beręsewicz (2021). Similarly, although the construction of an IT environment for the procedure of scanner data is one of the major challenges, we will not devote attention to it but only mention that all the analysis was done in the *PriceIndices* R package (Białek, 2022). The paper focuses on the final stage of the scanner data procedure, namely the determination of the correct price index.

1.1 Criteria for selecting a price index

The three main approaches in evaluating (but also constructing) a price index are: the axiomatic approach, the economic approach and the stochastic approach (von der Lippe, 2007).

According to the axiomatic approach, desirable index properties (the so called “tests”) are defined that an index may, or may not satisfy. A list of tests dedicated to bilateral indices, i.e., taking into account only the *current* and *base* period, can be found in Balk (1995) and von der Lippe (2007). A list of tests dedicated to multilateral indices, i.e., taking into account not only the *current* and *base* period but also all intermediate periods, can be found in ABS (2016), Eurostat (2018), and Białek (2023). Due to the high dynamics of scanner data, characterised by product churn, numerous appearances of new goods in supermarkets or the disappearance of goods previously sold, as well as the seasonality of products, multilateral indices are particularly recommended for determining changes in scanned prices (Eurostat, 2018). Thus, in the remainder of this paper, we will consider primarily multilateral indices. A key advantage of multilateral indices operating on the time window $[0, T]$ is their *transitivity*, which means that for any period $0 \leq s \leq t \leq T$, it holds that $P^{0,s}P^{s,t} = P^{0,t}$, where $P^{a,b}$ denotes a multilateral index that compares research period b to the base period a . By definition, *transitivity* eliminates the chain drift problem often associated with scanner data. Chain drift can be formalized as a violation of the *multi-period identity test*. This test expects that when all prices and quantities in the current period revert to their values from the base period, the index should indicate no price change and equal one. Thus, multilateral indices are free from chain drift within a given estimation time window. At this point it should be mentioned that one of the more restrictive tests against multilateral indices is the identity test, which is even a basic requirement for bilateral indices. The widely recognized and popular multilateral indices (GEKS, TPD or Geary-Khamis) do not meet this requirement. However, the identity test is a fairly natural axiom, since it expects the index to have a value of one when prices (for example, of seasonal goods) return to their initial values. In the following section, we will show that the proposed multilateral index formula satisfies the identity test.

The economic approach to the price index theory assumes that quantities are functions of prices (CPI Manual, 2004). This approach assumes that households optimise their purchase by maximising their utility under a budget constraint or by minimizing the cost of purchases for a given utility level. Multilateral methods involve different assumptions about the functional form of the utility function. For instance, the GEKS index is exact for a “flexible” functional form, i.e. it expresses the price differences experienced by optimising consumers without imposing restrictive assumptions about how they can substitute between products.

The main assumption of the stochastic approach is that prices or price changes can be estimated using an econometric model. The index formula is constructed based on appropriately estimated parameters from a panel regression equation, as is done at least for the TPD index (see Section 1.2). However, in the remainder of this paper, we will focus on the axiomatic approach, with particular attention to the identity test and the transitivity property.

1.2 Price indices considered for scanner data

As it was above-mentioned, the most promising group of indices in the context of scanner data appear to be multilateral indices. The multilateral price index is calculated for a given time window consisting of $T + 1$ consecutive months, which we number $0, 1, 2, \dots, T$ (typically $T = 12$ or $T = 24$). Well-known and widely used multilateral methods include the GEKS method (Gini, 1931; Eltetö and Köves, 1964), the Geary-Khamis (GK) method (Geary, 1958; Khamis, 1972), the CCDI method (Caves et al., 1982) or the Time Product Dummy Methods (de Haan and Krsinich, 2018). Other GEKS-type indices can also be found in the literature (e.g., GEKS-W or GEKS-L indices, cf. Section 2.1), or indices based on hedonic regression that explain price behaviour (e.g., Time Dummy Hedonic index, TDH). In what is known

as a dynamic approach, where the activity prior to calculating the index is to filter out relatively low sales, some EU countries determine the Jevons (1865) chain index. In this paper, however, we will limit our comparative analysis to multilateral indices only.

2 METHODS

2.1 The idea of GEKS-type indices

The popular GEKS method which is based on the Fisher price index and operates over the time interval $[0, T]$, can be expressed as follows:

$$P_{GEKS}^{0,t} = \prod_{\tau=0}^T \left(\frac{P_F^{\tau,t}}{P_F^{\tau,0}} \right)^{\frac{1}{T+1}}, \quad (1)$$

where $P_F^{a,b}$ denotes the Fisher (1922) price index that compares the research (current) period b to the base period a . As a rule, it is assumed that the base index in the GEKS formula is any *superlative* index (CPI Manual, 2004), which is also the Fisher index.

Accordingly, the literature also considers a GEKS index based on the superlative Törnqvist (1936) index and a GEKS index based on the superlative Walsh (1905) formula, which we will denote here as CCDI and GEKS-W, respectively. It is important to note that the GEKS, CCDI and GEKS-W indices do not satisfy the *identity test* (Eurostat, 2018).

Two new GEKS-type price indices have recently appeared in the literature that have broken the canons associated with assumptions towards the underlying index, located in the body of the GEKS formula. These GEKS-type indices are not based on a superlative price index at all, nor on an index that meets the *timere versal test*. The former (GEKS-L) is based on the Laspeyres (1871) price index, and the latter is based on the geometric Laspeyres index (CPI Manual, 2004). Please note that the GEKS-L and GEKS-GL indices satisfy the identity test (Bialek, 2023b).

2.2. Quality-adjusted GEKS-type indices

Before we provide definitions of quality-adjusted GEKS-type indices, let us denote the sets of homogeneous products that belong to the same product group in months 0 and t as G_0 and G_t respectively, and let $G_{0,t}$ denote a set of matched products available in both periods 0 and t . Let p_i^τ and q_i^τ denote the price and quantity of the i -th product at the time τ and let $N_{0,t}$ be the number of elements of set $G_{0,t}$.

In the unit value concept, prices of homogeneous products are equal to the ratio of expenditure and quantity sold (Chessa et al., 2017). However, quantities of different products cannot be added together as in the case of homogeneous products. That is why the idea of quality-adjusted unit value assumes that prices p_i^τ of different products $i \in G_t$ in month τ are transformed into “quality-adjusted prices” p_i^τ / v_i and quantities q_i^τ are converted into “common units” $v_i q_i^\tau$ using a set of factors $v = \{v_i : i \in G_t\}$. Thus, the “classical” quality-adjusted unit value $QUV_{G_t}^\tau$ of a set of products G_t in month τ can be expressed as follows:

$$QUV_{G_t}^\tau = \frac{\sum_{i \in G_t} q_i^\tau p_i^\tau}{\sum_{i \in G_t} v_i q_i^\tau}, \quad (2)$$

and it is used in the definition of the Geary-Khamis (GK) index (Chessa et al., 2017).

In practice, consumer responses to price changes can be delayed or even accelerated as consumers not only react to current price changes but also use their own “forecasts” or concerns about future price increases. For example, the consumption of thermophilic (seasonal) fruit tends to be higher in summer because it is cheaper than in winter, when the season is almost over. For instance, some interesting study

on “unconventional” consumer behaviour, such as stocking up and delayed quantity responses to price changes, and its impact on chain drift bias can be found in the paper by von Auer (2019). Since in practice we often observe prices and quantities that are not perfectly synchronised in time, the following form of the “asynchronous quality-adjusted unit value” was proposed (Białek, 2023a):

$$QUV_{G_\tau}^\tau = \frac{\sum_{i \in G_\tau} q_i^\tau p_i^\tau}{\sum_{i \in G_\tau} v_i q_i^\tau}, \quad (3)$$

which has more general form than (2) since we have $AQUV_{G_{\tau,\tau}}^{\tau,\tau} = QUV_{G_\tau}^\tau$. Let us define function $P^{\tau,s}(v, q^\tau, p^\tau, p^s)$ as follows:

$$P^{\tau,s}(v, q^\tau, p^\tau, p^s) = \frac{AQUV_{G_{\tau,s}}^{\tau,s}}{AQUV_{G_{\tau,\tau}}^{\tau,\tau}}. \quad (4)$$

Substituting Formula (4) into the GEKS formula and adopting the system of factors $v = \{v_i : i \in G_\tau\}$ corresponding to the augmented Lehr index (Lamboray, 2017; van Loon and Roels, 2018):

$$v_i = \frac{\sum_{\tau=0}^T p_i^\tau q_i^\tau}{\sum_{\tau=0}^T q_i^\tau}, \quad (5)$$

we obtain the GEKS-AQU index (Białek, 2023):

$$P_{GEKS-AQU}^{0,t} = \prod_{\tau=0}^T \left(\frac{P^{\tau,t}(v, q^\tau, p^\tau, p^t)}{P^{\tau,0}(v, q^\tau, p^\tau, p^0)} \right)^{\frac{1}{T+1}} = \prod_{\tau=0}^T \left(\frac{AQUV_{G_{\tau,t}}^{\tau,t}}{AQUV_{G_{\tau,0}}^{\tau,0}} \right)^{\frac{1}{T+1}}. \quad (6)$$

Białek (2023a) also considers the “asynchronous quality-adjusted price index” (AQI) which can be written as follows:

$$AQI_{G_{\tau,s}}^{\tau,s} = \frac{\sum_{i \in G_{\tau,s}} v_i q_i^\tau \frac{p_i^s}{p_i^\tau}}{\sum_{i \in G_{\tau,s}} v_i q_i^\tau}. \quad (7)$$

It can be treated as a weighted arithmetic mean of partial indices p_i^s / p_i^τ , where the weights are proportional to the relative share of each product's quality-adjusted quantities q_i^τ in the sum of all quality-adjusted quantities. Using AQI instead of AQUV in formula (6), we obtain the GEKS-AQI index (Białek, 2023a):

$$P_{GEKS-AQI}^{0,t} = \prod_{\tau=0}^T \left(\frac{AQI_{G_{\tau,t}}^{\tau,t}}{AQI_{G_{\tau,0}}^{\tau,0}} \right)^{\frac{1}{T+1}}. \quad (8)$$

Please note that the GEKS-AQU and GEKS-AQI indices meet transitivity and the identity test (Białek, 2023a).

2.3. The new quality-adjusted GEKS-type price index

Please note that the AQI defined in (7) can be viewed as a weighted version of the Carli (1804) price index. Currently, there is a trend in price statistics toward shifting from the Carli index to the Jevons (1865) index, which replaces the arithmetic mean of the sub-indices with the geometric mean. This shift serves as the foundation for our further considerations. Let us therefore define “the geometric version of asynchronous quality-adjusted price index” (GAQI) follows:

$$GAQI_{G_{\tau,s}}^{\tau,s} = \prod_{i \in G_{\tau,s}} \left(\frac{p_i}{p_i^\tau} \right)^{\frac{\sum_{k \in G_{\tau,s}} v_k q_k^\tau}{v_i q_i^\tau}}. \quad (9)$$

Using GAQI instead of AQUV in Formula (6), we obtain the proposed GEKS-GAQI index, i.e.

$$P_{GEKS-GAQI}^{0,t} = \prod_{\tau=0}^T \left(\frac{GAQI_{G_{\tau,t}}^{\tau,t}}{GAQI_{G_{\tau,0}}^{\tau,0}} \right)^{\frac{1}{T+1}}. \quad (10)$$

The properties of the index (10) will correspond to those of the GEKS-AQI and GEKS-AQU indices (Białek, 2023). It is noteworthy that the GEKS-GAQI index satisfies transitivity:

$$\begin{aligned} P_{GEKS-GAQI}^{0,s} P_{GEKS-GAQI}^{s,t} &= \prod_{\tau=0}^T \left(\frac{GAQI_{G_{\tau,s}}^{\tau,s}}{GAQI_{G_{\tau,0}}^{\tau,0}} \right)^{\frac{1}{T+1}} \prod_{\tau=0}^T \left(\frac{GAQI_{G_{\tau,t}}^{\tau,t}}{GAQI_{G_{\tau,s}}^{\tau,s}} \right)^{\frac{1}{T+1}} = \\ &= \prod_{\tau=0}^T \left(\frac{GAQI_{G_{\tau,t}}^{\tau,t}}{GAQI_{G_{\tau,0}}^{\tau,0}} \right)^{\frac{1}{T+1}} = P_{GEKS-GAQI}^{0,t}. \end{aligned} \quad (11)$$

We will now show that the index (10) satisfies the identity test. Let us assume that $G_0 = G_t$ and that for any $i \in G_{0,t}$ we have $p_i^0 = p_i^t$. From this assumption, we can conclude that $G_{\tau,0} = G_{\tau,t}$ for any $\tau \in \{0, 1, 2, \dots, T\}$. Consequently, we obtain:

$$GAQI_{G_{\tau,t}}^{\tau,t} = \prod_{i \in G_{\tau,0}} \left(\frac{p_i}{p_i^\tau} \right)^{\frac{\sum_{k \in G_{\tau,0}} v_k q_k^\tau}{v_i q_i^\tau}} = GAQI_{G_{\tau,0}}^{\tau,0}. \quad (12)$$

From (10) and (12) we have:

$$P_{GEKS-GAQI}^{0,t} = \prod_{\tau=0}^T \left(\frac{GAQI_{G_{\tau,0}}^{\tau,0}}{GAQI_{G_{\tau,0}}^{\tau,0}} \right)^{\frac{1}{T+1}} = 1,$$

which means that the GEKS-GAQI index satisfies the identity test.

3 EMPIRICAL STUDY

In this empirical study we use scanner data included in the R package *PriceIndices* (Białek, 2022). Two datasets are used: *milk* and *coffee*. The *milk* dataset contains scanner data on milk sales from a Polish supermarket, covering the period from December 2018 to December 2019. It consists of a data frame with 6 columns and 4 386 rows. The *coffee* dataset includes scanner data on coffee sales in a Polish supermarket

from December 2018 to December 2019. Both datasets contain the following columns: *time* – dates of transactions (Year-Month-Day); *prices* – prices of sold products (PLN); *quantities* – quantities of sold products; *prodID* – unique product codes obtained after product matching (i.e. the *milk* dataset contains 68 different product IDs, and the *coffee* dataset contains 79 different product IDs); *retID* – unique codes identifying outlets/retailer sale points (the *milk* dataset contains 5 different retIDs, the *coffee* dataset contains 20 different retIDs); *description* – descriptions of sold milk/coffee products (the *milk* dataset contains 6 different product descriptions corresponding to subgroups of the milk group, the *coffee* dataset contains 3 different product descriptions corresponding to subgroups of the coffee group).

3.1 Comparing index values for *milk* and *coffee* harvests

First, we conducted a comparison of index values separately for the *milk* and *coffee* datasets. In the analysis, the indices were compared in two variants: at the lowest level of data aggregation (i.e. at the GTIN code) and at a broader level, using the COICOP 6 data aggregation.

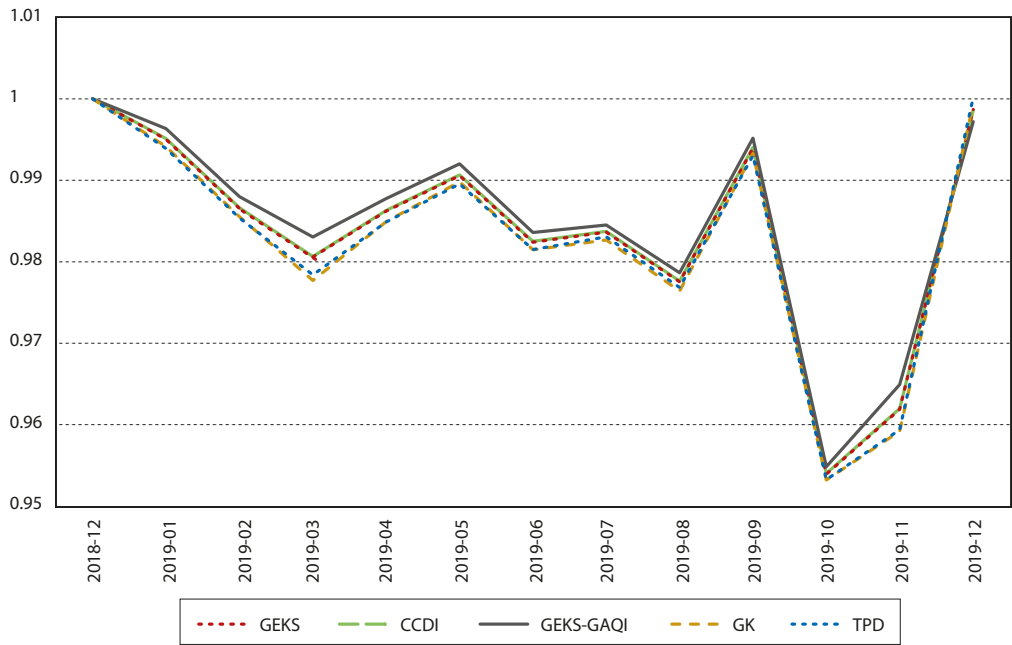
A comparison of the GEKS, the CCDI, Geary-Khamis (GK) and TPD indices for the *milk* dataset is shown in Figure 1 (at the GTIN level) and Figure 2 (at the COICOP 6 level). A comparison of the GEKS-L, GEKS-GL, GEKS-AQI, GEKS-AQU, and GEKS-GAQI indices for the *milk* dataset is shown in Figure 3 (at the GTIN level) and Figure 4 (at the COICOP 6 level). A similar analysis was conducted for the *coffee* dataset: Figure 5 and Figure 6 compare the GEKS, CCDI, GK, and TPD indices at the GTIN level, while Figure 7 and Figure 8 compare the GEKS-L, GEKS-GL, GEKS-AQI, GEKS-AQU, and GEKS-GAQI indices at the COICOP 6 data aggregation level. In addition, exact index values are presented for each dataset in Table 1 and Table 2.

Figure 1 Comparison of the GEKS, CCDI, GEKS-GAQI, TPD indices for the *milk* dataset at the GTIN level



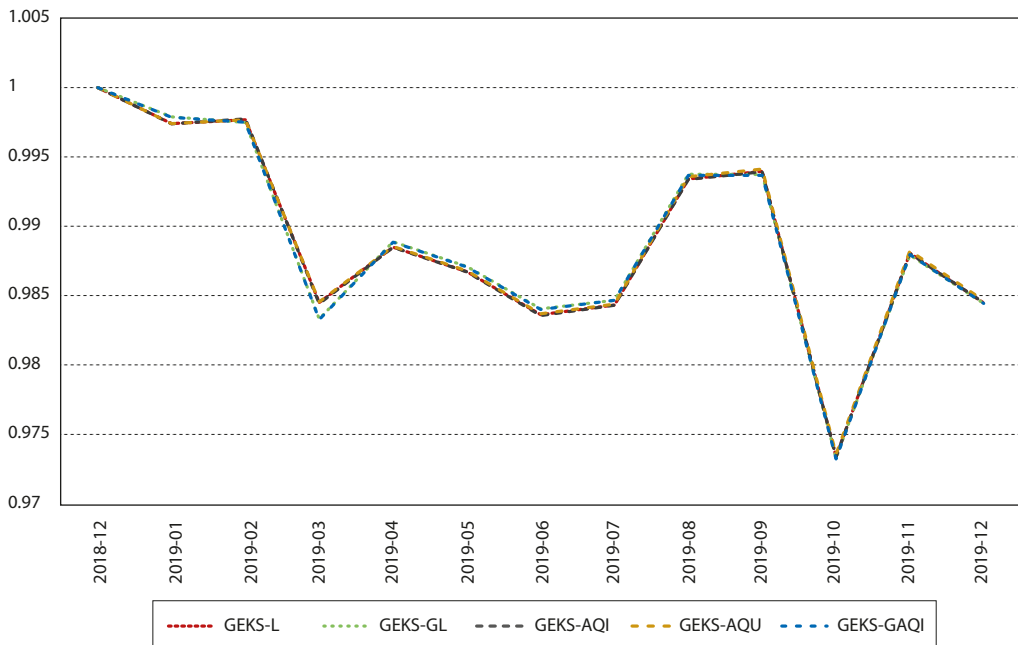
Source: Authorial computation

Figure 2 Comparison of the GEKS, CCDI, GEKS-GAQI, TPD indices for the *milk* dataset at the COICOP 6 level

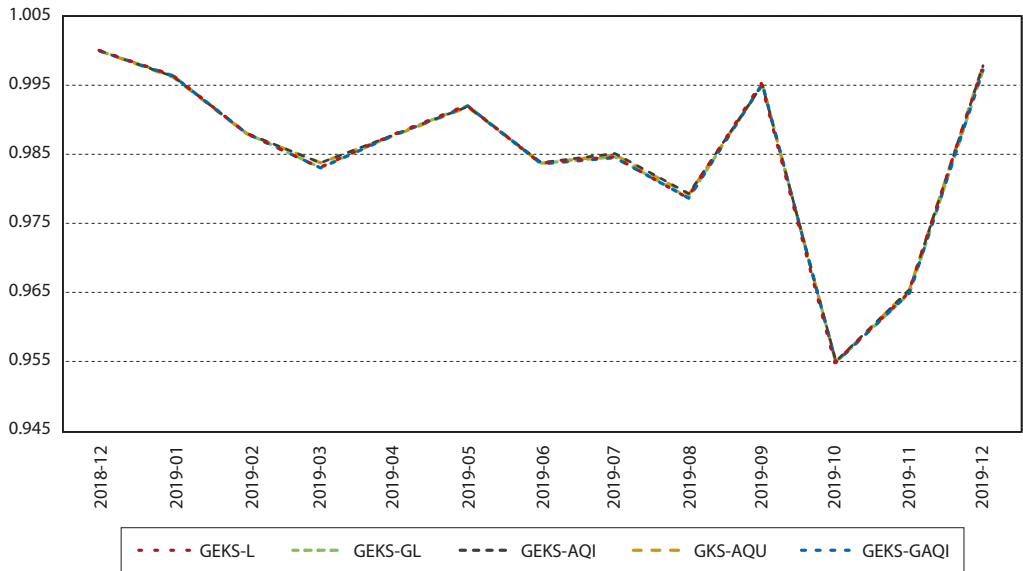


Source: Own computations

Figure 3 Comparison of the GEKS, GEKS-GL, GAQI, GEKS-AQU, GEKS-GAQI indices for the *milk* dataset at the GTIN level



Source: Own computations

Figure 4 Comparison of the GEKS, GEKS-GL, GAQI, GEKS-AQU, GEKS-GAQI indices for the *milk* dataset at the COICOP 6 level

Source: Own computations

At the GTIN level, smaller differences between index values are found for the indices presented in Figure 3 (GEKS-L, GEKS-GL, GEKS-AQI, GEKS-AQU, GEKS-GAQI). Similarly, at the COICOP 6 level, the differences between these indices are smaller than the differences between the other multilateral indices. A general observation is that smaller differences between the indices considered tend to occur at higher levels of data aggregation.

Table 2 Comparison of the GEKS-GAQI index and with the other indices for the *milk* dataset at two data aggregation levels

GTIN LEVEL									
TIME	GEKS	CCDI	GEKS-L	GEKS-GL	GEKS-AQI	GEKS-AQU	GEKS-GAQI	GK	TPD
2019-01	1.0020172	1.0018004	0.99739	0.9978795	0.9973837	0.9973899	0.9978555	1.0065835	1.0060062
2019-02	1.000133	0.9997978	0.9976675	0.9974851	0.9977317	0.9976676	0.9974946	1.0008587	1.0008605
2019-03	0.9839258	0.9840643	0.9845048	0.9833189	0.9844226	0.9845519	0.9832747	0.9818819	0.9838784
2019-04	0.9936427	0.9932822	0.9884879	0.9888732	0.9884773	0.9885357	0.9888407	0.9956176	0.995459
2019-05	0.9899234	0.9898612	0.9867213	0.987091	0.9867022	0.9867849	0.9870584	0.9919439	0.9918611
2019-06	0.9889829	0.9888433	0.9836283	0.9840436	0.9835558	0.9836914	0.9839941	0.9924056	0.9923642
2019-07	0.9862652	0.9864494	0.9843188	0.9846598	0.9843093	0.9843833	0.9846664	0.9887818	0.9887632
2019-08	0.9981114	0.9978518	0.993411	0.9937479	0.9933716	0.9935633	0.9936708	1.0001467	0.9999378
2019-09	0.9952078	0.9951481	0.9939553	0.9937319	0.9939327	0.9941192	0.993664	0.9941353	0.9949736
2019-10	0.9776535	0.9773428	0.9733868	0.9732912	0.9733652	0.9735608	0.9732242	0.9795284	0.979855
2019-11	0.9805743	0.9815496	0.9880499	0.9879278	0.9880425	0.9882264	0.9879558	0.9780526	0.9786244
2019-12	0.9876664	0.9876167	0.9844756	0.984444	0.9844791	0.9846515	0.9844307	0.9895606	0.9896739

Table 2

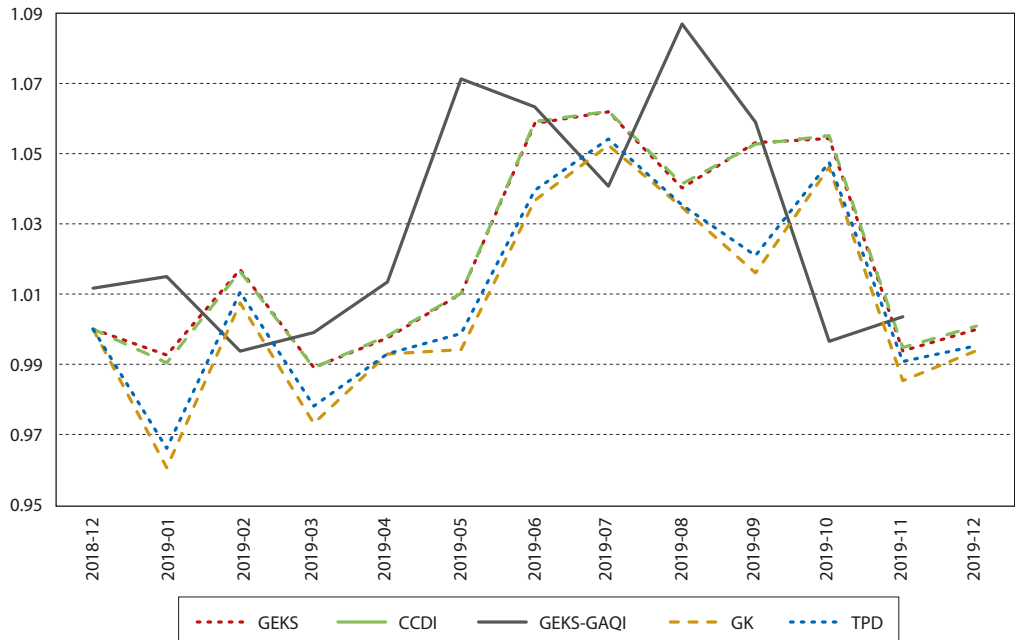
(continuation)

COICOP 6 LEVEL									
TIME	GEKS	CCDI	GEKS-L	GEKS-GL	GEKS-AQI	GEKS-AQU	GEKS-GAQI	GK	TPD
2019-01	0.9951258	0.9950253	0.9962738	0.9964043	0.9961855	0.9962738	0.9963452	0.9941479	0.9939163
2019-02	0.9866407	0.9865364	0.9880469	0.988047	0.9879654	0.9880469	0.987992	0.9855307	0.9853641
2019-03	0.980645	0.9805819	0.9837793	0.9830692	0.9837639	0.9837793	0.9830445	0.9777051	0.9783779
2019-04	0.986313	0.9862242	0.9878349	0.9878334	0.9877271	0.9878349	0.9877601	0.9849386	0.9848847
2019-05	0.9906549	0.9905792	0.9919102	0.992044	0.991862	0.9919102	0.9920148	0.9898024	0.9895561
2019-06	0.9825183	0.9824019	0.9838127	0.9836925	0.9836782	0.9838127	0.9835999	0.9814701	0.9815032
2019-07	0.98374	0.9836404	0.9851054	0.9846292	0.9849333	0.9851054	0.9845176	0.9826784	0.9830589
2019-08	0.9776267	0.9775384	0.9793169	0.9787175	0.9791766	0.9793169	0.9786326	0.9764093	0.9768421
2019-09	0.9940356	0.9939354	0.9950924	0.9952164	0.9950268	0.9950924	0.9951712	0.9933037	0.9930228
2019-10	0.9540006	0.9539199	0.9550102	0.9548949	0.9549045	0.9550102	0.9548293	0.953199	0.9532513
2019-11	0.9619743	0.9618896	0.9653588	0.964897	0.9653985	0.9653588	0.9649294	0.9591934	0.9593704
2019-12	0.998605	0.9987082	0.9978269	0.9971573	0.9979069	0.9978269	0.9972056	0.9993562	1.000172

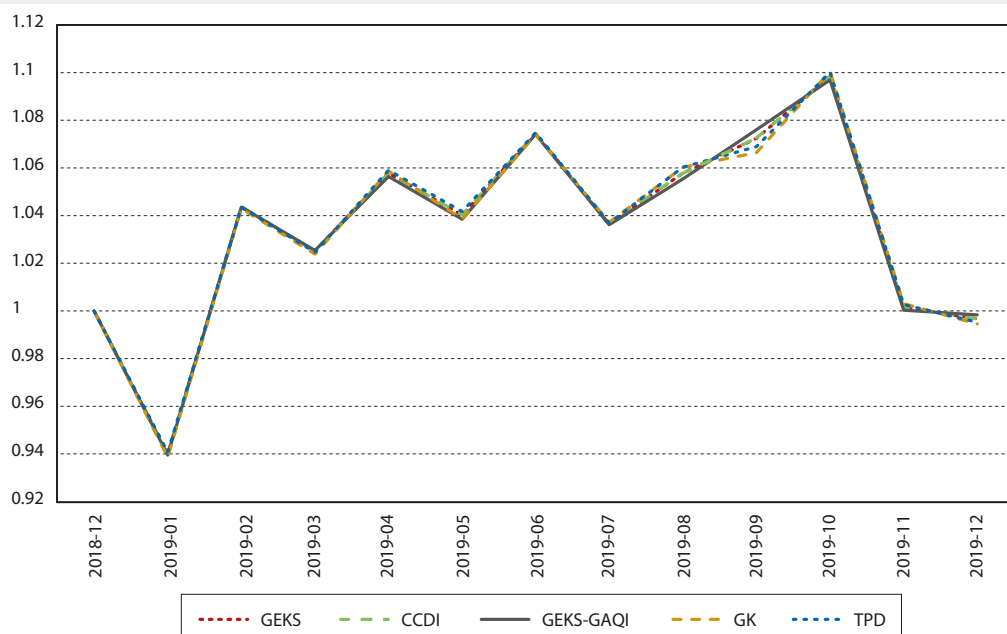
Source: Own calculations

The differences between the indices (GEKS, CCDI, etc.) are relatively minor. The lines closely follow each other, indicating small variations.

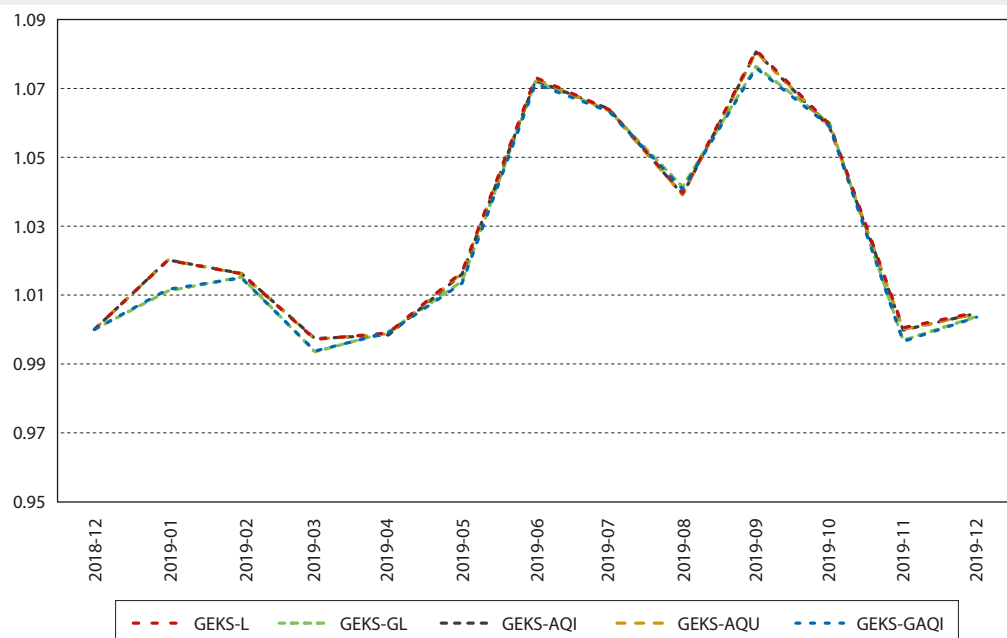
Figure 5 Comparison of the GEKS, CCDI, GEKS-GAQI, TPD indices for the *coffee* dataset at the GTIN level



Source: Own computations

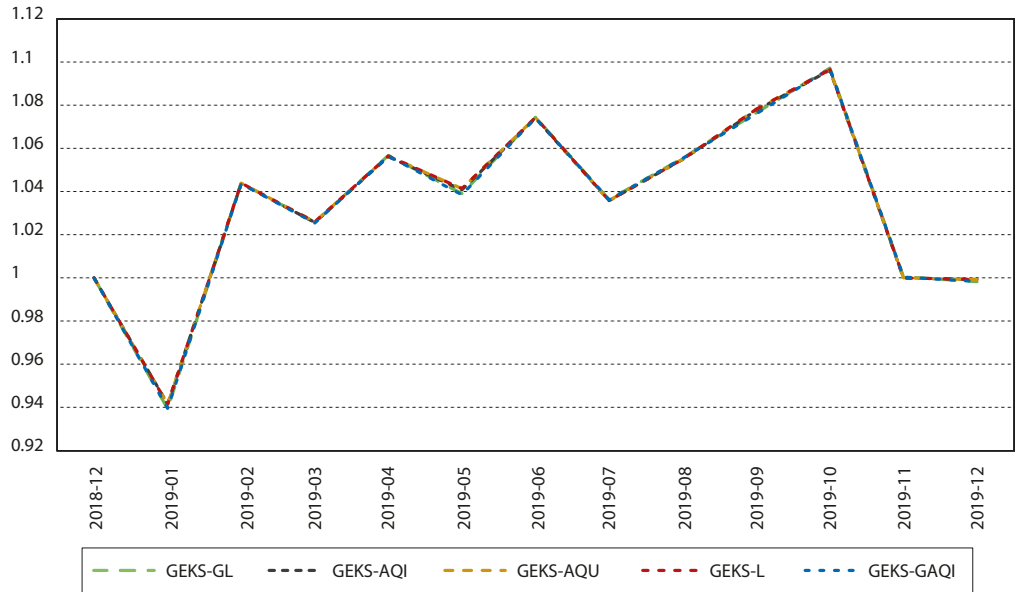
Figure 6 Comparison of the GEKS, CCDI, GEKS-GAQI, TPD indices for the *coffee* dataset at the COICOP 6 level

Source: Own calculations

Figure 7 Comparison of the GEKS, GEKS-GL, GAQI, GEKS-AQU, GEKS-GAQI indices for the *coffee* dataset at the GTIN level

Source: Own calculations

Figure 8 Comparison of the GEKS, GEKS-GL, GAQI, GEKS-AQU, GEKS-GAQI indices for the *coffee* dataset at the COICOP 6 level



Source: Own calculations

The largest differences for the *coffee* dataset were observed among the GEKS, CCDI, GEKS-GAQI, and TPD indices at the GTIN level. Moreover, at the COICOP 6 level, these differences decrease and mainly occur in the period 2019-07 to 2019-10. For GEKS, GEKS-GL, GAQI, GEKS-AQU, and GEKS-GAQI indices, the differences in values are also smaller at the COICOP level 6 compared to the GTIN level. Figure 7 and Figure 8 show greater stability in index values, indicating closer alignment between the considered indices. In contrast, Figure 5 and Figure 6 demonstrate larger fluctuations in index values and a more pronounced downward trend.

Table 3 Comparison of the GEKS-GAQI index with the other indices for the *coffee* dataset at two data aggregation levels

GTIN LEVEL									
TIME	GEKS	CCDI	GEKS-L	GEKS-GL	GEKS-AQI	GEKS-AQU	GEKS-GAQI	GK	TPD
2019-01	0.9926021	0.9903859	1.0201251	1.0111599	1.0201991	1.0201502	1.0116642	0.9605424	0.9660623
2019-02	1.0169978	1.0164447	1.0162029	1.0152675	1.0162685	1.0161897	1.0149682	1.0074853	1.010593
2019-03	0.9890729	0.9890071	0.9971429	0.9935341	0.9973624	0.997191	0.9937572	0.9731303	0.9780821
2019-04	0.9975123	0.998064	0.999012	0.9993254	0.9984175	0.9987502	0.9990596	0.9929061	0.9929982
2019-05	1.0101558	1.0101987	1.0167377	1.0139972	1.0160737	1.0164169	1.0134474	0.9941988	0.998733
2019-06	1.0585925	1.0591717	1.0731752	1.0720446	1.0723201	1.0728185	1.0713077	1.0366952	1.0395706
2019-07	1.0619259	1.0621074	1.0638811	1.0636037	1.064041	1.0635635	1.0633356	1.0524529	1.0541889
2019-08	1.0402685	1.0413992	1.0396131	1.041449	1.0391567	1.0391571	1.0407706	1.0346923	1.0353911
2019-09	1.0531954	1.0527549	1.0811057	1.0763807	1.0806033	1.0806315	1.0759171	1.0160246	1.0209771
2019-10	1.0543306	1.0551312	1.0595413	1.0599256	1.0587449	1.0590765	1.0590269	1.046137	1.0475339
2019-11	0.9937735	0.9947734	1.0005079	0.996838	0.9998289	1.000069	0.9965204	0.9852843	0.9907918
2019-12	0.9999563	1.0008778	1.004988	1.0037686	1.0046772	1.0045472	1.0035874	0.9939147	0.995287

Table 3

(continuation)

COICOP 6 level									
TIME	GEKS	CCDI	GEKS-L	GEKS-GL	GEKS-AQI	GEKS-AQU	GEKS-GAQI	GK	TPD
2019-01	0.9399788	0.9399845	0.9413398	0.9393232	0.9415082	0.9413398	0.9394949	0.9386217	0.9407134
2019-02	1.0436911	1.0436747	1.0438981	1.0438759	1.0435872	1.0438981	1.0435712	1.0429357	1.0434871
2019-03	1.025181	1.0251125	1.0259509	1.0256551	1.0257049	1.0259509	1.02541	1.0238247	1.0246086
2019-04	1.0578008	1.0578484	1.0564336	1.0567317	1.0560931	1.0564336	1.0564099	1.0586664	1.0590341
2019-05	1.0402501	1.0403372	1.0413519	1.0390375	1.0407758	1.0413519	1.0384942	1.0386484	1.0417217
2019-06	1.074446	1.0745063	1.0742963	1.0744018	1.0739492	1.0742963	1.0740775	1.0741201	1.0747576
2019-07	1.036507	1.0365144	1.0358323	1.0362459	1.0357696	1.0358323	1.0361837	1.0372845	1.0368635
2019-08	1.0577503	1.0577807	1.0549522	1.055459	1.0548623	1.0549522	1.055368	1.0602694	1.0602884
2019-09	1.0724209	1.0725615	1.0781131	1.0764596	1.0776213	1.0781131	1.0760158	1.0664161	1.0689335
2019-10	1.0985534	1.0985983	1.096479	1.0971887	1.0962252	1.096479	1.0969446	1.1004277	1.1001417
2019-11	1.0015007	1.0015133	0.9999942	1.0004513	0.9998789	0.9999942	1.0003346	1.0030366	1.002678
2019-12	0.9967013	0.9966486	0.9991691	0.9982211	0.9992552	0.9991691	0.998309	0.9945617	0.995269

Source: Own calculations

3.2 Distances of all indices from the benchmark GEKS-GAQI

Table 4 for the *milk* data set and Table 5 for the *coffee* dataset show the differences in distances between the target GEKS-GAQI index and the other indices.

Table 4 Distances between the benchmark GEKS-GAQI index and the other indices for the *milk* dataset at two data aggregation levels

Index	Distance	Underestimation	Overestimation
GTIN level			
GEKS	0.356	0.062	0.295
CCDI	0.336	0.053	0.282
GEKS-L	0.034	0.018	0.017
GEKS-GL	0.004	0.000	0.003
GEKS-AQI	0.035	0.019	0.016
GEKS-AQU	0.037	0.014	0.023
GK	0.550	0.094	0.456
TPD	0.540	0.078	0.462
COICOP 6 LEVEL			
Index	Distance	Underestimation	Overestimation
GEKS	0.141	0.130	0.012
CCDI	0.151	0.138	0.013
GEKS-L	0.032	0.002	0.030
GEKS-GL	0.006	0.001	0.005
GEKS-AQI	0.029	0.004	0.025
GEKS-AQU	0.032	0.002	0.030
GK	0.272	0.254	0.018
TPD	0.272	0.247	0.025

Source: Own calculations

For the *milk* dataset, the biggest differences in values were noted between the GEKS-GAQI and GK indices and between the GEKS-GAQI and TPD indices. The smallest differences were found between the GEKS-GAQI and GEKS-GL indices. This conclusion holds for both the GTIN and the COCOP 6 levels.

Table 5 Distances between the benchmark GEKS-GAQI index and the other indices for the *coffee* dataset at two data aggregation levels

Index	Distance	Underestimation	Overestimation
GTIN level			
GEKS	0.659	0.642	0.017
CCDI	0.644	0.626	0.018
GEKS-L	0.259	0.010	0.249
GEKS-GL	0.045	0.006	0.039
GEKS-AQI	0.245	0.021	0.224
GEKS-AQU	0.242	0.016	0.226
GK	2.082	2.082	0.000
TPD	1.776	1.776	0.000
COICOP 6 level			
Index	Distance	Underestimation	Overestimation
GEKS	0.125	0.045	0.080
CCDI	0.127	0.045	0.082
GEKS-L	0.086	0.013	0.073
GEKS-GL	0.025	0.002	0.022
GEKS-AQI	0.081	0.021	0.060
GEKS-AQU	0.086	0.013	0.073
GK	0.259	0.137	0.122
TPD	0.249	0.092	0.157

Source: Own calculations

For the *coffee* dataset, the largest differences in distances are observed for the GK and TPD indices at both the GTIN and COICOP 6 levels. Conversely, the smallest differences occur between the GEKS-GAQI and GEKS-GL indices.

CONCLUSIONS

The overall conclusion of the study is not surprising and is in line with predictions: the benchmarking of indices is strongly dependent on the level of aggregation at which it is conducted. The COICOP 6 level provides more stable and consistent patterns, while the GTIN provides more volatility in index values. However, the following observation detected is interesting and quite surprising: for the milk collection, at the GTIN code level, the considered multilateral indices seem to generate higher values than the proposed GEKS-GAQI index. The opposite situation is observed at the COICOP 6 level using the same scanner dataset: here the multilateral indices under consideration seem to mainly generate lower values than the GEKS-GAQI index value. In contrast, for the *coffee* dataset, the situation is reversed: at the GTIN code level, the existing multilateral indices generate lower values than the proposed GEKS-GAQI index. Yet, at the COICOP 6 level and using the same scanner dataset, the existing multilateral indices generate higher values than the proposed GEKS-GAQI index.

Overall, our proposal for the GEKS-GAQI index appears to be methodologically and axiomatically correct. This index satisfies the restrictive *identity test*, which makes it, along with the GEKS-L and GEKS-GL indices, a rather unique multilateral index. In an empirical study, it turned out that indices satisfying the *identity test* (the GEKS-GAQI and GEKS-GL indices) generate the most similar values (see Tables 3 and 4). In addition, since the index does *quality adjusting*, it bridges the gap between the idea of GEKS-type indices and the Geary-Khamis method.

References

- AUSTRALIAN BUREAU OF STATISTICS. (2016). *Making Greater Use of Transactions Data to Compile the Consumer Price Index*. Information Paper 6401.0.60.003, Canberra.
- BALK, M. (1995). Axiomatic price index theory: a survey. *Int. Statist. Rev.*, 63: 69–95.
- BIAŁEK, J., BERĘSEWICZ, M. (2021). Scanner data in inflation measurement: from raw data to price indices. *The Statistical Journal of the IAOS*, 37: 1315–1336.
- BIAŁEK, J., PANEK, T., KŁOPOTEK, M. (eds.). (2022). *Nowoczesne technologie i nowe źródła danych w pomiarze inflacji*. Warsaw: GUS.
- BIAŁEK, J. (2022). Scanner data processing in a newest version of the PriceIndices package. *Statistical Journal of the IAOS*, 38(4): 1369–1397.
- BIAŁEK, J. (2023a). Quality adjusted GEKS-type indices for price comparisons based on scanner data. *Statistics in Transition New Series*, 24: 151–169.
- BIAŁEK, J. (2023b). Improving quality of the scanner CPI: proposition of new multilateral methods [online]. *Quality and Quantity*, 57: 2893–2921. <<https://doi.org/10.1007/s11135-022-01506-6>>.
- CARLI, G. (1804). Del valore e della proporzione de'metalli monetati. In: *Scrittori Classici Italiani di Economia Politica*, 13: 297–336.
- CAVES, D. W., CHRISTENSEN, L. R., DIEWERT, W. E. (1982). Multilateral comparisons of output, input, and productivity using superlative index numbers. *Economic Journal*, 92(365): 73–86.
- CHESSA, A. G., VERBURG, J., WILLENBORG, L. (2017). A comparison of price index methods for scanner data. In: *15th Meeting of the Ottawa Group*, Eltville, 10–12.
- DE HAAN, J., KRSINICH, F. (2018). Time dummy hedonic and quality-adjusted unit value indexes: Do they really differ? *Review of Income and Wealth*, 64(4): 757–776.
- ELTETŐ, O., KÖVES, P. (1964). On a problem of index number computation relating to international comparison. *Statisztikai Szemle*, 42(10): 507–518.
- EUROSTAT. (2018). *Practical guide for processing supermarket scanner data*. Harmonised Index of Consumer Prices.
- FISHER, I. (1922). *The making of index numbers: a study of their varieties, tests, and reliability, Number 1*. Houghton Mifflin.
- GEARY, R. C. (1958). A note on the comparison of exchange rates and purchasing power between countries. *Journal of the Royal Statistical Society: Series A*, 121(1): 97–99.
- GINI, C. (1931). On the circular test of index numbers. *Metron*, 9(9): 3–24.
- JEVONS, W. S. (1865). On the variation of prices and the value of the currency since 1782. *Journal of the Statistical Society of London*, 28(2): 294–320.
- KHAMIS, S. H. (1972). A new system of index numbers for national and international purposes. *Journal of the Royal Statistical Society: Series A*, 135(1): 96–121.
- INTERNATIONAL LABOUR OFFICE. (2004). *Consumer Price Index Manual: Theory and Practice*. Geneva.
- LAMBORAY, C. (2017). The geary khamis index and the lehr index: how much do they differ. In: *15th Meeting of the Ottawa Group*, 10–12.
- LASPEYRES, K. (1871). Die Berechnung einer Mittleren Waarenpreisteigerung. *Jahrbücher für Nationalökonomie und Statistik*, 16(1): 296–318.
- TÖRNQVIST, L. (1936). The bank of Finland's consumption price index. *Bank of Finland Monthly Bulletin*, 1–8.
- WALSH, C. M. (1901). *The Measurement of General Exchange Value*. Macmillan and Co.
- VAN LOON, K. V., ROELS, D. (2018). Integrating big data in the Belgian CPI. In: *Meeting of the group of experts on consumer price indices*, Geneva, Switzerland, 8–9 May.
- VON AUER, L. (2019). The nature of chain drift. In: *17th Meeting of the Ottawa Group on Price Indices*, Rio de Janeiro, Brasil, 8–10 May.
- VON DER LIPPE, P. (2007). *Index Theory and Price Statistics*. Berlin, Germany: Peter Lang.

Artykuł 3

Artykuł 3 Estimating the elasticity of substitution when compiling the CES cost of living index on scanner data

opublikowany w

Quality & Quantity Vol. 58(6), 2024

Wkład współautorów:

Natalia Pawelec: 34%

Jacek Białek: 33%

Sylwia Roszkowska 33%

Wkład własny:

Wkład Autorki autoreferatu obejmuje udział w opracowaniu koncepcji badania, doborze danych i metod analizy oraz redakcji tekstu. W ramach tego równego podziału autorka niniejszej rozprawy odpowiadała za część analityczną związaną z przygotowaniem i przetworzeniem danych skanowanych (klasyfikacja, dopasowanie produktów w czasie, implementacja w pakiecie PriceIndices) oraz za współpracę przy opracowaniu i redakcji tekstu artykułu.



Estimating the elasticity of substitution when compiling the CES cost of living index on scanner data

Jacek Białek^{1,2} · Natalia Pawelec³ · Sylwia Roszkowska^{3,4}

Accepted: 15 June 2024 / Published online: 26 June 2024
© The Author(s) 2024

Abstract

Scanner data are electronic transaction data that specify turnover and the number of items sold by barcodes, e.g., the Global Trade Article Number. These data are of particular value and interest to theorists and practitioners who wish to measure the Cost of Living Index or the Consumer Price Index, since their complete content makes it possible to compute any price index formula, including superlative indices or CES (Constant Elasticity Substitution) indices. Since the CES index requires the estimation of the elasticity of substitution, this paper focuses on verifying various methods of estimating this parameter based on scanner data. The paper considers both algebraic methods and methods based on the panel regression approach. The main achievement of the paper is the separation of the main factors that affect the estimated value of the elasticity of substitution, i.e., the type of data filter used and the level of data aggregation. The paper also verifies how the elasticity of substitution estimates affect the differences between the values of the CES indices based on these estimates.

Keywords Cost of living index · CES model · Elasticity of substitution · Scanner data

✉ Jacek Białek
jacek.bialek@uni.lodz.pl

Natalia Pawelec
natalia.pawelec@uni.lodz.pl

Sylwia Roszkowska
sylwia.roszkowska@uj.edu.pl

¹ Department of Statistical Methods, University of Lodz, Narutowicza 68, 90-136 Lodz, Poland

² Department of Trade and Services, Statistics Poland, Aleja Niepodległości 208, 00-925 Warsaw, Poland

³ Department of Economic Mechanisms, University of Lodz, Narutowicza 68, 90-136 Lodz, Poland

⁴ Institute of Economics, Finance and Management, Jagiellonian University, Łojasiewicza 4, 30-348 Krakow, Poland

1 Introduction

The Consumer Price Index (CPI) is designed to provide an estimate of changes in the prices of goods and services that households acquire for consumption between two periods (International Labour Office 2004, 2020). Given that in the CPI methodology, the basket of goods is fixed, the CPI can be regarded as a Cost of Goods Index (COGI). However, over time, the US Congress embraced the view that the CPI should reflect cost changes to maintain a constant standard of living (see (Gordon 1999) or visit: <https://www.gao.gov/products/GGD-00-50>). Consequently, based on this alternative approach, the CPI evolved into a Cost of Living Index (COLI). Or, to put it another way, most economists and statisticians agree that COLI is a methodological foundation for constructing the CPI.

The COLI measures changes in the cost of living that result from changes in the price of goods and services, with the condition that other external influences that affect living standards remain unchanged. The COLI concept is consistent with an economic approach, i.e., it is assumed here that a consumer can maximise utility or minimise expenditure, subject to a budget constraint, to achieve a certain utility level.

Although the CPI is often treated as a proxy for the COLI, the latter contains the entire set of consumption goods and services consumed by the households from which they derive utility. Consequently, many differences separate the two. For instance, the COLI includes the goods and services received for free as in-kind transfers from governments or non-profit institutions, and it considers changes in environmental factors that affect consumers' well-being. However, the most easily recognised technical distinction is that the CPI is constructed using a Laspeyres-type index (International Labour Office 2004) formula that does not reflect the potential for consumer substitution in response to relative price changes.

With "traditional" data collection, the only reasonable and pragmatic measure of inflation is the Laspeyres index (International Labour Office 2004). This is because price data mostly come from interviewers in the field, and consumption level data come from the Household Budget Survey, which is never up to date. However, for more than 30 years, alternative data sources, such as scanner data or web-scraped data, have come to the fore in measuring inflation (International Labour Office 2020). Scanner data are of particular value and interest to theorists and practitioners who wish to measure the COLI since their content generates virtually no restrictions when applying any price index formula. Scanner data are transaction data that specify turnover, and numbers of items sold by barcodes, e.g., Global Trade Article Number (GTIN), formerly known as the European Article Number (EAN), or Stock Keeping Unit (SKU). The data are characterised by a huge volume of products and contain complete information about items sold by retail chains at all their locations (outlets), e.g., information about prices and quantities at the barcode level, and product attributes, like the size, weight, sale unit, colour, or package quantity (Australian Bureau of Statistics 2016; Eurostat 2018).

The use of scanner data brings many benefits to statistical offices, as automating its acquisition and processing leads to savings in time and money. Nevertheless, the use of this type of data also has its limitations and methodological challenges. First, an IT system upgrade or the construction of an entirely new IT environment to handle this type of data is most often required. Further, the automatic classification of products into appropriate categories and their matching over time, require the use of advanced machine learning and text analysis methods (Bialek and Beręsewicz 2021). This, in turn, requires building a competent team of people to handle these processes. From the point of view of inflation

measurement itself, we also have many challenges related to, for example, the high turnover of scanned products. The strong seasonality of products, new and disappearing goods with short life cycles make even the choice of an index formula problematic, not to mention about quality adjustments. The need to impute missing prices and thus the dilemma of choosing an imputation method, choosing the width of the time window when implementing multilateral indexes, or finally including products that have never before been observed in the dataset are just some of the potential problems (Eurostat 2018). And then there is the issue of the representativeness of the scanner data. Not all products in the CPI basket can be bought in a supermarket, and not every consumer shops in such places. Not all retailers or markets may have comprehensive scanner data available, and even when they do, there might be gaps or inconsistencies in the data due to various reasons such as technical glitches, data entry errors, or other operational issues. Additionally, scanner data, being transactional in nature, might not capture other relevant factors influencing price dynamics, such as promotional activities, loyalty programs, or other marketing strategies. Data from large supermarket chains might not capture the pricing dynamics in local mom-and-pop stores or online retailers. As a consequence, the lack of representativeness can lead to potential biases in the computed price indices on the basis of scanner data, making them less reflective of actual market conditions. Nevertheless, despite the mentioned difficulties, limitations and challenges, scanner data are gaining popularity and setting a new trend in inflation measurement.

As mentioned above, scanner data allow the calculation of any price index formula, including superlative indices (Diewert 1976), which are considered the best approximation of the COLI. Superlative indices (e.g., the Fisher index (Fisher 1922)) are unattainable in traditional data collection because there is no knowledge of current period consumption. Nevertheless, an interesting option is the ability to determine the CES (Constant Elasticity Substitution) index (International Labour Office 2004; Lloyd 1975; Moulton 1996), which uses the same consumption data as the Laspeyres index. However, it also approximates the superlative indices very well, provided that we accurately determine the elasticity of substitution (see Sect. 2.3). As suggested in Balk (2000), the value of the elasticity of substitution should be less than at a higher data aggregation level but greater than at a lower level. This paper evaluates the usefulness of scanner data in determining the elasticity of substitution for selected homogeneous scanner product groups. It also verifies the hypothesis that as the level of data aggregation increases, the elasticity of substitution decreases. The main purpose of the article is to evaluate selected methods of estimating the elasticity of substitution based on scanner data. It also detects the main factors that affect the quality of this estimation. There are few articles in the literature that examine the effectiveness of methods for estimating elasticities of substitution for scanner products. For example (Can 2022) examined elasticities of substitution for four groups of scanner products (soda, dairy, coffee, cheese) but it considered only one estimation method and only one data aggregation level. In the paper (Jeon et al. 2023), the impact of scanner data on demand elasticity estimates is examined. The same paper proposes methods to adjust these estimates for sales policy analyses. An interesting result from the paper (Jeon et al. 2023) is that scanner data generate statistically different elasticities (with more elastic demand) than other data types. The paper (Liesbeth et al. 2018), on the other hand, analyzes income elasticities for food, calories and nutrients using a regression approach, which is similar to what we called in the paper the **panel regression approach** (see Sect. 3.2). On the other hand, paper [?] leads a discussion on improving the method of estimating elasticity of substitution by approximating the Fisher index by using the CES index. This is the approach that we have labeled by **M-LM method** in this paper (see Sect. 3.1) and included among the *algebraic methods*.

This article, compared to the works cited above, expands the spectrum of methods used to estimate elasticity of substitution and additionally analyzes the impact of both the level of data aggregation and the data filters used on the final results of the elasticity of substitution estimates.

The article is organised as follows: Sect. 2 presents the theoretical background for the Cost of Living Index. Section 3 describes two groups of methods for estimating the elasticity of substitution: an algebraic approach and a panel regression approach. Section 4 empirically compares all discussed methods of estimating the elasticity of substitution. Section 5 lists the most important conclusions of the research.

2 Concept of the cost of living index

2.1 Theoretical background

The theory of the Cost of Living Index (COLI) for a single consumer (or household) was first developed by Konüs (1924), the Russian economist. In his economic approach, the period t quantity vector $q^t = [q_1^t, q_2^t, \dots, q_n^t]$ is determined by the consumer's preference function f and the period t price vector $p^t = [p_1^t, p_2^t, \dots, p_n^t]$ that the consumer faces when observing n commodities or items. The consumer's preferences over the given consumption vector q are assumed to be represented by a continuous, non-decreasing and concave utility function f (International Labour Office 2004). The main assumption here is that the consumer minimises the cost of achieving the period t utility level $u^t \equiv f(q^t)$ for the base period $t = 0$ and the current period $t = 1$ and consequently the consumer's cost function is defined as follows:

$$C(u^t, p^t) = \min_q \left\{ \sum_{i=1}^n p_i^t q_i : f(q) = u^t \right\}, t = 0, 1. \quad (1)$$

The Konüs family of true cost of living indices that compare the current period with the base one, is defined as the ratio of minimum costs of achieving the same utility level $u \equiv f(q)$:

$$P_K(p^0, p^1, q) = \frac{C(f(q), p^1)}{C(f(q), p^0)}, \quad (2)$$

where $q = (q_1, q_2, \dots, q_n)$ is a positive reference quantity vector. Assuming additionally that function f is *linearly homogeneous*, i.e., it satisfies $f(\lambda q) = \lambda f(q)$ for any positive λ , which is known in the economic literature as *homothetic preferences*, we get $C(u, p) = uc(p)$, where the unit cost function is defined as $c(p) \equiv C(1, p)$. The homothetic preferences assumption simplifies the family of true cost of living indices as follows (International Labour Office 2004):

$$P_K(p^0, p^1, q) = \frac{c(p^1)}{c(p^0)}, \quad (3)$$

Similarly, under the homothetic preferences assumption, the implicit quantity index that corresponds to the true cost of living index (3) is the utility ratio $f(q^1)/f(q^0)$. Under all

the above-mentioned assumptions, let us also assume that the consumer has the following quadratic mean of order r utility function:

$$f^r(q) = \sqrt[r]{\sum_{i=1}^n \sum_{k=1}^n a_{ik} q_i^{r/2} q_k^{r/2}}. \quad (4)$$

where the parameters a_{ik} satisfy $a_{ik} = a_{ki}$ for any i and k , and $r \neq 0$. Diewert (1976) demonstrate that the utility function f^r is a flexible functional form, i.e., it can approximate an arbitrary twice-differentiable, linearly homogeneous functional form to the second order.

Let us define the quadratic mean of order r quantity index as follows:

$$Q^r(p^0, p^1, q^0, q^1) = \frac{\sqrt[r]{\sum_{i=1}^n s_i^0 \left(\frac{q_i^1}{q_i^0}\right)^{r/2}}}{\sqrt[r]{\sum_{i=1}^n s_i^1 \left(\frac{q_i^1}{q_i^0}\right)^{-r/2}}}. \quad (5)$$

where

$$s_i^t = \frac{p_i^t q_i^t}{\sum_{k=1}^n p_k^t q_k^t}, t = 0, 1. \quad (6)$$

It is known (International Labour Office 2004, page 321) that

$$Q^r(p^0, p^1, q^0, q^1) = \frac{f^r(q^1)}{f^r(q^0)}. \quad (7)$$

Since the quantity index Q^r is exact for the functional form f^r , following to Diewerts's terminology (see (Diewert 1976) and also earlier work (Fisher 1922)), Q^r is a superlative price index for each $r \neq 0$. For each quantity index (5), we can define the corresponding implicit quadratic mean of order r price index as follows:

$$P_{im}^r(p^0, p^1, q^0, q^1) = \frac{\sum_{i=1}^n p_i^1 q_i^1}{Q^r(p^0, p^1, q^0, q^1) \sum_{i=1}^n p_i^0 q_i^0}. \quad (8)$$

It can be demonstrated that the P_{im}^r price index is also a superlative index formula for each $r \neq 0$. To establish this, it is necessary to prove its exactness for a flexible functional form c_{im}^r , where c_{im}^r is the unit cost function that corresponds to the aggregator function f^r (International Labour Office 2004). In analogy to the quadratic mean of order r quantity index (5), one can also define the quadratic mean of order r price index:

$$P^r(p^0, p^1, q^0, q^1) = \frac{\sqrt[r]{\sum_{i=1}^n s_i^0 \left(\frac{p_i^1}{p_i^0}\right)^{r/2}}}{\sqrt[r]{\sum_{i=1}^n s_i^1 \left(\frac{p_i^1}{p_i^0}\right)^{-r/2}}}. \quad (9)$$

It is also a superlative price index formula (Diewert 1976; International Labour Office 2004).

2.2 Proxies for the cost of living

In the economic and statistical literature, superlative indices are considered to be the best approximation of the Cost of Living Index (White 1999). From a theoretical standpoint, superlative indices should also best measure the CPI, which is used in practice to approximate the “true” COLI (von der Lippe 2007).

It is known (International Labour Office 2004) that $P_{im}^1 = P_W$ and $P_{im}^2 = P_F$, where P_W and P_F denote the Walsh price index (Walsh 1901) and the Fisher price index (Fisher 1922) respectively. It can also be shown that $P^{r \rightarrow 0} = P_T$ and $P^2 = P_F$, where P_T denotes the Törnqvist price index (Törnqvist 1936). Since all the indices above are superlative, they can be considered good proxies for the COLI. With the notions introduced in Sect. 2.1, the formulas of these superlative indices can be expressed as follows:

$$P_W = \frac{\sum_{i=1}^n p_i^1 \sqrt{q_i^0 q_i^1}}{\sum_{i=1}^n p_i^0 \sqrt{q_i^0 q_i^1}}, \quad (10)$$

$$P_T = \prod_{i=1}^n \left(\frac{p_i^1}{p_i^0} \right)^{\frac{s_i^0 + s_i^1}{2}}, \quad (11)$$

$$P_F = \left(\frac{\sum_{i=1}^n q_i^0 p_i^1 \sum_{i=1}^n q_i^1 p_i^1}{\sum_{i=1}^n q_i^0 p_i^0 \sum_{i=1}^n q_i^1 p_i^0} \right)^{0.5} = \sqrt{P_{La} P_{Pa}}, \quad (12)$$

where P_{La} and P_{Pa} denote the Laspeyres (1871) and Paasche (1874) price index respectively.

Considering the class of indices presented in (9) it was proven (Hill 2006) that all the commonly used superlative indices, i.e. P_W , P_T and P_F , fall into the interval $0 \leq r \leq 2$. It was also demonstrated that these superlative indices approximate each other (Diewert 1976). Thus, the remainder of the paper uses each index to determine the elasticity of substitution. It turns out that the choice of the superlative index in the “index” approach to determining the elasticity of substitution can make a difference (see Sects. 3.1 and 4).

2.3 CES model of consumer preferences and CES index

Let us consider the following unit cost function, which depends on prices and the additional parameter σ :

$$c_\sigma(p) = \alpha_0 \left(\sum_{i=1}^n \alpha_i p_i^{1-\sigma} \right)^{\frac{1}{1-\sigma}}, \quad \sigma \neq 1, \quad \text{or} \quad (13)$$

$$\ln(c_\sigma(p)) = \alpha_0 + \sum_{i=1}^n \alpha_i \ln(p_i), \quad \sigma = 1.$$

The unit cost function defined in (13) corresponds to the constant elasticity of substitution (CES) aggregator function, which was introduced in Arrow et al. (1961).

Parameter σ of the CES cost function (13) corresponds to the elasticity of substitution in the following CES utility function:

$$u(q) = \left(\sum_{i=1}^n \alpha_i q_i^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \quad (14)$$

The elasticity of a function of a single variable measures the percentage response of a dependent variable to a percentage change in the independent variable. By contrast, the elasticity of substitution between two inputs measures the percentage response of the relative marginal utility of the two goods to a percentage change in the ratio of their quantities (Saito 2012).

From Eq. (14), we conclude that the elasticity of substitution σ takes values from 0 to $+\infty$. If $\sigma \rightarrow +0$, the unit cost function (13) becomes linear and corresponds to the Leontief utility function which exhibits 0 substitutability between goods. Otherwise, if $\sigma \rightarrow 1$, the corresponding utility function is a Cobb-Douglas function. As σ tends to $+\infty$ utility function (14) is characterised by substitutability between each pair of inputs (Diewert and Fox 2022).

The CES cost function does not have a fully flexible functional form (if the number of commodities exceed two) but it is frequently used to aggregate commodities in a group of goods which are thought to be highly substitutable with each other (International Labour Office 2004).

In the rest of the paper, we will refer to the price index based on the elasticity of substitution as the CES index. This index requires estimation because it contains a parameter in the body, which is in line with the literature (e.g. Lent and Dorfman (2009)). Let us first consider the Lloyd-Moulton index (Lloyd 1975; Moulton 1996), which has the following form for $\sigma \neq 1$:

$$P_{LM}(\sigma) = \left(\sum_{i=1}^n s_i^0 \left(\frac{p_i^1}{p_i^0} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}. \quad (15)$$

It can be proven (International Labour Office 2004) that the Lloyd-Moulton index defined in (15) is accurate for CES preferences under the assumption of consumer cost-minimising behaviour, i.e. it holds that

$$P_{LM}(\sigma) = \frac{c_{\sigma}(p^1)}{c_{\sigma}(p^0)}. \quad (16)$$

Note that $P_{LM}(0) = P_{La}$ and that the Lloyd-Moulton price index (i.e., the CES index) use the same price and quantity information as the Paasche index (current-period weights are unnecessary), making it potentially of great practical value to statistical offices. With a correct estimate of the elasticity of substitution σ , assuming that it does not change too rapidly over time, the CES index should be a good proxy for the COLI.

Another index in the literature that approximates COLI, but using weights only from the base period, is the AG Mean index (Lent and Dorfman 2009). A rolling procedure is used to determine what parameter is in the body of this index, which must be constantly updated with new data. However, despite its computational simplicity and certain advantages, this

index will not be considered in the paper because it is not derived directly from the economic approach.

3 Estimates of the elasticity of substitution

Although there are many methods for estimating the elasticity of substitution, this paper uses two methods, or rather groups of methods, that appear most frequently in the literature (Ivancic et al. 2010; Haan et al. 2010). The first group compares bilateral indices, at least one of which depends on the elasticity of substitution. The second group uses a panel regression model that explains changes in the level of consumption through price changes based on the CES assumption. The first methods, which, following (Ivancic et al. 2010) we will refer to as the *algebraic approach*, compare only the current period with the base period in determining the elasticity of substitution. By contrast the second methods (*the panel regression approach*) consider the entire time interval between the base period and the current period.

3.1 The algebraic approach

The first method presented here was described by Haan et al. (2010), among others. Determining the elasticity of substitution for periods 0 and 1 involves comparing the CES index (15) and its "current weight counterpart" $P_{CW}(\sigma)$, where:

$$P_{CW}(\sigma) = \left(\sum_{i=1}^n s_i^1 \left(\frac{p_i^1}{p_i^0} \right)^{-\frac{1}{1-\sigma}} \right)^{-\frac{1}{1-\sigma}}. \quad (17)$$

Since the price index $P_{LM}(\sigma)$ monotonically decreases and $P_{CW}(\sigma)$ monotonically increases as σ increases, there is a unique value σ_0 for which it holds that $P_{LM}(\sigma_0) = P_{CW}(\sigma_0)$. Consequently, we have

$$P_{LM}(\sigma_0) = P^{2(1-\sigma_0)}, \quad (18)$$

where the right side of Eq. (18) means the quadratic mean of order r price index $P^r(p^0, p^1, q^0, q^1)$ defined in (9) but for $r = 2(1 - \sigma_0)$. Thus, we conclude that for the designated σ_0 value, the CES index is superlative and can be treated as a good approximation of the Cost of Living Index. However, that although we use the simplifying assumption that σ_0 is constant over time (regardless of the temporal distance of periods 0 and 1), we can only determine this parameter numerically. In the remainder of this paper, we will conventionally denote the method of estimating elasticity of substitution described above with the notation **M-LM**.

Another method of estimating the elasticity of substitution is to "simply" compare the Lloyd-Moulton index to one of the superlative indexes (see e.g. (Ivancic et al. 2010)). In particular, solving numerically (relative to σ) equations: $P_{LM}(\sigma) = P_F$, $P_{LM}(\sigma) = P_T$, and $P_{LM}(\sigma) = P_W$, we get methods labelled **M-F**, **M-T** and **M-W** respectively.

Balk (2000) suggested that the Sato-Vartia index (although not superlative—see Sato (1976); Vartia (1976)) can also be used to estimate the elasticity of substitution. The method of estimating the elasticity of substitution based on the Sato-Vartia index will be denoted by **M-SV**.

3.2 Panel regression approach

The theoretical foundation implies the notion of representative consumers. In other words, it is as if a single consumer buys all goods consumed in this market, substituting one good for another according to a rational decision resulting from optimising a utility function. This representative consumer is a price taker on the market. Although these hypotheses are somewhat unrealistic, the operating conditions are not very different when considering micro markets. To adopt a more realistic framework, we must work on a local market where consumers choose between really substitutable products (e.g., yoghurts sold in specific chain stores). In this market, the assumptions relative to the representative consumer are acceptable since it is possible to substitute a good for another. Moreover, for the consumer who consumes a good in a specific store, it is easier to substitute that good for another in this store rather than buy it in another store. Thus, an aggregate model of behaviour on the market of yoghurts sold in this store, if not the result of the aggregation of individual preferences, is nevertheless a plausible picture of how the micro-market works (Sillard 2012; Leclair et al. 2019).

One can apply a panel regression approach to extract the elasticities of substitution between quantities of goods under consideration and their prices. The most straightforward hedonic regression shows the relation between price (or log price) on the one hand side and product characteristics on the other-hand. Diewert (2003) develops sufficient conditions to allow a hedonic regression to be interpreted as a function of consumer preferences.

In the panel regression approach, we consider the demand function based on cost and utility functions (in logs) as follows:

$$\log(q_{it}) = \beta_0 - \sigma \log(p_{it}) + \epsilon_{it}, \quad (19)$$

The equation does not consider the household income level, the prices of other goods or the consumption structure. The unknown coefficient β_0 captures the σ and α parameters from Eq. (13). ϵ_{it} , and is the error term. The use of panel data implies the heterogeneity of the units analysed. Heterogeneity can be considered by fixed effects (FE) and captured by a parameter with dummy variables relating to the units. In the FE approach, individual effects can be thought of as different free expressions in the time series for individual units: they have the character of time-constant, deterministic parameters that vary for individual i . In contrast, unobserved heterogeneity in a random effects (RE) model can be captured differently. One assumption is that the free expression in the regression model varies randomly for the cross-section. Another assumption is that the random component of the regression equation is the sum of several components, including a random component that relates to a single unit. In demand analysis, the parameters of the distribution of this random component are interpreted as a random variation in preferences (Cameron and Trivedi 2010). Stated that heterogeneity needs attention in contemporary microeconomic research. In general, there is more heterogeneity than studies assume, and in addition, incorrect treatment of heterogeneity can lead to errors when estimating effects that are of interest to researchers (Browning and Carro 2007). The nature of these effects in a random effects model means that it is possible to introduce free expression into a random effects model, which was not the case with a fixed effects model. Individual effects are treated as deterministic, estimable parameters in the fixed effects model. In the random effects model, they are treated as random. As such, they do not constitute additional, potentially estimable parameters but extend the stochastic part of the model (Berry and Haile 2021).

An issue that should be addressed before proceeding is the potential endogeneity between prices and quantities (or expenditure). In the economic approach to index numbers, a consumer (or household) solves a utility maximisation problem. The consumer chooses a bundle of goods to maximise utility subject to some budget constraint. In this framework, prices are assumed to be exogenous. In essence, households regard the observed price data as given, while the quantity data are regarded as solutions to the various optimisation problems. This assumption is reasonable for our analysis as we use data on items that are bought in a store, where prices are generally taken as given by consumers (Ivancic et al. 2010).

However, it can be argued that prices are not exogenous. The primary assumption in those methods is that all prices p_i are exogenous. However, it is very likely they are endogenous (Freyberger 2015). One can imagine that the seller sets the price according to the reaction to the demand he expects. This generates a simultaneity bias when estimating the ordinary least square (OLS) and related methods (FE, RE). We then need an instrument to resolve this problem. The structural equation we are interested in is a demand equation. To identify this equation, we look for an exogenous instrument of the price in a demand equation (Birchall and Verboven 2022; Rilstone 1994). Adding the pricing equation clarifies questions about the optimal choice of instruments for our problem (Chamberlain 1987). The optimal instruments for the price will depend on the characteristics of rival products, but semiparametric analysis of optimal instruments is difficult, if not impossible (see (Newey 1990; Berry et al. 2004)). The correct choice in our case seems to be using variable related to production costs as instrumental variables and estimate parameters by two-stage least squares (2SLS) (Petrin and Train 2010; Beck and Lein 2020).

A synthetic comparison of the algebraic methods and the method using panel regression is summarized in Table 1.

4 Empirical study

4.1 Scanner data sets description

In the following empirical study, we used scanner data from one retail chain in Poland, i.e., monthly data on *yoghurt* (COICOP 5: 011441) and *other articles for personal hygiene* (COICOP 5: 121322) sold in over 500 outlets between December 2021 and December 2022 (i.e., 604,260 and 364,984 records, which means 36.4 MB and 19.7 MB of data in the csv format, respectively). The COICOP 5 *yoghurt* group consisted of the following local COICOP 6 product subgroups: Actimel (19 IDs), chocolate and walnut yoghurt (6 IDs), natural yoghurt (75 IDs), fruit yoghurt (119 IDs), drinking yoghurt (165 IDs), and the COICOP 5 *other articles for personal hygiene* group consists of the following local COICOP 6 product subgroups: tissues (60 products: IDs), wet wipes (88 IDs), toilet paper (117 IDs), baby diapers (193 IDs), sanitary pads (20 IDs), sanitary napkins (67 IDs), and tampons (22 IDs).

Before calculating the elasticities of substitution, the data sets were carefully prepared. Products were classified using the `data_selecting()` and `data_classification()` functions from the `PriceIndices` R package (Bialek 2021). The first function required the manual preparation of dictionaries of keywords and phrases that identified individual product groups. The second function was used for problematic, previously unclassified products and required manual preparation of learning samples based

Table 1 Comparison of properties of algebraic methods and a method using panel regression to estimate elasticity of substitution

Property	Algebraic approach	Panel regression approach
Calculation time and simplicity of implementation	Algebraic methods require less data than the panel regression approach because they only use data from the current and base periods. Consequently, algebraic methods are faster to estimate and simpler to implement	The panel approach requires a time-series sample, so we need more data here to estimate the elasticity of substitution than with algebraic methods. Parameter estimation here is slower and more complex, as it involves, among other things, choosing the right panel model and estimation method
Stability of results due to new data	When we expand the database to include data from the next month, algebraic methods can produce markedly different results from those previously determined (for the previous month). This phenomenon will be more pronounced the lower the level of data aggregation	Shifting the time window over which a panel regression model operates by one month usually does not lead to clear changes in the estimated elasticity of substitution (within the chosen and established model and estimation method)
Stability of results due to the choice of estimation method	Algebraic methods generally lead to much less varied values of the estimated elasticity of substitution than methods based on a panel regression approach	The results of estimating elasticities of substitution depend very much on the panel regression model and estimation method adopted

on historical data. The classification itself was based on machine learning using random trees and the XGBoost algorithm (Tianqi and Carlo 2016). Next, product matching was carried out based on the available GTIN (Global Trade Item Number) bar codes, internal retail chain codes and product labels. To match products, we used the `data_matching()` function from the `PriceIndices` package. To be more precise: products with two identical codes or one of the codes identical and an identical description were automatically matched. Products were also matched if they had identical one of the codes and the Jaro-Winkler distance (Jaro 1989) of their descriptions was smaller than the fixed precision value: 0.02.

4.2 Elasticity of substitution depending on the estimation method and data filter used

In the first phase of our empirical study, we set out to see how the measurement of elasticity of substitution is affected by the estimation method and the data filters used. The analysis included both the *algebraic approach* and *panel regression approach* (see Sect. 3). Three popular data filters were considered: the extreme price filter (**F1**), the dump price filter (**F2**) and the low sale filter (**F3**).

There is an ongoing discussion in the literature about whether or not to use data filters for scanner data, and if so, what kind of filters to apply. As a rule, scanner data indices are calculated using a dynamic approach, with most countries opting for the monthly chain Jevons index. This method is commonly referred to as the dynamic method (Eurostat 2018). The dynamic basket is determined using turnover figures of individual products in two adjacent months, i.e., the product is included in the sample if its turnover is above a fixed threshold determined by the number of products in a given product group. In van Loon and Roels (2018), the following condition for the above-mentioned rule can be met, which indicates whether the i -th product is taken into consideration when comparing months $t - 1$ and t :

$$\frac{s_i^{t-1} + s_i^t}{2} > \frac{1}{n\lambda}, \quad (20)$$

where s_i^t denotes the expenditure share of the i -th product at time t , n is the number of considered products and λ is a fixed parameter (usually and here set to 1.25). This kind of data filter can be called a *low sale filter* (**F3**). Those who support using filters also believe that products that display extreme price changes from one month to another should also be excluded from the sample (*extreme price filter*—**F1**). In this study, the **F1** filter eliminated products whose prices more than tripled or fell more than twice from month to month. As the list of possible data filters is longer. Statistics Belgium for example, implements a filter for products with a clearly decreasing price and substantially decreasing sales (*dump price filter*—**F2**) (van Loon and Roels 2018). The **F2** filter was also implemented in this survey, with the cutoff for dropping sales being -30% from month to month.

For the *algebraic methods*, the elasticities of substitution for successive months were determined for the two scanner datasets under consideration, with reference to a fixed base of December 2021. Observing the results for these methods (Figs. 1 and 2), the **M-F** method (based on the Fisher index) leads to a markedly different (and generally higher) elasticity of substitution results than the other methods. This is especially evident for the yoghurt dataset (see Fig. 1), where the differences (between the results of estimation by the **M-F** method versus the other methods that did not use filtering) exceeded 0.5 in June,

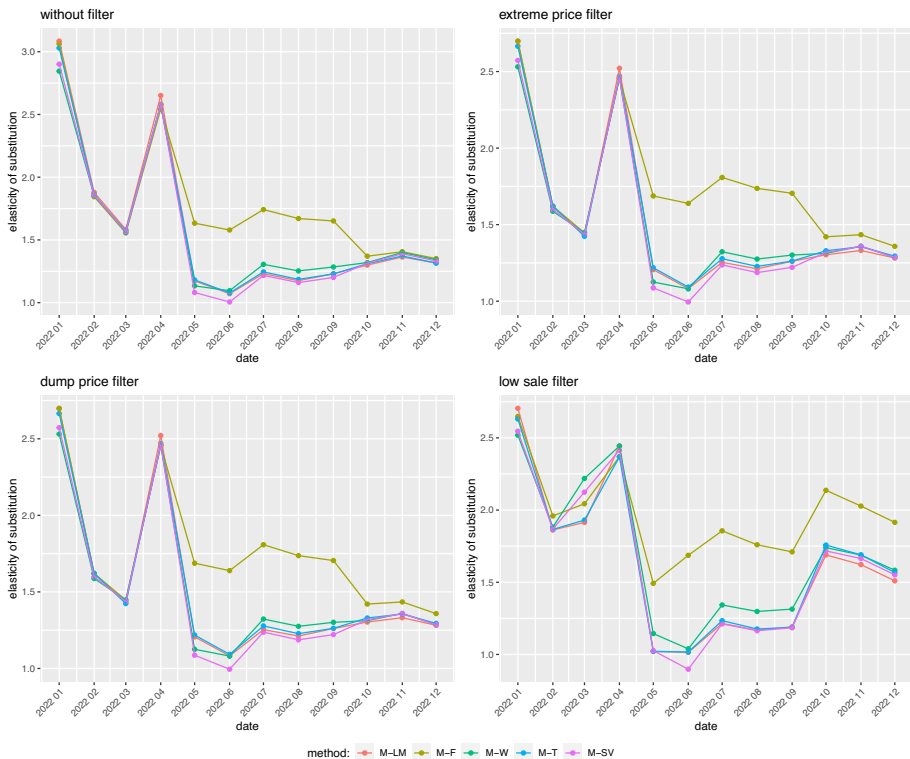


Fig. 1 Elasticity of substitution depending on data filters applied for yoghurt products

2022 (see Fig. 1 and Table 2). For other articles for personal hygiene, the differences in the elasticities of substitution were also close to 0.5 but this occurred in August 2022. This time, the largest values of the measured elasticity of substitution were generated by **M-LM** method (see Table 2). Thus, while it can be concluded that, in general, different algebraic methods lead to noticeably different values of elasticity of substitution, it is not possible to identify a method that always leads to the largest or smallest estimation results.

Additionally use of filters has an impact not only on the determined values of the elasticity of substitution, but also on the magnitude of differences created after applying different algebraic methods. For example, the **F1** and **F2** filters for the data set on yoghurt products (Fig. 1) lead to an even greater distancing of the estimation results obtained after applying algebraic methods.

Analogous results with panel regression approach discussed in Sect. *panel regression approach* discussed in Sect. 3.2 are presented in Table 3. Estimates of Eq. (13) are presented in that table hence when interpreting the sigma parameter, the sign of the parameter at $\log(\text{price})$ variable should be changed to the opposite sign. Additionally, Appendix A includes analogous panel regression estimates using different panel regression methods. The 2SLS method seems to be the most valid in case of analysed groups of products because of endogeneity of prices (and using cost indexes as instrumental variables). Comparing 2SLS estimates with other methods (OLS, FE or RE) one can observe differences confirming that primary assumption that all prices are exogenous generates a simultaneity

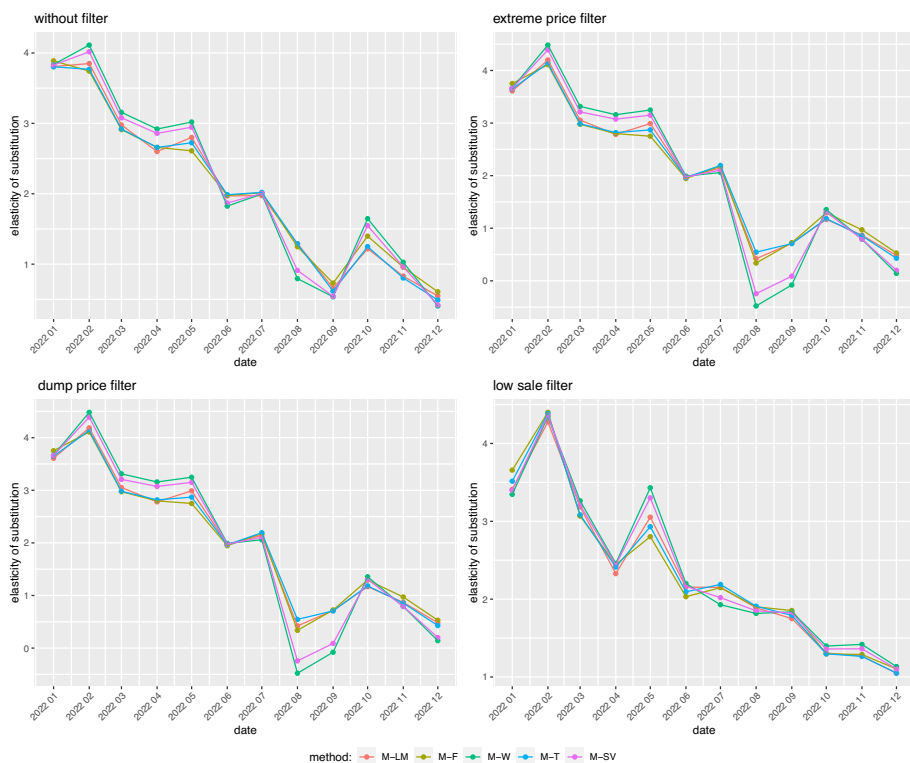


Fig. 2 Elasticity of substitution depending on data filters applied for other articles for personal hygiene

Table 2 The biggest differences in CES values depending on the algebraic method and the data set (*)

Method	Yoghurt (June, 2022)	Other articles for personal hygiene (August, 2022)
M-LM	1.071625	1.2912369
M-F	1.579094	1.2478638
M-W	1.095161	0.7947540
M-T	1.077728	1.2885284
M-SW	1.005402	0.9116364

(*) the case with no filtering

bias. The tests (F-test and Chi-sq), developed by Sanderson and Windmeijer (2016), suggest that the instrumental variables are valid (see (Sanderson and Windmeijer 2016)). The cluster option of standard errors takes heteroskedasticity into account. Focusing only on those results that apply to the most disaggregated case, i.e., results obtained at the GTIN level and signified in the figures by the description "before aggregation" (the impact of aggregation on the estimation results will be discussed in Sect. 4.3). After using the above-mentioned filters, we found that the estimated elasticity of substitution for yoghurt products decreased after using filters, it changed from 1.8 to 1.3–1.4 regardless of the filter (see

Table 3 CES 2SLS estimates on yoghurts and other articles for personal hygiene using the different price filter

Filter	Without	Dump price	Extreme price	Low sale
<i>Yoghurt</i>				
<i>log(price)</i>	-1.801*** (0.369)	-1.298*** (0.270)	-1.419*** (0.286)	-1.396*** (0.288)
SW Chi-sq	27.09	32.58	30.8	30.49
<i>p</i> -val	0.000	0.000	0.000	0.000
SW F-stat	27.02	32.48	30.71	30.4
<i>p</i> -val	0.000	0.000	0.000	0.000
Product fixed effect	Yes	Yes	Yes	Yes
Number of products	353	344	349	349
Number of observations	3521	3460	3497	3500
<i>Other articles for personal hygiene</i>				
<i>log(price)</i>	-0.992*** (0.198)	-0.815*** (0.192)	-0.813*** (0.193)	-1.054*** (0.245)
SW Chi-sq	389.82	420.26	414.79	333.78
<i>p</i> -val	0.000	0.000	0.000	0.000
SW F-stat	389.09	419.45	413.98	333.02
<i>p</i> -val	0.000	0.000	0.000	0.000
Product fixed effect	Yes	Yes	Yes	Yes
Number of products	532	516	516	440
Number of observations	5311	5155	5152	3129

Cluster standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; SW Chi-sq - Sanderson-Windmeijer first-stage chi-squared tests of underidentification SW F-stat - Sanderson-Windmeijer multivariate F test of excluded instruments

Table 3). Additionally, in the case of other articles for personal hygiene, we observed a decrease in the estimated sigma from 0.992 to approximately 0.81 after applying the dump price filter and the extreme price filter. In case of using the low prices filter the elasticity of substitution obtained by using the 2SLS method was approximately the same as in case of no filters.

4.3 Impact of the data aggregation level on CES estimates

The next stage of the study assessed the impact of data aggregation on the elasticity of substitution as measured by various methods. Again, both the algebraic and panel regression-based methods were considered (see Appendix). By the term "before aggregation" we mean the most disaggregated level of data, i.e., the GTIN code level. By contrast, the term "after aggregation" means that we define a homogeneous product at the COICOP 6 group level.

Table 8 presents the elasticities of substitution calculated using *algebraic methods* before and after aggregation for the two data sets and for two variants, i.e., one where elasticities of substitution were calculated by comparing December 2022 to December 2021, and another where the elasticities were computed for each pair of subsequent months, and the means of these elasticities were compared. Regardless of the method of calculation,

the estimated elasticities of substitution are noticeably smaller after data aggregation (as a rule, they are more than twice as small). This is particularly noticeable in the case of other articles for personal hygiene, where the average values of the elasticities of substitution dropped more than fourfold when moving from the GTIN code level to the COICOP 6 level (see Table 8).

With the *panel regression approach*, the parameters of Eq. (19) were estimated using several methods to determine the value of the sigma parameter and to check whether aggregation within individual ID codes or product groups influences the estimated magnitude of the sigma parameter of interest. The estimation used both standard methods for panel data, i.e., panel methods with fixed effects (FE) and random effects (RE); see Sect. 3.2 for a description. We have assumed that price need not be an exogenous variable; hence the two-stage method we have used, using instruments correlated with the explanatory variable. Regarding the yoghurt estimates, we use the price index of the consumer goods milk, cheese, and eggs as the instrument. For other articles for personal hygiene, our instrument is the price index of sold production in manufacturing paper and paper products. Both instruments are monthly single-basis indices. The instruments we used to approximate the development of production costs in the branches. Both instruments used in the regression equations are valid (Stock and Wright 2000; Stock and Yogo 2002; Cragg and Donald 1993). Additionally, to capture the heterogeneity of the products analysed within a given group, we use statistics to heteroskedasticity and clustering on products and groups (2SLS). Furthermore, we supplement our analysis following other researchers with estimates of the equation on first-difference increments (FD). However, in this case, the estimates are inferior and are also based on a shorter time series. For the yoghurts group (Table 4), the estimated sigma for the traditional FE and RE methods is underestimated due to endogeneity. Using instrumental shows causes the estimated sigma parameter to increase to 1.8. Aggregation caused the sigma parameter to be estimated on a relatively small sample (13 periods and 5 groups), which is additionally quite heterogeneous; hence, the estimation of the joint sigma is statistically insignificant (Table 6). We obtained slightly lower (in terms of modulus) sigma parameter estimates for other articles for personal hygiene. Our estimated parameter with instruments in the 2SLS method is 0.992 (Table 3). The post-aggregation estimates are significantly lower (Table 7).

4.4 Impact of the estimation method on CES index values

Section 4.2 showed that different methods of estimating elasticities of substitution can generate noticeably different estimates. In contrast Sect. 4.3 empirically showed that aggregating data leads to smaller values of estimated elasticities of substitution. However, this raises the natural question of whether differences between estimates of the elasticity of substitution lead to measurable differences among the CES indices that are based on them (see Sect. 3.1). To answer this question, it was decided to determine the CES indices for a full year of observations of the two data sets while taking into account: (a) monthly averaged elasticities of substitution determined for all the algebraic methods considered; (b) elasticities of substitution determined by the method based on panel regression (2SLS), which appears to be the most reliable (see Sect. 5).

Figure 3 presents the above-mentioned CES indices based on elasticities of substitution calculated for non-filtered yoghurt products and for two variants, before and after data aggregation. Figure 4 presents analogous results for other articles for personal hygiene. A surprising conclusion is that despite differences in the estimated elasticity of substitution,

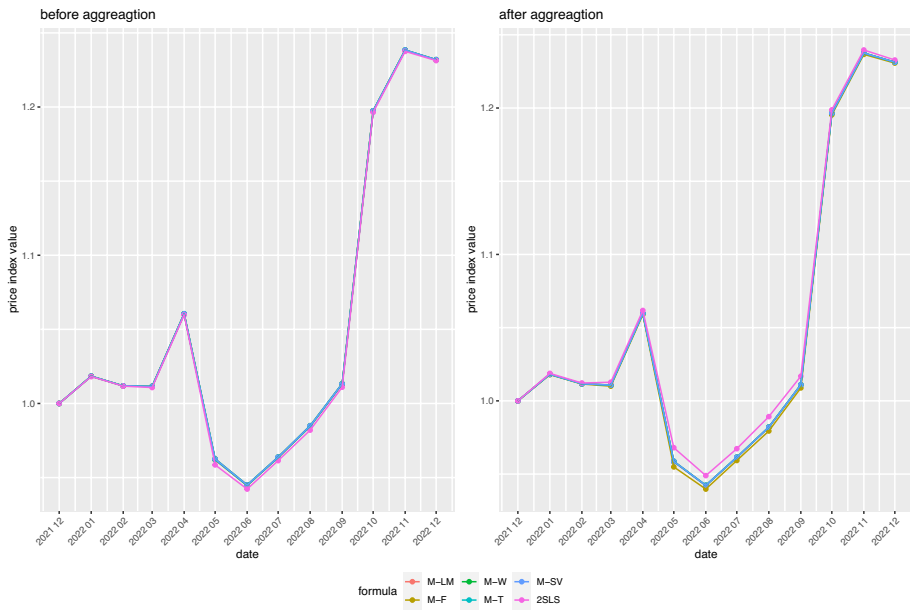


Fig. 3 CES indices based on elasticities of substitution calculated for yoghurt products

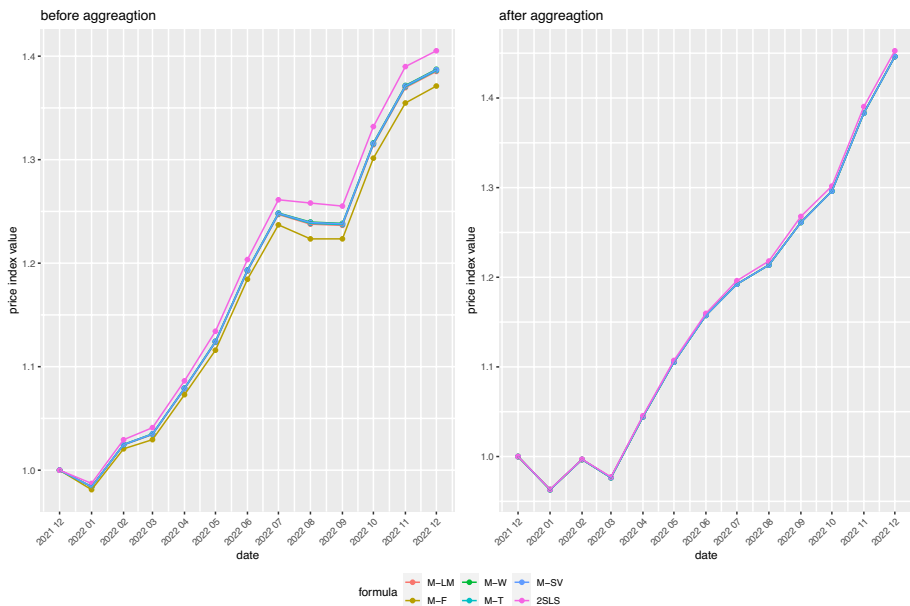


Fig. 4 CES indices based on elasticities of substitution calculated for other articles for personal hygiene

the differences, between the CES indices (the Lloyd-Moulton indices) are not noticeable. However, crucial differences between CES indices may occur with disaggregated input

data. At the level of the GTIN code, the estimated values of the CES index for other articles for personal hygiene differ by as much as 0.4 p.p. (June, 2022—see Fig. 4).

5 Conclusions

Both the *algebraic approach* and the *panel regression approach* are useful and valuable in estimating the elasticity of substitution. However, these approaches differ not only in the computational complexity or the set of data considered, but also their sensitivity to the choice of data filter or data aggregation level. Although the authors are aware that the results should be treated as preliminary, some observations regarding the methods of estimating the elasticity of substitution are presented as general conclusions listed below.

Firstly, algebraic methods generally lead to much less varied values of the estimated elasticity of substitution than methods based on a regression approach. As a rule, the maximum difference of the CES values obtained by algebraic methods does not exceed the level of 0.2–0.3 (Figs. 1 and 2). At the same time, methods based on panel regression lead to results that differ even several times regardless of the data filters used (Appendix A). Thus, one should be much more careful in choosing the regression rather than the algebraic method when estimating the elasticity of substitution based on scanner data. The algebraic **M-LM** method is an appropriate choice because it has the strongest theoretical foundation (see Sect. 3.1). Additionally, it generates stable estimates of the elasticity of substitution regardless of the data filter used. The **2SLS** method is the most useful because it allows us to consider endogeneity, product or group fixed effects, and the heterogeneity of the mentioned ones. We recognise that our estimates of CES elasticities using panel methods are not unweighted. Due to the lack of availability of relevant statistical data, we did not consider the disposable income of our consumer group, the prices of substitute and complementary goods, or volumes realised in other stores.

Secondly, for both algebraic and panel regression-based methods, data aggregation had a clear effect on the results of the elasticity of substitution. Moving from a disaggregated level (GTIN level) to a higher level of aggregation (COICOP level 6) each time led to a several-fold decrease in the value of the elasticity of substitution (Appendix B).

Thirdly, only for disaggregated data is there a fear that different estimates of the elasticity of substitution will lead to significant differences in the CES indices that are based on them. For data aggregated at the COICOP 6 level, the estimated substitution elasticities are lower, and the differences in these estimates obtained by various methods are also smaller. Therefore, the differences in the corresponding CES indices are also less important. The lack of product homogeneity within specific panels is also relevant here. However, the results are closer to reality than when working with aggregated data or broad product definitions.

It is worth adding at the end that the research results obtained can be important not only for the economy, but also for investment or pension fund managers and directors of banks or financial institutions. Better estimation of inflation is a profit for the state (indexation of pensions is based on the CPI) and for the banking sector (indexation of bank contracts is generally based on the inflation rate). Since the CES index does not require knowledge of the current level of consumption, proper estimation of the elasticity of substitution in effect can give more freedom to the governors of statistical offices in the organization of Household Budget Surveys (they can, for example, reduce their frequency and thus reduce the cost of this survey). Knowledge of the magnitude of the substitution effect of goods in

particular product segments can also be valuable to managers of retail chains, who can use this knowledge in determining the sales mix or in pricing policies (elasticity of substitution also shows how relative spending on goods changes when their relative prices change). In summary, the results obtained in the paper may find many potential beneficiaries. However, it is worth noting at this point that the survey we performed had certain limitations due to the availability of this type of data and also due to certain methodological assumptions we made. First, the study is limited to two sets of data from one retail chain. For this reason, the results obtained should be considered preliminary, and it is too early to generalize conclusions. Second, due to the multiplicity of data and hardware limitations, the study did not take into account the aggregation of results against the chain's outlets, i.e. the results from all outlets formed a single, consistent dataset. Third, the study focused on the post-pandemic period and perhaps, in the pre-pandemic period and almost certainly in the pandemic period, estimates of elasticity of substitution would have been quite different. Therefore, the authors intend to continue the study and expand it in the future by including more retail chains and new product groups.

Appendix A

See Tables 4, 5, 6, and 7.

Table 4 CES estimates on *yoghurts* before aggregation using different data filters and estimators

	Without	Dump price	Extreme price	Low sale
<i>Random effect (RE) estimator</i>				
<i>log(price)</i>	-0.731*** (0.034)	-0.726*** (0.032)	-0.678*** (0.033)	-0.695*** (0.033)
<i>constant</i>	3.804*** (0.091)	4.033*** (0.084)	3.911*** (0.085)	3.945*** (0.085)
Number of products	382	344	349	349
Number of observations	355	346	3497	35
<i>Fixed effect (FE) estimator</i>				
<i>log(price)</i>	-0.747*** (0.087)	-0.735*** (0.081)	-0.686*** (0.085)	-0.704*** (0.087)
<i>constant</i>	4.063*** (0.190)	4.166*** (0.176)	4.045*** (0.185)	4.085*** (0.189)
Number of products	382	344	349	349
Number of observations	355	346	3497	35
<i>First-difference (FD) estimator</i>				
<i>log(price)</i>	-0.398*** (0.047)	-0.388*** (0.046)	-0.260*** (0.046)	-0.318*** (0.046)
<i>constant</i>	-0.063*** (0.015)	-0.051*** (0.015)	-0.067*** (0.016)	-0.064*** (0.016)
Number of products	350	344	349	349
Number of observations	3139	3097	3134	3137

Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5 CES estimates on *other articles for personal hygiene* before aggregation using different data filters and estimators

	Without	Dump price	Extreme price	Low sale
<i>Random effect (RE) estimator</i>				
<i>log(price)</i>	−0.639*** (0.037)	−0.658*** (0.037)	−0.666*** (0.037)	−0.695*** (0.042)
<i>constant</i>	4.304*** (0.109)	4.664*** (0.108)	4.686*** (0.108)	4.947*** (0.124)
Number of products	567	516	516	447
Number of observations	5346	5155	5152	3136
<i>Fixed effect (FE) estimator</i>				
<i>log(price)</i>	−0.603*** (0.113)	−0.615*** (0.104)	−0.629*** (0.105)	−0.650*** (0.137)
<i>constant</i>	4.448*** (0.305)	4.729*** (0.278)	4.768*** (0.282)	5.176*** (0.372)
Number of products	567	516	516	447
Number of observations	5346	5155	5152	3136
<i>First-difference (FD) estimator</i>				
<i>log(price)</i>	−1.055*** (0.091)	−0.976*** (0.086)	−1.004*** (0.088)	−1.084*** (0.110)
<i>constant</i>	−0.016 (0.012)	−0.018 (0.012)	−0.018 (0.012)	−0.006 (0.016)
Number of products	516	516	516	440
Number of observations	4553	4553	4551	2616

Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 6 CES estimates on *yoghurts* after aggregation using different data filters and estimators

	Without	Dump price	Extreme price	Low sale
<i>Random effect (RE) estimator</i>				
<i>log(price)</i>	0.351 (0.348)	0.804*** (0.311)	0.794** (0.310)	0.793** (0.310)
<i>constant</i>	1.779** (0.903)	0.867 (0.809)	0.889 (0.806)	0.891 (0.806)
Number of products	5	5	5	5
Number of observations	65	65	65	65
<i>Fixed effect (FE) estimator</i>				
<i>log(price)</i>	0.632 (0.505)	1.102** (0.343)	1.088** (0.346)	1.088** (0.347)
<i>constant</i>	1.086 (1.245)	0.144 (0.832)	0.176 (0.839)	0.177 (0.841)
Number of products	5	5	5	5
Number of observations	65	65	65	65
<i>First-difference (FD) estimator</i>				
<i>log(price)</i>	0.296 (0.446)	0.725* (0.428)	0.782* (0.421)	0.773* (0.421)
<i>constant</i>	-0.023 (0.029)	-0.017 (0.027)	-0.019 (0.027)	-0.019 (0.027)
Number of products	5	5	5	5
Number of observations	60	60	60	60
<i>Two stage least squares (2SLS) estimator</i>				
<i>log(price)</i>	-5.618 (4.707)	-3.865 (4.927)	-4.087 (5.060)	-4.102 (5.097)
Number of products	5	5	5	5
Number of observations	65	65	65	65

Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 7 CES estimates on *other articles for personal hygiene* after aggregation using different data filters and estimators

	Without	Dump price	Extreme price	Low sale
<i>Random effect (RE) estimator</i>				
<i>log(price)</i>	−0.409*** (0.101)	−0.297*** (0.110)	−0.297*** (0.110)	−0.583*** (0.126)
<i>constant</i>	4.966*** (0.373)	4.981*** (0.378)	4.981*** (0.378)	5.863*** (0.398)
Number of products	7	7	7	7
Number of observations	91	91	91	83
<i>Fixed effect (FE) estimator</i>				
<i>log(price)</i>	−0.371*** (0.048)	−0.240** (0.072)	−0.240** (0.072)	−0.525*** (0.114)
<i>constant</i>	4.873*** (0.116)	4.849*** (0.166)	4.849*** (0.166)	5.827*** (0.266)
Number of products	7	7	7	7
Number of observations	91	91	91	83
<i>First-difference (FD) estimator</i>				
<i>log(price)</i>	−0.843*** (0.263)	−0.927*** (0.291)	−0.927*** (0.291)	−0.995*** (0.235)
<i>constant</i>	0.015 (0.027)	0.021 (0.028)	0.021 (0.028)	−0.001 (0.030)
Number of products	7	7	7	7
Number of observations	84	84	84	76
<i>Two stage least squares (2SLS) estimator</i>				
<i>log(price)</i>	−0.348*** (0.083)	0.001 (0.109)	0.001 (0.109)	−0.723** (0.308)
Number of products	7	7	7	7
Number of observations	91	91	91	83

Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Appendix B

See Table 8.

Table 8 Elasticities of substitution calculated before and after data aggregation for two considered data sets and two variants (yearly values vs mean of monthly values, non-filtered data)

Method	Before aggregation	After aggregation
<i>Yoghurt (Dec, 2021–Dec, 2022)</i>		
M-LM	1.3160	0.7763
M-F	1.3513	0.8593
M-W	1.3418	0.7714
M-T	1.3157	0.7617
M-SV	1.3320	0.7617
<i>Other articles for personal hygiene (Dec, 2021–Dec, 2022)</i>		
M-LM	0.5482	0.1770
M-F	0.6080	0.1857
M-W	0.4061	0.2059
M-T	0.4908	0.1681
M-SV	0.4208	0.1870
<i>Yoghurt (mean of monthly elasticities of substitution)</i>		
M-LM	1.8498	1.1744
M-F	2.4813	1.1355
M-W	1.7611	1.0321
M-T	1.8152	1.1392
M-SV	1.7787	1.0557
<i>Other articles for personal hygiene (mean of monthly elasticities of substitution)</i>		
M-LM	3.1939	0.7606
M-F	3.2221	0.7732
M-W	3.2334	0.7710
M-T	3.2198	0.7716
M-SV	3.2313	0.7713

Author contributions The contribution of each of the three co-authors to the creation of the article is equal (33.33(3) %), from the development of the concept, selection of data and methods of analysis, to editing the text.

Funding There is no funding concerning this paper.

Data availability There is no possibility to share the used data sets since there is no permission for that from the Statistics Poland (the data owner).

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arrow, K.J., Chenery, H.B., Minhas, B.S., Solow, R.M.: Capital-labor substitution and economic efficiency. *Rev. Econ. Stat.* **43**(3), 225–250 (1961)
- Australian Bureau of Statistics: Making Greater Use of Transactions Data to Compile the Consumer Price Index. Information Paper 6401.0.60.003, Canberra (2016)
- Balk, B.: On Curing the CPI's Substitution and New Goods Bias. Research Paper 0005, Statistics Netherlands (2000)
- Beck, G.W., Lein, S.M.: Price elasticities and demand-side real rigidities in micro data and in macro models. *J. Monet. Econ.* **115**, 200–212 (2020)
- Berry, S.T., Haile, P.A.: Foundations of demand estimation. In: *Handbook of Industrial Organization* vol. 4, pp. 1–62. Elsevier, Dordrecht (2021)
- Berry, S., Linton, O.B., Pakes, A.: Limit theorems for estimating the parameters of differentiated product demand systems. *Rev. Econ. Stud.* **71**(3), 613–654 (2004)
- Bialek, J.: Priceindices - a new R package for bilateral and multilateral price index calculations. *Stat. - Stat. Econ. J.* **36**(2), 122–141 (2021)
- Bialek, J., Beręsewicz, M.: Scanner data in inflation measurement: from raw data to price indices. *Stat. J. IAOS* **37**, 1315–1336 (2021)
- Birchall, C., Verboven, F.: Estimating substitution patterns and demand curvature in discrete-choice models of product differentiation. Discussion Papers DP16981 (2022)
- Browning, M., Carro, J.: Heterogeneity and microeconometrics modelling. In: *Advances in Economics and Econometrics. Theory and Applications: Ninth World Congress of the Econometric Society*, vol. 3, pp. 47–74. Cambridge University Press, Cambridge (2007)
- Cameron, A.C., Trivedi, P.K.: *Microeconometrics Using Stata*, vol. 2. Stata Press, College Station (2010)
- Can, T.: Challenging the CES assumption with scanner data - pitfalls of the fixed basket. In: *Paper Presented at the 17th Meeting of the Ottawa Group on Price Indices*, 6–10 June 2022, Rome (2022)
- Chamberlain, G.: Asymptotic efficiency in estimation with conditional moment restrictions. *J. Econom.* **34**(3), 305–334 (1987)
- Cragg, J.G., Donald, S.G.: Testing identifiability and specification in instrumental variable models. *Econom. Theory* **9**(2), 222–240 (1993)
- Diewert, W.E.: Hedonic regressions. a consumer theory approach. In: *Scanner Data and Price Indexes*, pp. 317–348. University of Chicago Press, Chicago (2003)
- Diewert, W.E.: Exact and superlative index numbers. *J. Econom.* **4**(2), 115–145 (1976)
- Diewert, W.E., Fox, K.J.: Substitution bias in multilateral methods for cpi construction. *J. Bus. Econ. Stat.* **40**(1), 355–369 (2022)
- Eurostat: Practical Guide for Processing Supermarket Scanner Data. Harmonised Index of Consumer Prices (2018)
- Fisher, I.: *The Making of Index Numbers: A Study of Their Varieties, Tests, and Reliability*, vol. XXXI, p. 526. Houghton Mifflin, Boston (1922)
- Freyberger, J.: Asymptotic theory for differentiated products demand models with many markets. *J. Econom.* **185**(1), 162–181 (2015)
- Gordon, R.: The boskin commission report and its aftermath. *Monet. Econ. Stud.* **17**, 41–68 (1999)
- Haan, J., Balk, B., Hansen, C.: Retrospective approximations of superlative price indexes for years where expenditure data is unavailable. In: *Price Indexes in Time and Space*, pp. 25–42. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-7908-2140-6_2
- Hill, R.: Superlative index numbers: not all of them are super. *J. Econom.* **130**(1), 25–43 (2006)
- International Labour Office: International Monetary Fund, Organisation for Economic Co-operation and Development, Statistical Office of the European Communities, United Nations, The World Bank: *Consumer Price Index Manual: Theory and Practice*. International Labour Organization, Geneva (2004)
- International Labour Office: International Monetary Fund, Organisation for Economic Co-operation and Development, Statistical Office of the European Union, The World Bank, The United Nations Economic Commission for Europe: *Consumer Price Index Manual, 2020: Concepts and Methods*. Washington D.C., United States of America, International Monetary Fund (2020)
- Ivancic, L., Diewert, E., Fox, K.J.: Using a Constant Elasticity of Substitution Index to Estimate a Cost of Living Index: From Theory to Practice. Discussion Papers 2010-15, School of Economics, The University of New South Wales (Oct 2010)
- Jaro, M.: Advances in record-linkage methodology as applied to matching the 1985 census of Tampa, Florida. *J. Am. Stat. Assoc.* **84**(406), 414–420 (1989). <https://doi.org/10.1080/01621459.1989.10478785>
- Jeon, Y., Hoang, H., Thompson, W., Abler, D.: A meta-analysis of us food demand elasticities to detect the impacts of scanner data. *Appl. Econ. Perspect. Policy* (2023)

- Konüs, A.A.: The problem of the true index of the cost of living. *Econ. Bull. Inst. Econ. Conjunct.* **9–10**, 64–71 (1924)
- Laspeyres, K.: IX. die berechnung einer mittleren waarenpreissteigerung. *Jahrbücher für Nationalökonomie und Statistik* **16**(1), 296–318 (1871)
- Leclair, M., Léonard, I., Rateau, G., Sillard, P., Varlet, G., Vernédal, P.: Scanner data: advances in methodology and new challenges for computing consumer price indices. *Economie et Statistique* **509**(1), 13–29 (2019)
- Lent, J.V., Dorfman, A.: Using a weighted average of base period price indexes to approximate a superlative index. *J. Off. Stat.* **25**, 139–149 (2009)
- Liesbeth, C., Melo, P.C., Abdul-Salam, Y., Roberts, D., Mary, S., Gomez y Paloma, S.: Income elasticities for food, calories and nutrients across Africa: a meta-analysis. *Food Policy* (2018). <https://doi.org/10.1016/j.foodpol.2018.04.002>
- Lloyd, P.: Substitution effects and biases in nontrue price indices. *Am. Econ. Rev.* **65**(3), 301–313 (1975)
- Moulton, B.R.: Constant Elasticity Cost-of-Living Index in Share Relative Form. Bureau of Labor Statistics, Washington (1996)
- Newey, W.K.: Efficient instrumental variables estimation of nonlinear models. *Econometrica: Journal of the Econometric Society*, 809–837 (1990)
- Paasche, H.: Über die Preisentwicklung der letzten Jahre nach den Hamburger Borsennotirungen. *Jahrbücher für Nationalökonomie und Statistik* **12**(1), 168–178 (1874)
- Petrin, A., Train, K.: A control function approach to endogeneity in consumer choice models. *J. Mark. Res.* **47**(1), 3–13 (2010)
- Rilstone, P.: Efficient instrumental variables estimation of nonlinear dependent processes. *J. Nonparametric Stat.* **4**(1), 103–106 (1994)
- Saito, T.: How do we get cobb-douglas and leontief functions from ces function: a lecture note on discrete and continuum differentiated object models. *J. Ind. Organ. Educ.* **6**(1), 0000101515193550411037 (2012)
- Sanderson, E., Windmeijer, F.: A weak instrument f-test in linear iv models with multiple endogenous variables. *J. Econom.* **190**(2), 212–221 (2016)
- Sato, K.: The ideal log-change index number. *Rev. Econ. Stat.* **58**(2), 223–228 (1976)
- Sillard, P.: Scanner data: Towards constant utility indexes? Discussion Papers 2012, Institut National de la Statistique et des Études Économiques, Division des prix à la consommation (2012). https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2012/France_Geneva-Sillard-Eng.pdf
- Stock, J.H., Wright, J.H.: Gmm with weak identification. *Econometrica* **68**(5), 1055–1096 (2000)
- Stock, J.H., Yogo, M.: Testing for weak instruments in linear IV regression. National Bureau of Economic Research Cambridge, Mass (2002)
- Tianqi, C., Carlo, G.: Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, August 13–17, 2016. ACM, pp. 785–794 (2016). <https://doi.org/10.1145/2939672.2939785>
- Törnqvist, L.: The Bank of Finland's Consumption Price Index. *Bank of Finland Monthly Bulletin*, 1–8 (1936)
- van Loon, K.V., Roels, D.: Integrating big data in the Belgian CPI. In: Paper Presented at the Meeting of the Group of Experts on Consumer Price Indices, 8–9 May 2018, Geneva (2018)
- Vartia, Y.O.: Ideal log-change index numbers. *Scand. J. Stat.* **3**, 121–126 (1976)
- von der Lippe, P.: Index Theory and Price Statistics. Peter Lang, Berlin (2007)
- Walsh, C.M.: The Measurement of General Exchange Value. Macmillan and Co., New York (1901)
- White, A.G.: Measurement biases in consumer price indexes. *Int. Stat. Rev.* **67**(3), 301–325 (1999)

Artykuł 4

Artykuł 4 The use of transitive Montgomery indicators for scanner data analysis

opublikowany w

Argumenta Oeconomica, 53(2), 2024

Wkład współautorów:

Natalia Pawelec: 80%

Jacek Białek: 20%

Wkład własny:

Autorka przeprowadziła całość analizy empirycznej: obliczenie i porównanie bilateralnych oraz multilateralnych indyktorów Montgomery'ego w dwóch wariantach doboru produktów (produkty dopasowane i wszystkie dostępne), z zastosowaniem trzech rodzajów filtrów danych i na dwóch poziomach agregacji (GTIN i COICOP 6), w środowisku R z wykorzystaniem pakietu *PriceIndices*. Autorka opracowała tabele i rysunki wynikowe oraz przygotowała i zredagowała manuskrypt artykułu.

The use of transitive Montgomery indicators for scanner data analysis

Jacek Białek

Department of Statistical Methods, University of Lodz, Poland

e-mail: jacek.bialek@uni.lodz.pl

ORCID: [0000-0002-0952-5327](https://orcid.org/0000-0002-0952-5327)

Natalia Pawelec

Department of Economic Mechanisms, University of Lodz, Poland

e-mail: natalia.pawelec@uni.lodz.pl

ORCID: [0009-0008-5289-6204](https://orcid.org/0009-0008-5289-6204)

©2024 Jacek Białek, Natalia Pawelec

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>

Quote as: Białek, J., & Pawelec, N. (2024). The use of transitive Montgomery indicators for scanner data analysis. *Argumenta Oeconomica*, 2(53), 1-13.

DOI: [10.15611/aoe.2024.2.01](https://doi.org/10.15611/aoe.2024.2.01)

JEL: C43, E31

Abstract: Although the modern price index theory is based on an analysis of ratios of prices and quantities, one may be often more interested in working with differences in these values in many economic areas, e.g.: revenue change decompositions, profit and cost change decompositions, or an analysis of changes in consumer surplus. The benefit of using these differences is that there is no problem associated with the occurrence of zero prices and quantities, a problem that arises when working with ratios. In practice, one mostly cares about decomposing the value difference into indicators of contributions from price and quantity differences. The well-known price and quantity indicators are the Bennet and the Montgomery indicators, which are not transitive. This paper revises the price and quantity Montgomery indicators and their multilateral versions for the analysis of scanner data. Specifically, instead of considering ‘classical’ comparisons across firms, countries or regions, the transitive versions of the Montgomery indicators were adapted to work on scanner data sets observed over a fixed time window. One of the objectives of the study was to compare bilateral and multilateral Montgomery indicator values for different data aggregation levels and three main types of data filters. To the best of the authors’ knowledge, this study is pioneering on the grounds of implementing the multilateral Montgomery indicators in scanner data analysis.

Keywords: scanner data, Montgomery indicators, transitivity, multilateral indicators

1. Introduction

The contemporary price index theory is based on comparisons of ratios of prices, quantities and expenditures of goods and services (von der Lippe, 2007; International Labour Office, 2004; International Monetary Fund, 2020). These index numbers are used to build various economic measures, such as gross domestic product (GDP), producer price index (PPI) and consumer price index (CPI). Nevertheless, in many business contexts one may be more interested in the magnitude of differences in prices, quantities and sale values. This may concern many economic areas, including revenue change decompositions, profit and cost change decompositions, or the analysis of changes in consumer surplus (Diewert, 2005). An important benefit of using such differences is that there is no problem associated with the occurrence of zero prices and quantities, a problem that arises when working with ratios. The issue of zero prices or zero quantities may be very serious in many business applications where not all goods are produced and purchased in every period.

Zero prices and quantities are also common in scanner data sets due to the high product churn observed in supermarkets. Scanner data, which support CPI calculations in many countries, mean transaction data that specify turnover and numbers of items sold by barcodes, e.g. Global Trade Item Number (GTIN), European Article Number (EAN), and Stock Keeping Unit (SKU) codes. Scanner data have numerous advantages compared to traditional survey data collection because such data sets are much bigger than traditional ones and they contain complete transaction information, namely information about prices and quantities at barcode level and also many other product characteristics (i.e. product weight sales unit, product label, VAT level, etc.). Nevertheless, at barcode level, the occurrence of zero-scanner prices is quite common because supermarkets often offer a variety of discounts and price reductions (even down to zero), as well as a significant turnover of products. This causes analytical problems both for statistical offices and for supermarket owners wishing to compare the sales performance of different product segments in two time periods. Thus, an approach based on the differences in price values, quantities and expenditure can also be very useful in analysing scanner data.

The difference-based approach to index numbers is well established in the economic literature, where it was introduced in the early 20th century (Bennet, 1920; Montgomery, 1929). Note that index numbers, expressed in terms of differences, are referred to as indicators (Diewert, 2005). Recently, one can witness a return of interest in this approach on the part of statisticians and economists (Balk *et al.*, 2004; Diewert, 2005; Fox, 2006; Cross and Färe, 2009; de Boer and Rodrigues, 2020). However, to the best of the authors' knowledge, there is a lack of studies that apply the indicators to the analysis of scanner data, which is the main objective of this paper focused on the Montgomery indicators because they receive much less attention in the literature than the Bennet indicators.

The added value of the paper and its contribution to the existing knowledge about indicators can be summarised as follows: (1) the Montgomery indicators and their transitive versions defined for comparisons across companies (or regions) in cross-section or panel context are adopted to work on scanner data sets observed over a fixed time window; (2) the axiomatic properties of the Montgomery multilateral indicators are verified; (3) variants of the Montgomery indicators based on matched samples are considered, which may be more accurate due to the high turnover of scanner products; (4) the impact of data filters used and the level of data aggregation on the price and quantity Montgomery indicators is also examined.

The structure of the paper is as follows: Section 2 presents the bilateral Montgomery indicators, Section 3 adopts the transitive Montgomery indicators from the field of comparisons between companies and examines their axiomatic properties, Section 4 presents the results of the empirical study in which the bilateral and multilateral Montgomery indicators were applied to the analysis of scanner data, and Section 5 lists the most important conclusions drawn from the empirical study.

2. The Montgomery indicators

The axiomatic and economic index theory is derived from Fisher (1922) and Konús (1939), respectively. Although until the end of the 20th century there was a lack of studies providing a foundation for the construction of price and quantity indices, more recently there has been a marked interest among statisticians and economists in the indicators theory. For instance, Chambers (2001) proposed a new economic framework for indicators by using Diewert's (1976) quadratic lemma. Balk *et al.* (2004) developed the theory of economic price and quantity indicators by deriving an exact relation between indicators and directional distance functions. Finally, in the paper by Diewert (2005), an additive test approach was developed.

As a rule, price and quantity indicators are calculated using firm-level price and quantity data. Authors use the context of the production theory or concentrate on the input side of companies or regions (Balk *et al.*, 2004; Cross and Färe, 2009; Fox, 2006). To the best of the authors' knowledge, this study is pioneering in the area of implementing the multilateral Montgomery indicators in scanner data analysis. In this section, the bilateral Montgomery indicator formula is expressed with the additional distinction between available and matched products. The above-mentioned distinction does matter in the case of scanner data due to the high product churn. Section 3 goes a step further by proposing a multilateral approach, where the transitive version of the Montgomery indicators takes the entire given time window into account.

2.1. The case of a matched product sample

Due to the high turnover of scanner products, a frequently used approach in determining price indices is the matched sample approach. In order to correctly define the Montgomery indicators in the matched sample approach, first enter the necessary notation.

Denote sets of homogeneous products which belong to the same product groups (category) in months 0 and t by G_0 and G_t respectively, with $G_{0,t}$ denoting a set of matched products in both considered periods, i.e. $G_{0,t} = G_0 \cap G_t$. Let p_i^τ and q_i^τ denote the price (more precisely: *unit value*) and quantity of product i at time $\tau \in \{0, t\}$, which are assumed to be positive. Note that the *Vartia mean* (Vartia, 1976) of two positive real numbers x and y , which is known in the mathematical literature as the *logarithmic mean* (von der Lippe, 2007), is defined as below.

$$L(x, y) = \begin{cases} \frac{x-y}{\ln(x)-\ln(y)}; & x \neq y, \\ x; & x = y. \end{cases} \quad (1)$$

Under the above, the bilateral Montgomery price indicators, which compare current period t to base period 0 and consider only the matched products, can be defined as follows:

$${}_mIP_{0,t}^M = \sum_{i \in G_{0,t}} L(p_i^t q_i^t, p_i^0 q_i^0) \ln\left(\frac{p_i^t}{p_i^0}\right), \quad (2)$$

and the bilateral Montgomery quantity indicator can be expressed as follows:

$${}_mIQ_{0,t}^M = \sum_{i \in G_{0,t}} L(p_i^t q_i^t, p_i^0 q_i^0) \ln\left(\frac{q_i^t}{q_i^0}\right). \quad (3)$$

It is easy to show (Montgomery, 1929) that the consumer's value change over the two periods under consideration, observed for the matched products, can be decomposed by using the Montgomery price and quantity indicators, i.e.

$$V_{G_{0,t}}^t - V_{G_{0,t}}^0 \equiv \sum_{i \in G_{0,t}} p_i^t q_i^t - \sum_{i \in G_{0,t}} p_i^0 q_i^0 = {}_mIP_{0,t}^M + {}_mIQ_{0,t}^M. \quad (4)$$

2.2. The case of unmatched products

Note that in the case of scanner data, the set of products available in the base and current periods, i.e. $G_0^t = G_0 \cup G_t$, is generally an oversampling of the set of matched products $G_{0,t} = G_0 \cap G_t$. Regarding the Bennet indicators, this does not generate technical problems in their calculation since it is simply assumed that a product that is available in period t_1 , but is not available in period t_2 , has zero price and quantity in period t_2 . Such a procedure does not apply to the Montgomery indicators since they use logarithms of prices and quantities in their syntax. However, it can be shown (see Appendix 1) that prices and quantities of mismatched products in the periods being compared can effectively be determined in line with the following procedure:

$$p_i^t = q_i^t = \varepsilon, \text{ for } i \in G_0 \setminus G_t, \text{ and } p_i^0 = q_i^0 = \varepsilon, \text{ for } i \in G_t \setminus G_0, \quad (5)$$

which holds for a sufficiently small, real number $\varepsilon > 0$.

With the above-presented procedure for treating mismatched products, analogously to Section 2.1, one can define the Montgomery indicators for all available products (including mismatched products) as follows:

$$IP_{0,t}^M = \sum_{i \in G_0^t} L(p_i^t q_i^t, p_i^0 q_i^0) \ln \left(\frac{p_i^t}{p_i^0} \right), \quad (6)$$

$$IQ_{0,t}^M = \sum_{i \in G_0^t} L(p_i^t q_i^t, p_i^0 q_i^0) \ln \left(\frac{q_i^t}{q_i^0} \right). \quad (7)$$

Note that in general, when $G_0^t \neq G_{0,t}$, the Montgomery indicators (6) and (7) are functions of parameter ε (under assumption (5)), and thus they can be denoted by $IP_{0,t}^M(\varepsilon)$ and $IQ_{0,t}^M(\varepsilon)$. The observation in Appendix 1 leads to the conclusion that the Montgomery indicators satisfy the *sum test* (otherwise known as the *value change test*) with a good approximation if procedure (5) is taken into consideration, i.e.

$$V_{G_0^t}^t - V_{G_0^t}^0 \equiv \sum_{i \in G_t \setminus G_0} p_i^t q_i^t - \sum_{i \in G_0 \setminus G_t} p_i^0 q_i^0 + \sum_{i \in G_{0,t}} p_i^t q_i^t - \sum_{i \in G_{0,t}} p_i^0 q_i^0 = \lim_{\varepsilon \rightarrow 0^+} (IP_{0,t}^M(\varepsilon) + IQ_{0,t}^M(\varepsilon)). \quad (8)$$

Remark 1. In the empirical study (see Section 4), the *PriceIndices* package (Bialek, 2022) is used, in which the implementation of the Montgomery indicators takes $\varepsilon = 0.000001$. Such a small value of parameter ε means that the *sum test* is approximately satisfied in the case of these indicators. However, it should be noted that for the static item universe, i.e. when $G_0^t = G_{0,t}$, the *sum test* is fully satisfied and no proxy is needed (Diewert and Mizobuchi, 2009). One should also keep in mind that the Montgomery indicators satisfy most of the required tests except for the *monotonicity tests* and some *symmetry tests*. Those interested in the mathematical properties of the Montgomery indicators are referred to Diewert (2005).

3. The multilateral Montgomery indicators

For international or inter-regional comparisons, *transitivity* means that estimates of price dynamics and quantities of selected attributes do not depend on the choice of the underlying country or region. Similarly, for comparisons across companies, computing transitive price and quantity indices (indicators) do not depend on the choice of the company benchmark. A lack of index transitivity is a well-known problem in the literature on international comparisons and scanner data (Gini, 1931; Eltetö and Köves, 1964; Szulc, 1964; Ivancic *et al.*, 2011; Chessa, 2016). By definition, *transitivity* eliminates the chain drift problem. The chain drift can be formalised in terms of the violation of the *multi-period identity test*, according to which one can expect that when all prices and quantities in a current period revert back to their values from the base period, the index should indicate no price change and equal

one. In the same situation, the price indicator should equal zero. Even chain superlative indices may suffer from chain drift (Białek and Roszko-Wójtowicz, 2023). Multilateral price indices are *transitive*, which means that the calculation of the price dynamics for any two moments in the time window does not depend on the choice of the base period (Chessa, 2016). Thus, this section focuses on forming the multilateral Montgomery indicators.

In the case of any price indicator IP and quantity indicator IQ , *transitivity* means that the following relations occur for any $0 < s < t$:

$$IP_{0,s} + IP_{s,t} = IP_{0,t} \text{ and } IQ_{0,s} + IQ_{s,t} = IQ_{0,t}. \quad (9)$$

It is easy to verify that bilateral Montgomery indicators are not transitive. In fact, using the *coffee* data set from the *PriceIndices* R package (Białek, 2022), taking *January 2019* as base period 0, *February 2019* as internal period s , and *March 2019* as base period t , and running the *Montgomery()* function, one obtains: $IP_{0,s}^M + IP_{s,t}^M = -49493.63 \neq IP_{0,t}^M = -26389.21$. The following design of the multilateral Montgomery indicators is an adaptation of Fox's (2006) transformation for the scanner data case, considering some time windows with more than two periods instead of some number of companies or regions.

Let us denote by $[0, T]$ the considered time interval, which typically (while computing multilateral indices) consists of 13 or 25 months (Eurostat, 2022), while $G_{[0,T]}$ denotes the set of available products in the whole interval $[0, T]$, i.e. $G_{[0,T]} = \bigcup_{\tau=0}^T G_\tau$, and let $G_{[0,T]}^m$ denote the set of matched products in the whole interval $[0, T]$, i.e. $G_{[0,T]}^m = \bigcap_{\tau=0}^T G_\tau$. Thus one can introduce the additional notations:

$${}_T IP_{\tau,t}^M = \sum_{i \in G_{[0,T]}} L(p_i^t q_i^t, p_i^\tau q_i^\tau) \ln \left(\frac{p_i^t}{p_i^\tau} \right), \quad (10)$$

$${}_{mT} IP_{\tau,t}^M = \sum_{i \in G_{[0,T]}^m} L(p_i^t q_i^t, p_i^\tau q_i^\tau) \ln \left(\frac{p_i^t}{p_i^\tau} \right), \quad (11)$$

$${}_T IQ_{\tau,t}^M = \sum_{i \in G_{[0,T]}} L(p_i^t q_i^t, p_i^\tau q_i^\tau) \ln \left(\frac{q_i^t}{q_i^\tau} \right), \quad (12)$$

$${}_{mT} IQ_{\tau,t}^M = \sum_{i \in G_{[0,T]}^m} L(p_i^t q_i^t, p_i^\tau q_i^\tau) \ln \left(\frac{q_i^t}{q_i^\tau} \right). \quad (13)$$

It should be noted that the previously assumed procedure (5) still applies to formulas (10) and (12). Defining a multilateral Montgomery indicator for all available products from a time window requires first introducing the following averaged price and quantity indicators:

$$I\bar{P}_{t_0}^M = \frac{1}{T+1} \sum_{\tau=0}^T {}_T IP_{\tau,t_0}^M, \quad (14)$$

$$I\bar{Q}_{t_0}^M = \frac{1}{T+1} \sum_{\tau=0}^T {}_T IQ_{\tau,t_0}^M. \quad (15)$$

Now, using (14) and (15), the price and quantity multilateral Montgomery indicator can be defined respectively:

$$MIP_{0,t}^M = I\bar{P}_t^M - I\bar{P}_0^M, \quad (16)$$

$$MIQ_{0,t}^M = I\bar{Q}_t^M - I\bar{Q}_0^M. \quad (17)$$

Note that the multilateral Montgomery indicators (16) and (17) are *transitive*, i.e.

$$MIP_{0,s}^M + MIP_{s,t}^M = I\bar{P}_s^M - I\bar{P}_0^M + I\bar{P}_t^M - I\bar{P}_s^M = I\bar{P}_t^M - I\bar{P}_0^M = MIP_{0,t}^M, \quad (18)$$

$$MIQ_{0,s}^M + MIQ_{s,t}^M = I\bar{Q}_s^M - I\bar{Q}_0^M + I\bar{Q}_t^M - I\bar{Q}_s^M = I\bar{Q}_t^M - I\bar{Q}_0^M = MIQ_{0,t}^M. \quad (19)$$

These multilateral Montgomery indicators satisfy the *time reversal test* since (20) and (21) hold:

$$MIP_{0,t}^M = I\bar{P}_t^M - I\bar{P}_0^M = -(I\bar{P}_0^M - I\bar{P}_t^M) = -MIP_{t,0}^M, \quad (20)$$

$$MIQ_{0,t}^M = I\bar{Q}_t^M - I\bar{Q}_0^M = -(I\bar{Q}_0^M - I\bar{Q}_t^M) = -MIQ_{t,0}^M. \quad (21)$$

It can be proven that the multilateral Montgomery indicators also satisfy the *value change test* and the *multi-period identity test* (see Appendix 2).

Remark 2. Let us conclude this section by pointing out that all the axiomatic properties valid for the multilateral Montgomery indicators calculated on the set of all available products $G_{[0,T]}$ also carry over to the analogous multilateral Montgomery indicators estimated on the set of matched products $G_{[0,T]}^m$. The transitive Montgomery indicators based on matched products can be written as

$${}_mMIP_{0,t}^M = \frac{1}{T+1} \sum_{\tau=0}^T ({}_mTIP_{\tau,t}^M - {}_mTIP_{\tau,0}^M), \quad (22)$$

$${}_mMIQ_{0,t}^M = \frac{1}{T+1} \sum_{\tau=0}^T ({}_mTIQ_{\tau,t}^M - {}_mTIQ_{\tau,0}^M), \quad (23)$$

where, similarly to the proof from **Appendix 2**, it can be proven that

$${}_mMIP_{0,t}^M + {}_mMIQ_{0,t}^M = V_{G_{[0,T]}^m}^t - V_{G_{[0,T]}^m}^0. \quad (24)$$

4. Empirical study

This empirical study utilised scanner data obtained from a single retail chain operating in Poland. Specifically, the authors analysed monthly data pertaining to *stationery and hygiene products* (COICOP 5: 121322), taken from December 2021 to December 2022 and include more than 500 outlets. The COICOP 5 product group consists of the following local COICOP 6 product subgroups: tissues (60 products, i.e. 66 product identifiers (IDs)), wet wipes (88 IDs), toilet paper (117 IDs), baby diapers (193 IDs), sanitary pads (20 IDs), sanitary napkins (67 IDs), and tampons (22 IDs).

To proceed with the calculations, the data had to be properly prepared, and functions from the *PriceIndices* R package (Białek, 2021) were used for this purpose. One of the first functions was *data_selecting()*, which required the creation of dictionaries containing the relevant keywords and phrases that were necessary to identify distinct product groups. The second function – *data_classification()* was utilised to handle problematic products that were previously unclassified, necessitating the manual preparation of learning samples using historical data. Next, product matching was carried out by leveraging available Global Trade Item Number (GTIN) barcodes, internal retail chain codes, and product labels. The *PriceIndices* package facilitated this task through the use of the *data_matching()* function.

Different variants were considered for comparing the sales value difference: the Montgomery price indicator (referred to as the “price effect” in the figure) and the Montgomery quantity indicator (referred to as the “quantity effect”). Calculations were performed to analyse and assess these indicators across the various variants.

Figures 1 and 2 show the bilateral versions of the Montgomery indicators, while Figures 3 and 4 present the multilateral versions which are equivalent to the first two figures. The bilateral indicators, similar to the multilateral indicators, were examined in two scenarios: one without filtering the original data (as depicted in Figures 1 and 3), and the other with the implementation of a low sales filter using parameter $\lambda = 1.25$ (Loon and Roels, 2018).

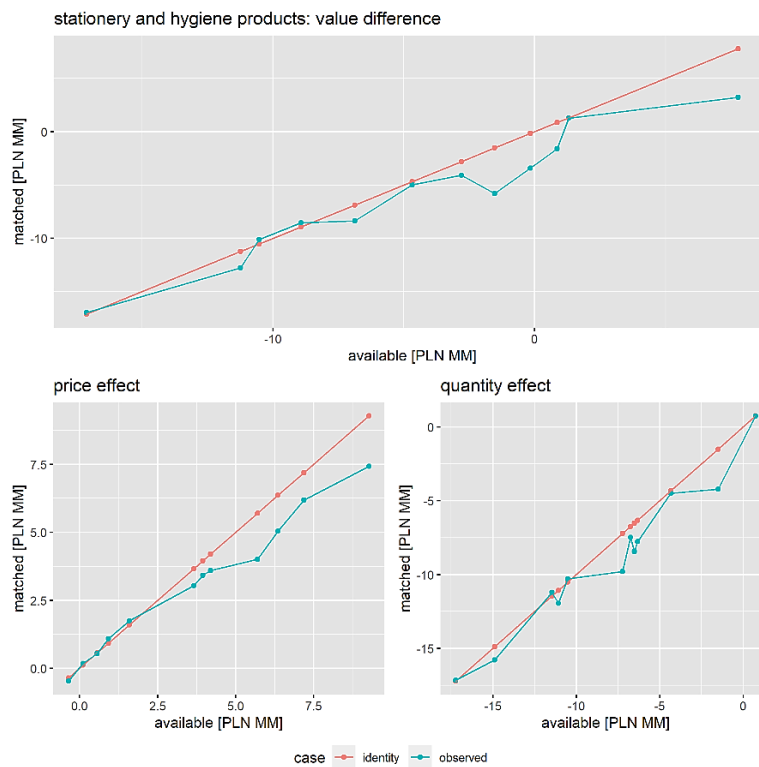


Fig. 1. Comparison of the difference in the value of sales and the bilateral Montgomery indicators calculated for all available products and for matched products (no data filters applied)

Source: own calculations in the *PriceIndices* R package.

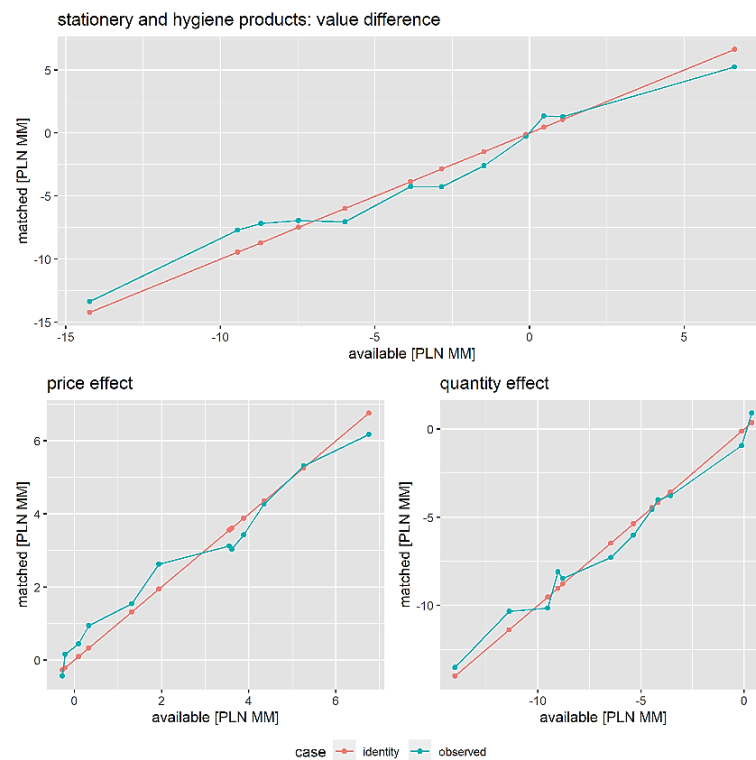


Fig. 2. Comparison of the difference in the value of sales and the bilateral Montgomery indicators calculated for all available products and for matched products (data filters applied)

Source: own calculations in the *PriceIndices* R package.

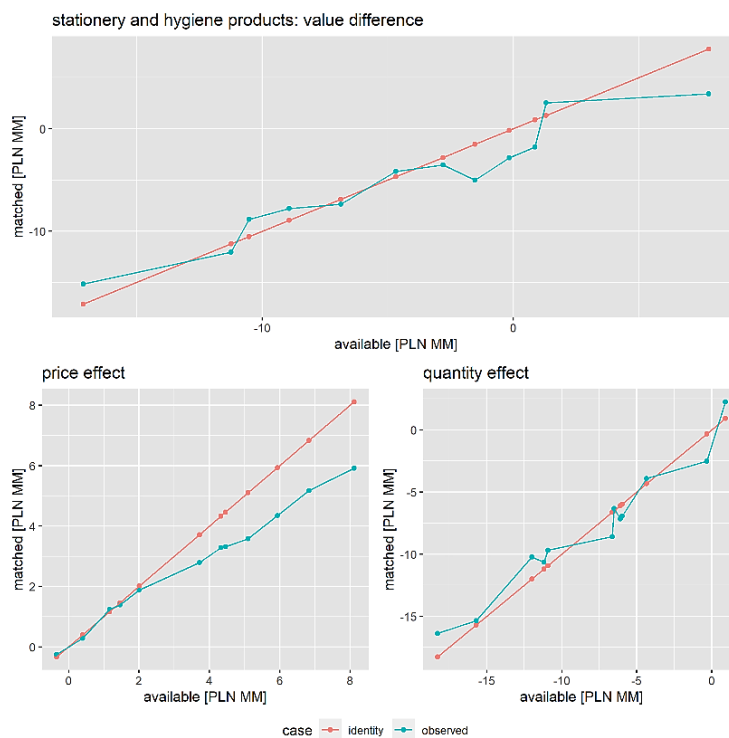


Fig. 3. Comparison of the difference in the value of sales and the multilateral Montgomery indicators calculated for all available products and for matched products (no data filters applied)

Source: own calculations in the *PriceIndices* R package.

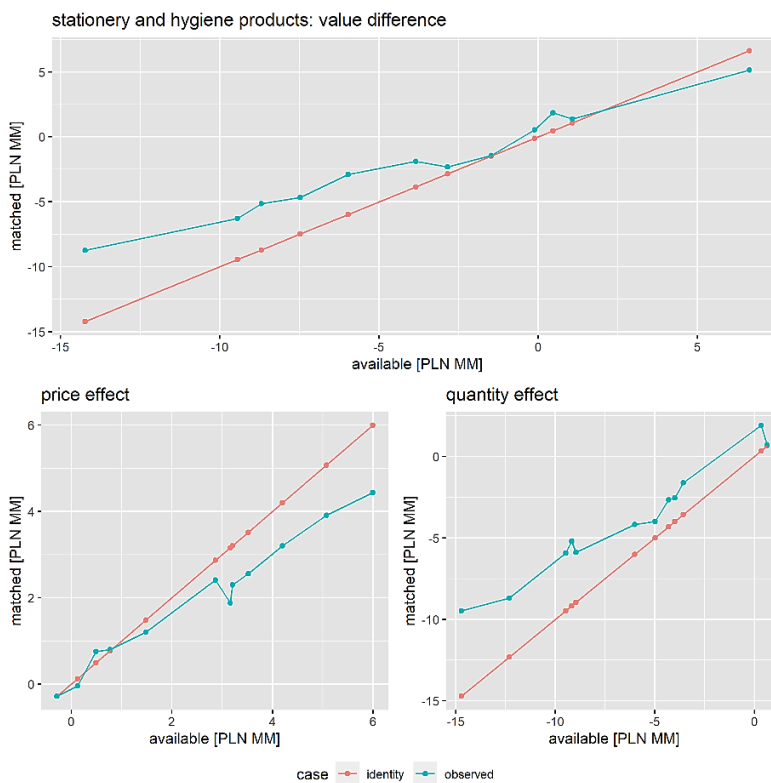


Fig. 4. Comparison of the difference in the value of sales and the multilateral Montgomery indicators calculated for all available products and for matched products (data filters applied)

Source: own calculations in the *PriceIndices* R package.

On the OX axis, the values calculated in ascending order for all months within the analysed time window, representing the characteristics under study (value difference, price or quantity effect), were plotted for all available products within that time frame. On the OY axis, similar values for the corresponding months were indicated, but this time only the products that match the criteria were included.

The presented figures can be interpreted as follows: if the green line (called “empirical”) is below the red line (called “identity”), this means that product matching caused a decrease in the values of the characteristics under study. The greater the deviation between the empirical line and the theoretical one, the stronger the effect mentioned above.

In the bilateral approach, the impact of data filtering appeared to go in the opposite direction than in the multilateral approach. For the bilateral approach, one can see that data filtering decreases the product matching effect (Figures 1 and 2). At the same time, data filtering increases the product matching effect in the multilateral approach (Figures 3 and 4). In other words, for the Montgomery’s bilateral indicators, after filtering, smaller differences were observed between the results of price and quantity indicators calculated for matched and available products. In contrast, the differences were greater when filtering was applied to multilateral indicators (see Figure 4). In this case, the red and green lines in the figure were clearly separated, indicating that the matched and unmatched approaches generated markedly different scores for price and quantity indicators and, consequently, for the total value change. It can be further noted that the price effect was weakened after filtering in the multilateral approach when only matched products were included, but simultaneously the quantity effect was strengthened.

To be more specific, the Montgomery indicators’ values (in all the versions) were calculated for the product precisely defined at two levels: the narrow level using the GTIN code and the broader level using COICOP 6 classification. For a better and more accurate view, a series of comparisons were made for the current period set at the end of the time window (December 2022). Their results are shown in Tables 1 and 2.

Table 1. Comparison of the Montgomery indicators across the data aggregation level and data filtering (all the available products were considered, the normalised values are in brackets)

GTIN level: bilateral approach				
Characteristics	No data filtering		With data filtering	
Sales value difference	7769806.28	(100)	6624556.59	(100)
Price Montgomery indicator	9275020.81	(119.37)	6755626.25	(101.98)
Quantity Montgomery indicator	-1505214.53	(-19.37)	-131069.66	(-1.98)
GTIN level: multilateral approach				
Sales value difference	7769806.28	(100)	6624556.59	(100)
Price Montgomery indicator	8110048.49	(104.38)	5990950.99	(90.44)
Quantity Montgomery indicator	-340242.21	(-4.38)	633605.60	(9.56)
COICOP 6 level: bilateral approach				
Characteristics	No data filtering		With data filtering	
Sales value difference	7769806.28	(100)	6624556.59	(100)
Price Montgomery indicator	9895893.58	(127.36)	8018053.51	(121.04)
Quantity Montgomery indicator	2126087.30	(-27.36)	-1393496.92	(-21.04)
COICOP 6 level: multilateral approach				
Characteristics	No data filtering		With data filtering	
Sales value difference	7769806.28	(100)	6624556.59	(100)
Price Montgomery indicator	8298307.45	(106.8)	6743673.56	(101.8)
Quantity Montgomery indicator	-528501.17	(-6.8)	-119116.97	(-1.8)

Source: own calculations in the *PriceIndices* R package.

Table 2. Price and quantity contributions across the bilateral and multilateral approach (all the available products were considered, the normalised values are in brackets: total sales value difference = 100)

Product contributions: bilateral approach						
COICOP 6 subgroup	Sales value difference		Price contributions		Quantity contributions	
Tissues	2423488.86	(36.58)	1745602.68	(26.35)	677886.18	(10.23)
Toilet paper	-149921.79	(-2.26)	91729.22	(1.38)	-241651.01	(-3.65)
Baby diapers	3200407.98	(48.31)	5897172.95	(89.02)	-2696764.97	(-40.71)
Sanitary pads	1142085.37	(17.24)	207101.43	(3.13)	934983.94	(14.11)
Sanitary napkins	112069.03	(1.69)	45314.16	(0.68)	66754.87	(1.01)
Tampons	-103572.86	(-1.56)	31133.07	(0.47)	-134705.93	(-2.03)
Product contributions: multilateral approach						
COICOP 6 subgroup	Sales value difference		Price contributions		Quantity contributions	
Tissues	2423488.86	(36.58)	1412934.66	(21.33)	1010554.20	(15.25)
Toilet paper	-149921.79	(-2.26)	71219.97	(1.08)	-221141.76	(-3.34)
Baby diapers	3200407.98	(48.31)	4963225.03	(74.92)	-1762817.05	(-26.61)
Sanitary pads	1142085.37	(17.24)	220984.24	(3.34)	921101.13	(13.90)
Sanitary napkins	112069.03	(1.69)	50149.98	(0.76)	61919.05	(0.93)
Tampons	-103572.86	(-1.56)	25159.68	(0.38)	-128732.54	(-1.94)

Source: own calculations in the *PriceIndices* R package.

When calculating the Montgomery indicators for data specified with GTIN accuracy, it can be seen that the absolute values of the bilateral indicators are greater than the absolute values of the multilateral ones when the data are not filtered (Table 1). For filtered data, the price and quantity effects are accentuated when switching from the bilateral to multilateral approaches. A slightly different conclusion applied to the data aggregated at COICOP 6 level. In this case, in both the unfiltered and filtered data variants, the transition to the multilateral indicators had flattened assessments of the price and quantity effect, i.e. the absolute values of price and quantity multilateral indicators were smaller. Since multilateral indicators take into account a much wider range of data, this may mean that bilateral indicators overestimate the price and quantity effect, especially for unfiltered data.

At the same time, the normalised shares of individual COICOP 6 product subgroups in creating the price and quantity effect appeared to be similar if one compared the bilateral and multilateral approaches (see Table 2). Importantly, for each of the six product subgroups, the sign for the assessment of the price and quantity effect did not change when switching from the bilateral approach to the multilateral one, although the nominal values (non-normalised) were radically different.

5. Conclusions

The adaptation of the Montgomery transitive indicators to multilateral versions, operating on a fixed time window, appears to be a valuable addition to the analysis of scanner data due to the product churn that occurs here as a rule. A valuable result obtained from the study was the conclusion that the bilateral and multilateral Montgomery indicators differ not only in terms of tests (axioms) they satisfy, but also due to the fact that these indicators generate different values regardless of the level of data aggregation. As shown in Appendix 1, the use of the Montgomery indicators is also possible when considering all products available in the periods being compared by using a simple procedure of replacing zero prices and quantities with sufficiently small numbers. To the best of the authors' knowledge, this article constitutes the first application of the Bennet multilateral indicators (in "matched" and "available" versions) in the analysis of scanner data.

The main practical conclusion is that the relationship between the bilateral price and the quantity Montgomery indicators depends on the level of data aggregation, the choice between matched products and all available products which are to be considered, and the potential use of data filters.

The main theoretical conclusion is that the multilateral Montgomery indicators, while gaining *transitivity* and satisfying the *multi-period identity test*, lose one of the leading axioms (*identity test*). Unfortunately, it is still an open question whether it is possible to construct multilateral (and transitive at the same time) price and quantity indicators which simultaneously meet the *identity test*. According to the authors' opinion, however, even if this is not possible, multilateral indices still seem attractive from the point of view of the axiomatic approach since they meet most of the requirements. In addition, since they use the entire time window, and thus a wider range of data than their bilateral counterparts, their application to dynamically changing scanner products seems to be justified.

Appendix 1

Let us denote by $V_i^t - V_i^0$ the value change of product i calculated for the considered current period t and base period 0, i.e. $V_i^t - V_i^0 = p_i^t q_i^t - p_i^0 q_i^0$ and denote by ${}^{jt}IP_{0,t}^M(\varepsilon)$ and ${}^{j0}IQ_{0,t}^M(\varepsilon)$ the Montgomery price and quantity indicators concerning product j observed in only one of these considered periods, calculated in line with procedure (5). Let us assume that product jt is available only in period t and product $j0$ is available only in period 0. From the retailer's perspective, not observing product jt in period 0 and product $j0$ in period t , the value changes of products jt and $j0$ are as follows:

$$V_{jt}^t - V_{jt}^0 = p_{jt}^t q_{jt}^t \text{ and } V_{j0}^t - V_{j0}^0 = -p_{jt}^0 q_{jt}^0. \quad (\text{A1})$$

The following lemma is the base for statement (8):

Lemma

Using the above notation yields

$$V_{jt}^t - V_{jt}^0 = \lim_{\varepsilon \rightarrow 0^+} ({}^{jt}IP_{0,t}^M(\varepsilon) + {}^{jt}IQ_{0,t}^M(\varepsilon)), \quad (\text{A2})$$

and

$$V_{j0}^t - V_{j0}^0 = \lim_{\varepsilon \rightarrow 0^+} ({}^{j0}IP_{0,t}^M(\varepsilon) + {}^{j0}IQ_{0,t}^M(\varepsilon)). \quad (\text{A3})$$

Proof. According to assumption (5), one has $p_{jt}^0 = q_{jt}^0 = p_{j0}^t = q_{j0}^t = \varepsilon$, and as a consequence obtains

$$\begin{aligned} {}^{jt}IP_{0,t}^M(\varepsilon) + {}^{jt}IQ_{0,t}^M(\varepsilon) &= \frac{(p_{jt}^t q_{jt}^t - \varepsilon^2) \left[\ln\left(\frac{p_{jt}^t}{\varepsilon}\right) + \ln\left(\frac{q_{jt}^t}{\varepsilon}\right) \right]}{\ln(p_{jt}^t q_{jt}^t) - \ln(\varepsilon^2)} = \\ &= \frac{(p_{jt}^t q_{jt}^t - \varepsilon^2) [\ln(p_{jt}^t q_{jt}^t) - \ln(\varepsilon^2)]}{\ln(p_{jt}^t q_{jt}^t) - \ln(\varepsilon^2)} = p_{jt}^t q_{jt}^t - \varepsilon^2, \end{aligned} \quad (\text{A4})$$

and thus, from (A1) and (A4), there is an immediate proof of (A2) because

$$\lim_{\varepsilon \rightarrow 0^+} ({}^{jt}IP_{0,t}^M(\varepsilon) + {}^{jt}IQ_{0,t}^M(\varepsilon)) = \lim_{\varepsilon \rightarrow 0^+} (p_{jt}^t q_{jt}^t - \varepsilon^2) = p_{jt}^t q_{jt}^t = V_{jt}^t - V_{jt}^0. \quad (\text{A5})$$

Note that

$$\begin{aligned} {}^{j0}IP_{0,t}^M(\varepsilon) + {}^{j0}IQ_{0,t}^M &= \frac{(\varepsilon^2 - p_{j0}^0 q_{j0}^0) \left[\ln\left(\frac{\varepsilon^2}{p_{j0}^0}\right) + \ln\left(\frac{\varepsilon^2}{q_{j0}^0}\right) \right]}{\ln(\varepsilon^2) - \ln(p_{j0}^0 q_{j0}^0)} = \\ &= \frac{(\varepsilon^2 - p_{j0}^0 q_{j0}^0) [\ln(\varepsilon^2) - \ln(p_{j0}^0 q_{j0}^0)]}{\ln(\varepsilon^2) - \ln(p_{j0}^0 q_{j0}^0)} = \varepsilon^2 - p_{j0}^0 q_{j0}^0, \end{aligned} \quad (A6)$$

and thus, from (A1) and (A6), there is an immediate proof of (A3) since

$$\lim_{\varepsilon \rightarrow 0^+} ({}^{j0}IP_{0,t}^M(\varepsilon) + {}^{j0}IQ_{0,t}^M(\varepsilon)) = \lim_{\varepsilon \rightarrow 0^+} (\varepsilon^2 - p_{j0}^0 q_{j0}^0) = -p_{j0}^0 q_{j0}^0 = V_{j0}^t - V_{j0}^0 \quad (A7)$$

holds.

Appendix 2

Proof of the value change test (sum test):

From (16) and (17):

$$\begin{aligned} MIP_{0,t}^M + MIQ_{0,t}^M &= \frac{1}{T+1} \sum_{\tau=0}^T ({}_TIP_{\tau,t}^M - {}_TIP_{\tau,0}^M + {}_TIQ_{\tau,t}^M - {}_TIQ_{\tau,0}^M) = \\ &= \frac{1}{T+1} \sum_{\tau=0}^T \sum_{i \in G_{[0,T]}} \left\{ L(p_i^t q_i^t, p_i^\tau q_i^\tau) \ln\left(\frac{p_i^t}{p_i^\tau}\right) - L(p_i^0 q_i^0, p_i^\tau q_i^\tau) \ln\left(\frac{p_i^0}{p_i^\tau}\right) + \right. \\ &\quad \left. + L(p_i^t q_i^t, p_i^\tau q_i^\tau) \ln\left(\frac{q_i^t}{q_i^\tau}\right) - L(p_i^0 q_i^0, p_i^\tau q_i^\tau) \ln\left(\frac{q_i^0}{q_i^\tau}\right) \right\} = \\ &= \frac{1}{T+1} \sum_{\tau=0}^T \sum_{i \in G_{[0,T]}} \left\{ L(p_i^t q_i^t, p_i^\tau q_i^\tau) [\ln(p_i^t q_i^t) - \ln(p_i^\tau q_i^\tau)] - \right. \\ &\quad \left. - L(p_i^0 q_i^0, p_i^\tau q_i^\tau) [\ln(p_i^0 q_i^0) - \ln(p_i^\tau q_i^\tau)] \right\} = \\ &= \frac{1}{T+1} \sum_{\tau=0}^T \sum_{i \in G_{[0,T]}} (p_i^t q_i^t - p_i^\tau q_i^\tau + p_i^\tau q_i^\tau - p_i^0 q_i^0) = \\ &= \sum_{i \in G_{[0,T]}} (p_i^t q_i^t - p_i^0 q_i^0) = \sum_{i \in G_{[0,T]}} p_i^t q_i^t - \sum_{i \in G_{[0,T]}} p_i^0 q_i^0 = V_{G_{[0,T]}}^t - V_{G_{[0,T]}}^0 \end{aligned} \quad (A8)$$

Proof of the multi-period identity test:

Let us assume that there is the following relationship between prices and quantities of the current and base periods: $p_i^0 = p_i^t$ and $q_i^0 = q_i^t$. Since the multilateral price Montgomery indicator satisfies *transitivity*, one has:

$$\begin{aligned} MIP_{0,1}^M + MIP_{1,2}^M + \dots + MIP_{t-1,t}^M &= MIP_{0,t}^M = \\ &= \frac{1}{T+1} \sum_{\tau=0}^T \sum_{i \in G_{[0,T]}} \left[L(p_i^t q_i^t, p_i^\tau q_i^\tau) \ln\left(\frac{p_i^t}{p_i^\tau}\right) - L(p_i^0 q_i^0, p_i^\tau q_i^\tau) \ln\left(\frac{p_i^0}{p_i^\tau}\right) \right] = \\ &= \frac{1}{T+1} \sum_{\tau=0}^T \sum_{i \in G_{[0,T]}} \left[L(p_i^0 q_i^0, p_i^\tau q_i^\tau) \ln\left(\frac{p_i^0}{p_i^\tau}\right) - L(p_i^0 q_i^0, p_i^\tau q_i^\tau) \ln\left(\frac{p_i^0}{p_i^\tau}\right) \right] = 0. \end{aligned} \quad (A9)$$

Since the equality of prices and quantities from the current and base periods entails relationship (A9), one can conclude that the Montgomery price indicator satisfies the multi-period identity test (in the additive version for indicators). The analogical conclusion can be drawn for the Montgomery quantity indicator.

References

- Balk, B. M., Färe, R. & Grosskopf, S. (2004). The theory of economic price and quantity indicators. *Economic Theory*, 23, 149-164. <https://doi.org/10.1007/s00199-003-0364-5>
- Bennet, T. L. (1920). The theory of measurement of changes in cost of living. *Journal of the Royal Statistical Society*, 83(3), 455-462. <https://doi.org/10.2307/2340960>
- Białek, J. (2021). Price indices – a new R package for bilateral and multilateral price index calculations. *Statistika – Statistics and Economy Journal*, 36(2), 122-141.
- Białek, J. (2022). Scanner data processing in a newest version of the PriceIndices package. *Statistical Journal of the IAOS*, 38(4), 1369-1397. <https://doi.org/10.3233/sji-220963>
- Białek, J., & Roszko-Wójtowicz, E. (2023). Potential reasons for CPI chain drift bias while using electronic transaction data. *Technological and Economic Development of Economy*, 29(2), 564-590. <https://doi.org/10.3846/tede.2023.18467>
- Chambers, R. G. (2001). Consumers' surplus as an exact and superlative cardinal welfare indicator. *International Economic Review*, 42(1), 105-119. <https://doi.org/10.1111/1468-2354.00102>
- Chessa, A. (2015). Towards a generic price index method for scanner data in the Dutch CPI. The 14th meeting of the Ottawa Group, Tokyo, 20-22 May 2015. <https://stats.unece.org/ottawagroup/download/f412.pdf>
- Chessa, A. G. (2016). A new methodology for processing scanner data in the Dutch CPI. *Eurostat Review of National Accounts and Macroeconomic Indicators*, 1(2016), 49-69.
- Cross, R. M. & Färe, R. (2009). Value data and the Bennet price and quantity indicators. *Economics Letters*, 102(1), 19-21. <https://doi.org/10.1016/j.econlet.2008.10.003>
- De Boer, P., & Rodrigues, J. F. (2020). Decomposition analysis: When to use which method? *Economic Systems Research*, 32(1), 1-28. <https://doi.org/10.1080/09535314.2019.1652571>
- Diewert, W. E. (2005). Index number theory using differences rather than ratios. *American Journal of Economics and Sociology*, 64(1), 311-360. <https://doi.org/10.1111/j.1536-7150.2005.00365.x>
- 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., & Mizobuchi, H. (2009). Exact and superlative price and quantity indicators. *Macroeconomics Dynamics*, 13(2), 335-380. <https://doi.org/10.1017/s136510050909018x>
- Eltető, O., & Köves, P. (1964). On a problem of index number computation relating to international comparison. *Statisztikai Szemle*, 42(10), 507-518.
- Eurostat (2022). Guide on multilateral methods in the harmonised index of consumer prices. Publications Office of the European Union.
- Fisher, I. (1922). The making of index numbers: A study of their varieties, tests, and reliability (No. 1). Houghton Mifflin Company, 1923 [c1922].
- Fox, K. J. (2006). A method for transitive and additive multilateral comparisons: A transitive Bennet indicator. *Journal of Economics*, 87, 73-87. <https://doi.org/10.1007/s00712-005-0160-8>
- Gini, C. (1931). On the circular test of index numbers. *Metron*, 9(9), 3-24.
- International Labour Office (2004). *Consumer price index manual: Theory and practice*. Geneva.
- International Monetary Fund (2020). *Consumer price index manual: Concepts and methods*. Washington, D.C.
- 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>
- Konüs, A. A. (1939). The problem of the true index of the cost of living. *Econometrica: Journal of the Econometric Society*, 10-29.
- Montgomery, J. K. (1929). *Is there a theoretically correct price index of a group of commodities?* L'Universale Tipogr. Poliglotta
- Montgomery, J. K. (1937). *The mathematical problem of the price index*, Orchard House. P.S. King & Son.
- Szulc, B. (1964). Indices for multiregional comparisons. *Przegląd Statystyczny*, 3, 239-254.
- Van Loon, K. & Roels, D. (2018, May). Integrating big data in the Belgian CPI. *Meeting of the group of experts on consumer price indices* (pp. 7-9).
- Vartia, Y. O. (1976). Ideal log-change index numbers. *Scandinavian Journal of Statistics*, 3, 121-126.
- Von der Lippe, P. (2007). *Index theory and price statistics*. Peter Lang Verlag. <https://doi.org/10.3726/978-3-653-01120-3>

Artykuł 5

Artykuł 5 Zastosowanie indykatorów Benneta i Montgomery’ego w analizie danych skanowanych – porównanie własności i oszacowań efektów

opublikowany w

Wiadomości Statystyczne. The Polish Statistician, 70(3), 2025

Zastosowanie indyktorów Benneta i Montgomery'ego w analizie danych skanowanych – porównanie własności i oszacowań efektów¹

Natalia Pawelec^a

Streszczenie. Na początku XXI w. nastąpił renesans zainteresowania indykatorami cen i ilości zaproponowanymi w latach 20. XX w., co wynika z tego, że indykatory są mniej niż indeksy wrażliwe na występowanie zerowych cen i ilości. Własność ta ma szczególne znaczenie w badaniach segmentów rynku charakteryzujących się dużą rotacją produktów oferowanych w sprzedaży. Indykatory cen i ilości mogą znaleźć zastosowanie w analizie danych skanowanych, czyli pochodzących bezpośrednio z sieci handlowych. Analityk staje wówczas przed wyborem odpowiedniej formuły indykatora. Dwie najpopularniejsze to formuła Benneta i formuła Montgomery'ego. Celem badania omawianego w artykule jest wskazanie potencjalnych różnic między wartościami bilateralnych indyktorów (cenowych i ilościowych) Benneta i Montgomery'ego, a także sprawdzenie, w jaki sposób filtrowanie i agregacja danych skanowanych wpływają na skalę tych różnic. Porównano zarówno własności aksjomatyczne indyktorów, jak i potencjalne różnice wartości uzyskane dla wybranych grup produktów. Rezultaty badania wskazują, że różnice wartości indyktorów Benneta i Montgomery'ego sięgają kilku punktów procentowych. Wyniki oszacowań efektów – przede wszystkim cenowych, ale także ilościowych – zależą od poziomu agregacji danych, a zastosowanie filtrów danych uwydatnia zarówno efekt cenowy, jak i ilościowy.

Słowa kluczowe: dane skanowane, indyktor Benneta, indyktor Montgomery'ego

JEL: C43

Application of Bennet and Montgomery indicators in scanner data analysis: comparison of properties and effect estimates

Abstract. Price and quantity indicators, proposed in the 1920s, have since the early 2000s experienced a kind of revival, which results from the fact that, contrary to indices, these indicators are less sensitive to the occurrence of zero prices and quantities. This property is particularly important in the analysis of market segments that have a high turnover of products offered for sale. Price and quantity indicators can be applied in the analysis of scanned data, i.e. data coming directly from retail chains. The analyst is then faced with the dilemma of choosing an appropriate formula of the indicator, the two most popular being the Bennet and Montgomery formulas. The

¹ Artykuł został opracowany na podstawie referatu wygłoszonego na XLI Międzynarodowej Konferencji Naukowej „Wielowymiarowa analiza statystyczna”, która odbyła się w dniach 6–8 listopada 2023 r. w Łodzi. / The article is based on a paper delivered at 41st International Conference 'Multivariate Statistical Analysis' held on 6–8 November 2023 in Lodz, Poland.

^a Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, Instytut Ekonomii, Polska / University of Lodz, Faculty of Economics and Sociology, Institute of Economics, Poland.
ORCID: <https://orcid.org/0009-0008-5289-6204>. E-mail: natalia.pawelec@uni.lodz.pl.

aim of the research described in the paper is to identify the potential differences between the values of the Bennet and Montgomery bilateral indicators (price and quantity) and also to see how the filtering of scanner data and their aggregation affect the scale of these differences. The paper compares both the axiomatic properties of these indicators and the potential differences in their values obtained for selected groups of scanner products. The results of the study indicate that the differences in the values between the Bennet and Montgomery indicators reach several percentage points. Particularly the estimates of the price effects, but also the quantity effects depend on the level of data aggregation and the use of data filters highlights both the price and quantity effects.

Keywords: scanner data, Bennet indicator, Montgomery indicator

1. Wprowadzenie

Teoria indeksów (wskaźników) cen polega na porównywaniu cen, ilości i wydatków na towary i usługi między dwoma okresami: bazowym i bieżącym. Indeksy cen są wykorzystywane do konstruowania różnych miar aktywności gospodarczej, takich jak wskaźnik cen producentów (Producer Price Index – PPI) i wskaźnik cen konsumpcyjnych (Consumer Price Index – CPI, znany też jako wskaźnik cen towarów i usług konsumpcyjnych), oraz do urealniania produktu krajowego brutto (Fox, 2006). Nieuwzględnianie w pełni zmian jakościowych oraz towarów nowych i znikających to jedno z ważniejszych źródeł potencjalnego obciążenia pomiaru CPI. Podczas przeprowadzania analiz można przyjąć podejście wskaźnikowe (preferowane przez ekonomistów) – wykorzystywać indeksy, lub podejście biznesowe – określać różnice pomiędzy cenami i/ilością towarów i usług, czyli posługiwać się indykatorami². Analiza oparta na różnicach sprawia, że w sytuacji, gdy w niektórych okresach pojawią się zerowe ilości i/albo ceny, nie wystąpi problem dzielenia przez zero, który daje o sobie znać podczas pracy z indeksami opartymi na ilorazach (Fox, 2006). Nabiera to szczególnego znaczenia w przypadku danych skanowanych, ponieważ występowanie zerowych cen jest tu powszechnym zjawiskiem, które przysparza trudności analitycznych zarówno statystykom, jak i np. właścicielom supermarketów chcącym porównywać wyniki sprzedaży różnych grup (segmentów) produktów w dwóch okresach. Indykatory² oparte na różnicach wartości cen, ilości i wydatków okazują się zatem przydatne przede wszystkim w analizie danych skanowanych. Należy jednak podkreślić, że w badaniach zmian cen produktów indykatory nie zastępują indeksów cen, lecz stanowią ich cenne uzupełnienie.

Jeden z postulatów teorii indeksów – odwróconego czasu (ang. *factor reversal test*) – wymaga, aby indeks wartości dekomponował się na iloczyn indeksu cen i indeksu

² Indykator (ang. *indicator*) nie jest synonimem wskaźnika/indeksu (ang. *index*). W literaturze wyraźne rozróżnienie tych pojęć – w teorii pomiaru inflacji – wprowadził Diewert (2005). Jako zasadniczą różnicę pomiędzy nimi wskazał to, że indeks cen (ang. *price index*) uwzględnia ilorazy cen, a indykator cen (ang. *price indicator*) – różnice cen, co ma wpływ na ich własności i zastosowanie.

ilości. To restrykcyjny wymóg, którego nie spełnia np. indeks Laspeyresa (von der Lippe, 2007). W przypadku indykatorów, np. Benneta (1920) czy Montgomery'ego (1929), zamysł jest taki, aby różnicę wartości sprzedaży w porównywanych okresach zdekomponować na sumę dwóch efektów: cenowego (bazującego na różnicy cen produktów) i ilościowego (opartego na różnicy pomiędzy ilością sprzedanych produktów w porównywanych okresach). O ile w teorii indeksów jest to wymóg nieosiągalny dla większości formuł indeksowych, o tyle w teorii indykatorów stanowi on punkt wyjścia ich konstrukcji (Balk i in., 2004; Cross i Färe, 2009). Warto podkreślić, że w ostatnich 20 latach tematyka dotycząca indykatorów ponownie zagościła w literaturze ekonomicznej i wciąż zyskuje na popularności. Świadczą o tym prace Chambersa (2001), Balka i in. (2004) oraz Ivancica i in. (2011), w których zaproponowano nowe ujęcie indykatorów cen i ilości.

Celem badania omawianego w artykule jest wskazanie potencjalnych różnic między wartościami bilateralnych indykatorów (cenowych i ilościowych) Benneta i Montgomery'ego, a także sprawdzenie, w jaki sposób filtrowanie i agregacja danych skanowanych wpływają na skalę tych różnic.

2. Podejście aksjomatyczne w ocenie indykatorów

W teorii indeksów wyróżniono dwa główne podejścia: aksjomatyczne i ekonomiczne (Balk i in., 2004). W pierwszym z nich określa się pożądane własności matematyczne indeksów (postulaty, aksjomaty, nazywane powszechnie *testami*) i proponuje indeksy, które spełniają możliwie najwięcej z tych aksjomatów. Drugie podejście odwołuje się do optymalizacji, polegającej na zminimalizowaniu kosztu zakupu koszyka dóbr w obrębie założonej użyteczności tego koszyka. Źródło podejścia ekonomicznego wiąże się z problemem właściwego zdefiniowania, a następnie pomiaru indeksu kosztów utrzymania (Cost of Living Index – COLI). W tym podejściu przyjmuje się, że pomiędzy cenami i ilościami zachodzą relacje wyjaśnione przez modele mikroekonomiczne. Warto wspomnieć również o ujęciu stochastycznym, w którym ostateczny wybór indeksu cenowego opiera się na własnościach statystycznych. W tym podejściu przyjmuje się, że indeks cenowy stanowi nieznaną parametr w ekonometrycznym modelu wyjaśniającym zachowanie cen, szacowany za pomocą odpowiedniego estymatora. Pozwala to na wyznaczenie różnych charakterystyk estymatora (i tym samym – indeksu), np. poziomu jego obciążenia i poziomu zmienności mierzonej wariancją, a nawet skonstruowanie odpowiadającego mu przedziału ufności. W podejściach stochastycznym i aksjomatycznym uwzględniane są jedynie zmienne obserwowalne (ceny i ilości komponentów); zakłada się również, że ceny i ilości komponentów są od siebie niezależne (Białek, 2021).

Teoretyczne podstawy budowania indykatorów opracowali Bennet (1920) i Montgomery (1929). Do końca XX w. nie mieli oni kontynuatorów, ale na początku XXI w. wzrosło zainteresowanie teorią indykatorów i aspektami praktycznymi (Fox, 2006). W tabl. 1 przedstawiono wybrane testy (postulaty), które – w odpowiednio zmienionej formie – dotyczą zarówno indykatorów, jak i indeksów statystycznych³.

Tabl. 1. Zestawienie najważniejszych testów wymaganych w teorii indeksów i indykatorów

Testy	Warianty testów w przypadku	
	indykatorów	indeksów
Identyczności (ang. <i>identity test</i>)	$IP_{0,t}(p_i^0, p_i^0, q_i^0, q_i^t) = 0$	$P_{0,t}(p_i^0, p_i^0, q_i^0, q_i^t) = 1$
Granic (ang. <i>bounding test</i>)	$\min\{q_i^0(p_i^t - p_i^0), q_i^t(p_i^t - p_i^0)\} \leq IP_{0,t}(p_i^0, p_i^t, q_i^0, q_i^t) \leq \max\{q_i^0(p_i^t - p_i^0), q_i^t(p_i^t - p_i^0)\}$	$\min\{(p_i^t/p_i^0), i=1, 2, \dots, n\} \leq P_{0,t}(p_i^0, p_i^t, q_i^0, q_i^t) \leq \max\{(p_i^t/p_i^0), i=1, 2, \dots, n\}$
Monotoniczności (ang. <i>monotonicity</i>)	$IP_{0,t}(p_i^0, p_i^s, q_i^0, q_i^t) < IP_{0,t}(p_i^0, p_i^t, q_i^0, q_i^t)$ jeśli $p_i^s < p_i^t$	$P_{0,t}(p_i^0, p_i^s, q_i^0, q_i^t) < P_{0,t}(p_i^0, p_i^t, q_i^0, q_i^t)$ jeśli $p_i^s < p_i^t$
Rosnących cen (ang. <i>positivity of price change</i>)	$IP_{0,t}(p_i^0, p_i^t, q_i^0, q_i^t) > 0$ jeśli $p_i^t > p_i^0$	$P_{0,t}(p_i^0, p_i^t, q_i^0, q_i^t) > 1$ jeśli $p_i^t > p_i^0$
Linijowej homogeniczności (ang. <i>linear homogeneity</i>)	$IP_{0,t}(\lambda p_i^0, \lambda p_i^t, q_i^0, q_i^t) = \lambda IP_{0,t}(p_i^0, p_i^t, q_i^0, q_i^t)$ dla dowolnej wartości $\lambda > 0$	$P_{0,t}(p_i^0, \lambda p_i^t, q_i^0, q_i^t) = \lambda P_{0,t}(p_i^0, p_i^t, q_i^0, q_i^t)$ dla dowolnej wartości $\lambda > 0$
Odwróconego czasu (ang. <i>time reversal test</i>)	$IP_{0,t}(p_i^0, p_i^t, q_i^0, q_i^t) = -IP_{t,0}(p_i^t, p_i^0, q_i^t, q_i^0)$	$P_{0,t}(p_i^0, p_i^t, q_i^0, q_i^t) = 1/P_{t,0}(p_i^t, p_i^0, q_i^t, q_i^0)$

Źródło: opracowanie własne na podstawie: Diewert (2005).

3. Zastosowanie indykatorów cen i ilości w analizie danych skanowanych

3.1. Charakterystyka danych skanowanych

Dane skanowane pochodzą z elektronicznych terminali wykorzystywanych w punktach sprzedaży sieci handlowych i dostarczają informacji o sprzedawanych produktach zawartych w kodzie kreskowym, np. w systemie GTIN (Global Trade Item Number) lub EAN (European Article Number). Pozwalają na zdobycie wiedzy o cenach i poziomie konsumpcji produktów już na najniższym poziomie agregacji danych (Eurostat, 2017). Metody wykorzystywania danych skanowanych – opracowane w latach 70. XX w. – zmieniały się na przestrzeni lat. Zwiększały się także możliwości zdobywania tego typu informacji, co spowodowało, że stosowanie danych skanowanych do obliczania CPI stało się w ostatnim dwudziestoleciu powszechne (Białek i Roszkowski-Wójtowicz, 2023). Dane skanowane zawierają pewne metadane, zapisane w kodzie

³ Objasnienie symboli znajdujących się w tabl. 1 zamieszczono w części 3.

kreskowym, które pozwalają na rozpoznawanie produktów i w rezultacie przypisanie ich do odpowiednich grup COICOP (Classification of Individual Consumption by Purpose – Klasyfikacja Spożycia Indywidualnego według Celu).

Dane skanowane mają więcej zalet niż dane uzyskiwane tradycyjnie. Są dostarczane przez sieci handlowe (a zatem relatywnie tanie), a ich przetwarzanie odbywa się z reguły w sposób niemal w pełni zautomatyzowany. Poza tym najczęściej mają bardzo duży wolumen i niosą informacje nie tylko o cenie produktu, lecz także o kodzie sprzedawcy, etykiecie produktu, jednostce (np. szt., kg) i wartości sprzedaży oraz liczbie sprzedanych jednostek. Taki bogaty zasób wiedzy jest przydatny w procesie klasyfikacji produktów do jednorodnych grup (segmentów), a także podczas dopasowywania produktów w czasie (ang. *product matching*). Co ważne, dane skanowane mają charakter transakcyjny – informują wyłącznie o rzeczywistych wydatkach poniesionych przez konsumentów (nie uwzględniają niezakupionych produktów).

3.2. Indykator Benneta

Zbiory jednorodnych produktów należących do tej samej grupy produktów w miesiącach 0 i t oznaczono odpowiednio jako G_0 i G_t . Niech $G_{0,t}$ oznacza zbiór dopasowanych produktów w obu momentach: 0 i t , a G_0^t – zbiór dostępnych produktów w miesiącach 0 i t , tj. $G_0^t = G_0 \cup G_t$. Niech p_i^τ i q_i^τ oznaczają odpowiednio cenę (dokładnie: wartość jednostkową) i ilość i -tego produktu w czasie $\tau \in \{0, t\}$, przy czym zakładamy, że $p_i^\tau = q_i^\tau = 0$, jeśli i -ty produkt nie jest dostępny w czasie τ . Zgodnie z powyższymi oznaczeniami cenowe i ilościowe indykatory Laspeyresa i Paaschego, które są addytywnymi odpowiednikami indeksów cen i ilości Laspeyresa i Paaschego (International Labour Office i in., 2004), można przedstawić następująco:

$$IP_{0,t}^L = \sum_{i \in G_0^t} q_i^0 (p_i^t - p_i^0), \quad IQ_{0,t}^L = \sum_{i \in G_0^t} p_i^0 (q_i^t - q_i^0), \quad (1)$$

$$IP_{0,t}^P = \sum_{i \in G_0^t} q_i^t (p_i^t - p_i^0), \quad IQ_{0,t}^P = \sum_{i \in G_0^t} p_i^t (q_i^t - q_i^0). \quad (2)$$

Addytywnym odpowiednikiem cenowego i ilościowego indeksu Fishera (1922) są cenowe i ilościowe indykatory Benneta (1920), zdefiniowane jako średnia arytmetyczna indykatorów Laspeyresa i Paaschego:

$$IP_{0,t}^B = \frac{1}{2} (IP_{0,t}^L + IP_{0,t}^P) = \sum_{i \in G_0^t} \frac{q_i^0 + q_i^t}{2} (p_i^t - p_i^0), \quad (3)$$

$$IQ_{0,t}^B = \frac{1}{2}(IQ_{0,t}^L + IQ_{0,t}^P) = \sum_{i \in G_0^t} \frac{p_i^0 + p_i^t}{2} (q_i^t - q_i^0). \quad (4)$$

Niech $V_{G_0}^0$ i $V_{G_t}^{0t}$ oznaczają całkowite wydatki na wszystkie dostępne produkty odpowiednio w okresach 0 i t , tj. $V_{G_0}^0 = \sum_{i \in G_0} p_i^0 q_i^0$ oraz $V_{G_t}^{0t} = \sum_{i \in G_0} p_i^t q_i^t$. Indykatory Benneta pozwalają rozłożyć bezwzględną zmianę całkowitej wartości sprzedaży addytywnie na efekt cenowy i ilościowy:

$$V_{G_t}^{0t} - V_{G_0}^0 = IP_{0,t}^B + IQ_{0,t}^B. \quad (5)$$

Równanie (5) oznacza, że indykatory Benneta spełniają aksjomat sumy, znany z podejścia aksjomatycznego (Diewert, 2005). Łatwo wykazać, że indykatory Benneta spełniają także aksjomaty: identyczności, monotoniczności w cenach/iłościach, jednorodności w cenach/iłościach i odwracalności w czasie (Diewert, 2005).

W przeciwieństwie do przedsiębiorstw, w przypadku których analizowany zbiór pozostaje względnie stabilny w czasie (oznacza to, że obserwacje dotyczą zazwyczaj tej samej grupy firm, a zmiany w jej składzie są rzadkie, podobnie jak zerowe ceny lub ilości), w przypadku danych skanowanych zmienność asortymentu oraz występowanie zerowych wartości są częstym zjawiskiem – co wynika ze znacznej rotacji produktów spowodowanej sezonowością artykułów, trendami w sprzedaży i polityką dostawców – i stanowią duże wyzwanie analityczne.

Poniżej przedstawiono wzory na indykatory cen i ilości Benneta zdefiniowane tylko dla dopasowanych produktów. Chodzi o porównywanie produktu homogenicznego⁴, obserwowanego jednocześnie w okresie bieżącym i bazowym:

$$mIP_{0,t}^B = \sum_{i \in G_{0,t}} \frac{q_i^0 + q_i^t}{2} (p_i^t - p_i^0), \quad (6)$$

$$mIQ_{0,t}^B = \sum_{i \in G_{0,t}} \frac{p_i^0 + p_i^t}{2} (q_i^t - q_i^0). \quad (7)$$

Niech $V_{G_{0,t}}^0$ i $V_{G_{0,t}}^{0t}$ oznaczają całkowite wydatki na wszystkie dopasowane produkty, odpowiednio w okresach 0 i t , tj. $V_{G_{0,t}}^0 = \sum_{i \in G_{0,t}} p_i^0 q_i^0$ oraz $V_{G_{0,t}}^{0t} = \sum_{i \in G_{0,t}} p_i^t q_i^t$. Indykatory Benneta oparte na dopasowanych produktach umożliwiają dekompozycję

$$V_{G_{0,t}}^{0t} - V_{G_{0,t}}^0 = mIP_{0,t}^B + mIQ_{0,t}^B. \quad (8)$$

⁴ Produkt homogeniczny to taki, który nie zmienił się między porównywanymi okresami i który zachowuje identyczne cechy, takie jak np. marka, rozmiar, skład i kod EAN.

3.3. Indykator Montgomery'ego

Bilateralny indyktor cenowy Montgomery'ego, który porównuje bieżący okres t z okresem bazowym 0 i uwzględnia tylko dopasowane produkty, można – przy użyciu oznaczeń wprowadzonych w sekcji 3.2 – zdefiniować następująco:

$$IP_{0,t}^M = \sum_{i \in G_{0,t}} L(p_i^t q_i^t, p_i^0 q_i^0) \ln \left(\frac{p_i^t}{p_i^0} \right), \quad (9)$$

a bilateralny indyktor ilościowy Montgomery'ego jako:

$$IQ_{0,t}^M = \sum_{i \in G_{0,t}} L(p_i^t q_i^t, p_i^0 q_i^0) \ln \left(\frac{q_i^t}{q_i^0} \right). \quad (10)$$

Można wykazać (Montgomery, 1929), że zmiana wartości sprzedaży w dwóch rozważanych okresach (zaobserwowana dla dopasowanych produktów) może być zdekomponowana również przy użyciu indyktorów cenowych i ilościowych Montgomery'ego. Innymi słowy zachodzi

$$V_{G_{0,t}}^{0t} - V_{G_{0,t}}^0 = mIP_{0,t}^M + mIQ_{0,t}^M. \quad (11)$$

Warto podkreślić, że w przypadku danych skanowanych zbiór produktów dostępnych w okresach bazowym i bieżącym (tj. $G_0^t = G_0 \cup G_t$) jest zazwyczaj nadzbiorem zbioru dopasowanych produktów. W przypadku indykatora Benneta nie stwarza to problemów obliczeniowych, ponieważ zakłada się, że jeśli produkt jest dostępny w okresie t_1 , ale nie jest dostępny w okresie t_2 , to ma zerową ilość i cenę w okresie t_2 . W przypadku indykatora Montgomery'ego zastosowanie takiej procedury nie jest możliwe, ponieważ wykorzystuje się tu logarytmy z cen i ilości. Można jednak wykazać, że następująca procedura umożliwia skuteczne określenie cen i ilości niedopasowanych produktów w porównywanych okresach:

$$p_i^t = q_i^t = \varepsilon \text{ dla } i \in G_0 \setminus G_t \text{ oraz } p_i^0 = q_i^0 = \varepsilon \text{ dla } i \in G_0 \setminus G_t, \quad (12)$$

przy czym ε jest z założenia niewielką liczbą rzeczywistą dodatnią (w badaniu przyjęto $\varepsilon = 0,000001$).

Przyjęcie dostatecznie małej wartości ε nie tylko nie powoduje żadnych obciążeń oceny dynamiki zmian cen i ilości, ale jest wymogiem koniecznym w przypadku stosowania indyktorów Montgomery'ego, gdy występują zerowe ceny lub ilości. Białek i Pawelec (2024) wykazali, że dekompozycja zmiany całkowitej wartości sprzedaży na efekt cenowy i ilościowy z wykorzystaniem indyktorów Montgomery'ego jest osiągnięta dopiero asymptotycznie, czyli przy $\varepsilon \rightarrow 0^+$. Przeprowadzone przez nich liczne

eksperymenty dowodzą, że wartości $\varepsilon = 0,000001$ lub niższe w zasadzie gwarantują poprawną ocenę efektu cenowego i ilościowego.

Dzięki powyższej procedurze indyktor Montgomery'ego dla wszystkich dostępnych (w tym niedopasowanych) produktów można zdefiniować następująco:

$$IP_{0,t}^M = \sum_{i \in G_{0,t}} L(p_i^t q_i^t, p_i^0 q_i^0) \ln \left(\frac{p_i^t}{p_i^0} \right), \quad (13)$$

$$IQ_{0,t}^M = \sum_{i \in G_{0,t}} L(p_i^t q_i^t, p_i^0 q_i^0) \ln \left(\frac{q_i^t}{q_i^0} \right). \quad (14)$$

4. Badanie empiryczne

4.1. Charakterystyka wykorzystanych zbiorów danych skanowanych

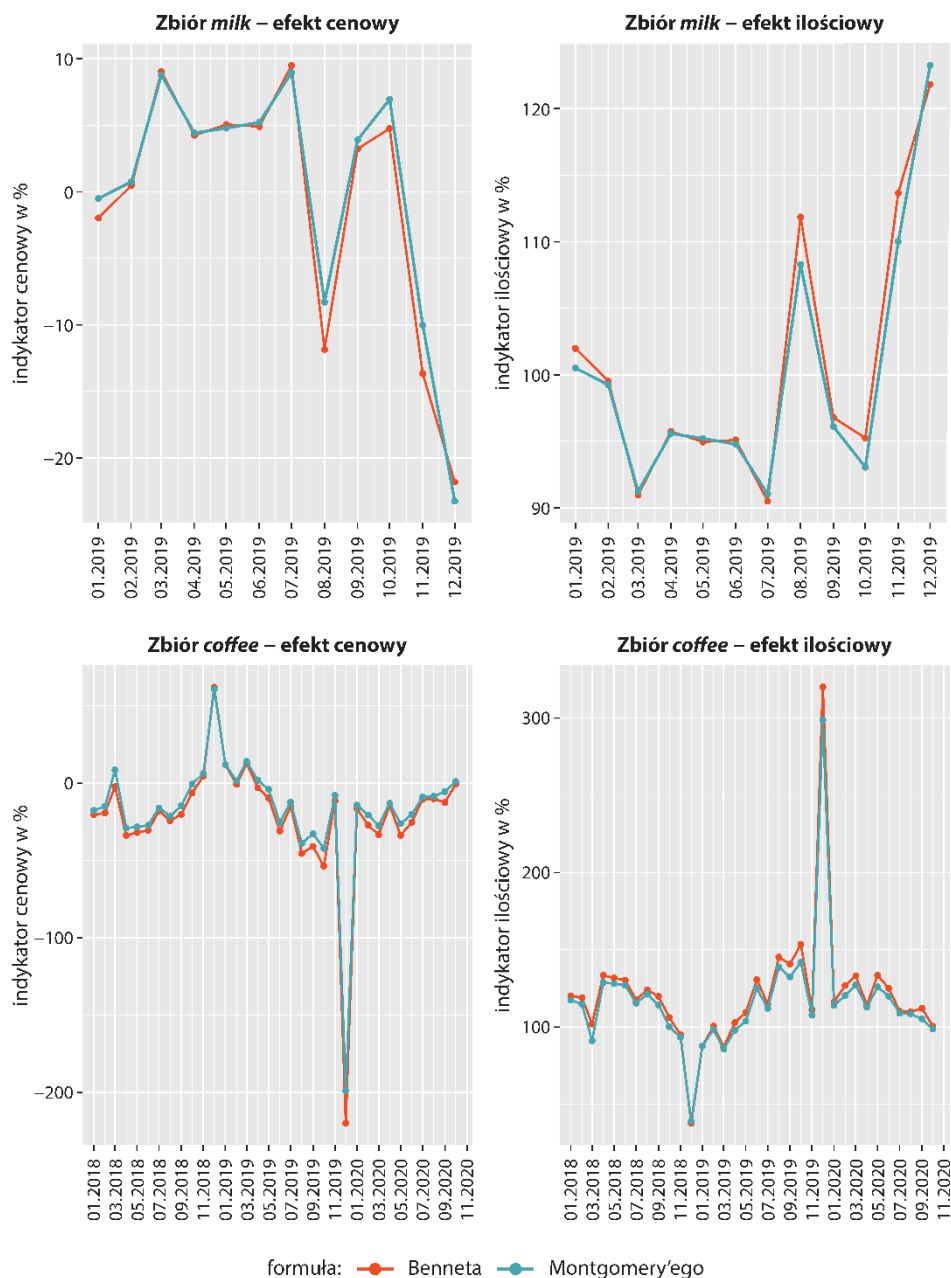
W badaniu empirycznym wykorzystano dane skanowane będące składową pakietu *PriceIndices* (Białek, 2021) napisanego w środowisku R. Dane pochodzą z sieci handlowej prowadzącej działalność w Polsce⁵. Zbiór *milk* składa się z 4386 rekordów dotyczących sprzedaży mleka i obejmuje okres od grudnia 2018 r. do sierpnia 2020 r. Zbiór *coffee* zawiera 42 561 rekordów dotyczących sprzedaży kawy i obejmuje okres od grudnia 2017 r. do października 2020 r. Zmienne w układzie kolumnowym, które znajdują się w zbiorach *milk* i *coffee*, są następujące:

- *time* – data wykonania transakcji;
- *prices* – cena sprzedanego produktu w zł;
- *quantities* – ilość sprzedanego produktu w mg/ml;
- *prodID* – unikalny kod produktu nadany po przeprowadzeniu dopasowania produktów w czasie;
- *retID* – unikalny kod poszczególnych punktów sprzedaży detalicznej (outletów sieci handlowej);
- *description* – opis sprzedanych produktów (tzw. etykieta produktu).

4.2. Porównanie indyktorów cen i ilości

Na wykresie zestawiono efekty cenowy i ilościowy oszacowane z wykorzystaniem indyktorów Benneta oraz Montgomery'ego w podziale na zbiory *milk* i *coffee*.

⁵ Nie uzyskano zgody na ujawnienie nazwy sieci handlowej, z której pochodzą dane.

Wykres. Porównanie efektów cenowego i ilościowego pomiędzy indykatorami

Źródło: opracowanie własne na podstawie danych pochodzących z pakietu *PricelIndices*.

Z wykresu wynika, że zarówno dla zbioru *milk*, jak i dla zbioru *coffee* cenowy indyktor Benneta daje nieco mniejsze wartości niż cenowy indyktor Montgomery'ego – stanowi to regułę w całym oknie czasowym. Z kolei efekt ilościowy w przypadku indykatora Benneta jest większy, niż wskazuje indyktor Montgomery'ego – to również reguła obserwowana w całym analizowanym oknie czasowym.

4.3. Wpływ filtrowania danych i agregacji na różnice pomiędzy indykatorami Benneta i Montgomery'ego

W tabl. 2 i 3 zamieszczono wartości indykatorów Benneta i Montgomery'ego w dwóch przypadkach: bez użycia filtrów danych i po zastosowaniu filtrów danych (ściśle: filtra niskiej sprzedaży i filtra ekstremalnych zmian cen). Uwzględniono dwa warianty definicji produktu homogenicznego:

- określanie produktu na najniższym poziomie agregacji danych, tj. z zastosowaniem kodu GTIN;
- ujęcie szersze, czyli z wykorzystaniem kategorii COICOP 6.

Uzyskane wyniki przedstawiono w tabl. 2.

Tabl. 2. Porównanie indykatorów Benneta oraz Montgomery'ego wyznaczonych dla różnych poziomów agregacji danych i w zależności od opcjonalnego filtrowania danych dla zbioru *milk*

Wyszczególnienie	Indyktor Benneta		Indyktor Montgomery'ego	
	bez filtrów	z filtrami	bez filtrów	z filtrami
Różnica wartości sprzedaży: w zł	9 859,35	3 290,01	859,35	3 290,01
w %	100,00	100,00	100,00	100,00
Poziom GTIN				
Efekt cenowy: w zł	-2 151,48	-13 017,97	-2 293,52	-11 092,52
w %	-21,82	-395,68	-23,26	-337,16
Efekt ilościowy: w zł	1 010,83	16 307,98	12 152,87	14 382,53
w %	121,82	495,68	123,26	437,16
Poziom COICOP 6				
Efekt cenowy: w zł	-103,43	-1 518,77	-97,50	-1 515,30
w %	-1,05	-46,16	-0,99	-46,06
Efekt ilościowy: w zł	9 962,78	4 808,78	9 956,84	4 805,31
w %	101,05	146,16	100,99	146,06

Źródło: opracowanie własne na podstawie danych pochodzących z pakietu *PriceIndices*.

Na poziomie GTIN – bez względu na rodzaj zastosowanego indykatora – użycie filtrów danych powoduje, że efekt cenowy staje się bardziej odseparowany od efektu ilościowego. Podobną zależność można zaobserwować na poziomie agregacji danych COICOP 6, ale nie jest ona już tak wyraźna.

Zastosowanie filtrów danych nie zmienia znaków efektów cenowego i ilościowego. Oznacza to, że jeśli przed zastosowaniem filtrów wartość indykatora

cenowego/iłościowego była ujemna, to po ich zastosowaniu pozostawała ujemna, a jeśli była dodatnia – to pozostawała dodatnia. Oczywiście różnica wartości sprzedaży produktów jest jednakowa bez względu na poziom agregacji danych. Obserwując wartości indykatorów wyrażone liczbowo (nie procentowo), można zauważyć, że tylko w przypadku poziomu COICOP 6 zastosowanie filtrów prowadzi do zmniejszenia efektu ilościowego; dotyczy to obu indykatorów. Na poziomie GTIN zarówno efekt cenowy, jak i ilościowy uwydatniał się po zastosowaniu filtrów danych, ale zachowywał kierunek zmian.

Tabl. 3. Porównanie indykatorów Benneta oraz Montgomery'ego wyznaczonych dla różnych poziomów agregacji danych i w zależności od opcjonalnego filtrowania danych dla zbioru *coffee*

Wyszczególnienie	Indykator Benneta		Indykator Montgomery'ego	
	bez filtrów	z filtrami	bez filtrów	z filtrami
Różnica wartości sprzedaży: w zł	-535 494,52	-416 645,82	-535 494,52	-416 645,82
w %	100,00	100,00	100,00	100,00
Poziom GTIN				
Efekt cenowy: w zł	2 834,32	-38 480,34	-6084,94	-43 557,37
w %	-0,53	9,24	1,14	10,45
Efekt ilościowy: w zł	-538 328,84	-378 165,48	-529 409,58	-373 088,45
w %	100,53	90,76	98,86	89,55
Poziom COICOP 6				
Efekt cenowy: w zł	116 264,89	64 714,01	112 272,90	60 838,91
w %	-21,27	-15,53	-20,97	-14,60
Efekt ilościowy: w zł	-651 759,41	-481 359,83	-647 767,39	-477 484,73
w %	121,71	115,53	120,97	114,60

Źródło: opracowanie własne na podstawie danych pochodzących z pakietu *PricelIndices*.

W przypadku zbioru danych o sprzedaży kawy wnioski tylko częściowo pokrywają się z tymi dotyczącymi sprzedaży mleka. Prawdopodobnie, że zastosowanie filtrów danych wyraźniej odseparowuje efekt cenowy i ilościowy, dotyczy jedynie niskiego poziomu agregacji danych (GTIN). Na poziomie agregacji COICOP 6 zauważalny jest odwrotny kierunek zmian – po wprowadzeniu filtrów efekty cenowy i ilościowy zbliżają się do siebie.

5. Podsumowanie

Dane skanowane mają wiele zalet i powinny być traktowane w szczególny sposób, ponieważ znacząco różnią się od tradycyjnych danych wykorzystywanych w badaniach statystycznych. Charakteryzują się wysokim stopniem szczegółowości – zawierają informacje o konkretnych produktach (np. na poziomie kodu EAN), dokładnej liczbie sprzedanych jednostek, rzeczywistej cenie transakcyjnej oraz czasie i miejscu

sprzedaży. Dzięki temu umożliwiają precyzyjną analizę zmian cen i zachowań konsumentów. Jednak ze względu na dużą rotację asortymentu oraz ogromną objętość danych te wymagają specjalnych metod przetwarzania i analizy. Jednym z problemów związanych z szacowaniem indeksów cen na najniższym poziomie agregacji danych są zerowe ceny i ilości, charakterystyczne dla produktów rotujących (np. sezonowych). W artykule poddano więc pod dyskusję zastosowanie indyktorów, które w przeciwieństwie do indeksów nie są wrażliwe na zerowe ceny i ilości.

Głównym wynikiem badania omawianego w artykule jest wykazanie praktycznej wartości indyktorów w pracy na danych skanowanych. Możliwość prostej i szybkiej dekompozycji efektów cenowego i ilościowego może okazać się przydatna zarówno statystykom pracującym w urzędach statystycznych, jak i np. właścicielom sieci handlowych przy określaniu popytu na produkty.

Nasuwać się również wnioski szczegółowe. W przypadku indykatora Benneta efekt cenowy jest mniejszy (oszacowanie na niższym poziomie), a efekt ilościowy – większy (oszacowanie na wyższym poziomie) niż w przypadku indykatora Montgomery’ego. Różnice wartości tych indyktorów nie przekraczają kilku punktów procentowych.

Wyniki oszacowań efektów cenowych i ilościowych zależą od poziomu agregacji danych, ale nie jest możliwe dokonanie jednoznacznej oceny kierunku tego wpływu (jest on różny w przypadku różnych zbiorów danych). Zastosowanie filtrów danych najczęściej uwypatnia (ale nie stanowi to reguły) zarówno efekt cenowy, jak i ilościowy w przypadku obu porównywanych indyktorów i obu poziomów agregacji danych.

Bibliografia

- Balk, B. M., Färe, R., Grosskopf, S. (2004). The theory of economic price and quantity indicators. *Economic Theory*, 23(1), 149–164. <https://doi.org/10.1007/s00199-003-0364-5>.
- Bennet, T. L. (1920). The Theory of Measurement of Changes in Cost of Living. *Journal of the Royal Statistical Society*, 83(3), 455–462. <https://doi.org/10.2307/2340960>.
- Białek, J. (2012). Propozycja indeksu cen. *Wiadomości Statystyczne*, 57(7), 14–24. <https://doi.org/10.59139/ws.2012.07.1>.
- Białek, J. (2021). PriceIndices – a New R Package for Bilateral and Multilateral Price Index Calculations. *Statistika. Statistics and Economy Journal*, 101(2), 122–141. https://csu.gov.cz/docs/107508/74c31da1-5b28-d24e-2763-fe355180540c/32019721q2_bialek.pdf?version=1.0.
- Białek, J., Pawelec, N. (2024). The use of transitive Montgomery indicators for scanner data analysis. *Argumenta Oeconomica*, 53(2), 1–13. <https://doi.org/10.15611/aoe.2024.2.01>.
- Białek, J., Roszko-Wójtowicz, E. (2023). Potential reasons for CPI chain drift bias while using electronic transaction data. *Technological and Economic Development of Economy*, 29(2), 564–590. <https://doi.org/10.3846/tede.2023.18467>.
- Chambers, R. G. (2001). Consumers’ surplus as an exact and superlative cardinal welfare indicator. *International Economic Review*, 42(1), 105–119. <https://doi.org/10.1111/1468-2354.00102>.

- Cross, R. M., Färe, R. (2009). Value data and the Bennet price and quantity indicators. *Economics Letters*, 102(1), 19–21. <https://doi.org/10.1016/j.econlet.2008.10.003>.
- Diewert, W. E. (2005). Index Number Theory Using Differences Rather Than Ratios. *The American Journal of Economics and Sociology*, 64(1), 311–360. <https://doi.org/10.1111/j.1536-7150.2005.00365.x>.
- Eurostat. (2017). *Harmonised Index of Consumer Prices. Practical Guide for Processing Supermarket Scanner Data*. <https://circabc.europa.eu/ui/group/7b031f10-ac19-4da3-a36f-58708a70133d/library/8e1333df-ca16-40fc-bc6a-1ce1be37247c/details?download=true>.
- Fisher, I. (1922). *The making of index numbers: A study of their varieties, tests, and reliability*. Houghton Mifflin.
- Fox, K. J. (2006). A Method for Transitive and Additive Multilateral Comparisons: A Transitive Bennet Indicator. *Journal of Economics*, 87(1), 73–87. <https://doi.org/10.1007/s00712-005-0160-8>.
- International Labour Office, International Monetary Fund, Organisation for Economic Co-operation and Development, Eurostat, United Nations, World Bank. (2004). *Consumer Price Index Manual. Theory and Practice*. https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@dgreports/@stat/documents/presentation/wcms_331153.pdf.
- 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>.
- von der Lippe, P. (2007). *Index Theory and Price Statistics*. Peter Lang. <https://doi.org/10.3726/978-3-653-01120-3>.
- Montgomery, J. K. (1929). *Is There a Theoretically Correct Price Index of a Group of Commodities?*. L'Universale Tipografia poliglotta.

Artykuł 6

Artykuł 6 Comparison of transitive Bennet and Montgomery indicators on the basis of scanner data sets

opublikowany w

European Research Studies Journal, Vol 27, 2024

Wkład współautorów:

Natalia Pawelec: 80%

Jacek Białek: 20%

Wkład własny:

Autorka przeprowadziła całość analizy empirycznej na dwóch zbiorach danych skanowanych, obliczenie tranzytywnych indykatorów Benneta i Montgomery'ego w pięciu wariantach filtrowania danych i na dwóch poziomach agregacji (GTIN i COICOP 6) z wykorzystaniem pakietu *PriceIndices*, identyfikację przypadków zmiany znaku indykatorów pod wpływem filtrowania oraz analizę zbieżności wyników obu formuł. Autorka opracowała tabele wynikowe oraz przygotowała i zredagowała manuskrypt artykułu.

Comparison of Transitive Bennet and Montgomery Indicators on the Basis of Scanner Data Sets

Submitted 04/09/24, 1st revision 14/09/24, 2nd revision 16/10/24, accepted 30/10/24

Jacek Bialek¹, Natalia Pawelec²

Abstract:

Purpose: The main purpose of the study is to assess the practical applicability of Bennet and Montgomery indicators for analysing electronic transaction data in the form of scanner data. The specific goal is to compare transitive versions of these indicators on the basis of dynamic scanner data, where there is a high turnover of products.

Design/Methodology/Approach: The model for considerations is developed based on a comprehensive review of existing literature. Authors adopt the transitive Bennet and Montgomery indicators from the field of inter-firm comparisons and use the Fox's transformation for obtaining multilateral versions of these indicators being applicable in the scanner data case.

Findings: The transitive Bennet formula leads to greater differences in values between the price and quantity indicators than the transitive Montgomery formula. This conclusion is identical for each type of data filter and for each data aggregation level. However, it can be seen that for a higher level of data aggregation (COICOP level 6), the differences between the corresponding multilateral Bennet and Montgomery indicators are much smaller than for the barcode level (GTIN level). In principle, it can be said that at a higher level of aggregation, the Bennet and Montgomery multilateral indicators lead to fairly similar values and any differences between their values are substantial at the disaggregated level.

Practical Implications: The Bennet and Montgomery transitive indicators can be useful in analysing scanner data if retail chain managers want to decompose sales value into a price and quantity factor. The framework provides a roadmap for practical using these multilateral indicators for dynamic scanner data sets. The adoption of the transitive indicators for the scanner data analysis can lead to substantial improvements in the management of product sales by the retail chain. Areas of potential application for the approach based on price and quantity differences could be also, for example: profit and cost change decompositions or the analysis of changes in consumer surplus.

Originality/Value: The added value of the work is the identification of potential differences between the multilateral Bennet and Montgomery indicators calculated on scanner data sets obtained from retail chains. In particular, it is pioneering to determine the magnitude of these differences depending on the level of data aggregation and the type of data filters used.

Keywords: Scanner data, price and quantity indicators, Bennet indicator, Montgomery indicator, transitivity.

¹University of Lodz, Department of Statistical Methods, Poland,
e-mail: jacek.bialek@uni.lodz.pl;

²University of Lodz, Department of Economic Mechanisms, Poland,
e-mail: natalia.pawelec@uni.lodz.pl;

JEL codes: C43, E31.

Paper type: Research article.

Acknowledgements: Publication co-financed from the funds of state budget within the program of the Minister of Education and Science under the title „Doskonała nauka – Wsparcie konferencji naukowych”, project no. DNK/SP/549670/2022, amount of funding PLN 147,570.90, total value of the project PLN 165,839.60.

1. Introduction

The modern price index theory is based on comparisons of ratios of prices, quantities and expenditures of goods and services (von der Lippe, 2007; International Labour Office, 2004; International Monetary Fund, 2020). These index numbers are used to build various economic measures, such as the Gross Domestic Product (GDP), the Producer Price Index (PPI) or the Consumer Price Index (CPI).

Nevertheless, in many business contexts we may be more interested in the magnitude of differences in prices, quantities and sale values. Areas of potential application for the approach based on price and quantity differences could be, for example, revenue change decompositions, profit and cost change decompositions, or the analysis of changes in consumer surplus (Diewert, 2005).

An important benefit of using such differences is that there is no problem associated with the occurrence of zero prices and quantities, a problem that arises when we work with ratios. The problem with zero prices or zero quantities may be very serious in many business applications where not all goods are produced and purchased in every period. Zero prices and quantities are common in scanner data sets obtained from retail chains due to the high product churn observed in supermarkets.

In the case of scanner data and the multilateral index (or indicator) approach in which we observe price and quantity processes over a fixed time window, the construction of a price and quantity indicators is complicated. The added value of the work is the identification of potential differences between the multilateral Bennet and Montgomery indicators calculated on scanner data sets obtained from retail chains. In particular, it is pioneering to determine the magnitude of these differences depending on the level of data aggregation and the type of data filters used.

The main conclusion obtained from empirical considerations based on real scanner data sets is that the transitive Bennet formula leads to greater differences in values between the price and quantity indicators than the transitive Montgomery formula. This conclusion is identical for each type of data filter and for each data aggregation

level. However, it can be seen that for a higher level of data aggregation (COICOP level 6), the differences between the corresponding multilateral Bennet and Montgomery indicators are much smaller than for the barcode level (GTIN level). In principle, it can be said that at a higher level of aggregation, the Bennet and Montgomery multilateral indicators lead to fairly similar values and any differences between their values are substantial at the disaggregated level.

2. Literature Review

The price and quantity indicators theory is well established in the economic literature, where it was introduced in the early 20th century (Bennet, 1920; Montgomery, 1929). Professor Diewert (2005) was the first to call index numbers expressed in terms of difference as "indicators". Recently, one can see a return of interest in this approach on the part of statisticians and economists (Balk *et al.*, 2004; Diewert, 2005; Fox, 2006; Cross and Färe, 2009; de Boer and Rodrigues, 2020).

Most often, papers on indicators revolve around issues related to company operations (e.g., Fox, 2006; Cross and Färe, 2009). Nevertheless, economic papers which apply price and quantity indices to areas of scanner data analysis can be also found (e.g., Białek, Pawelec, 2024).

3. Methodology: Bilateral Indicators

In a static world of products (i.e., when the set of products is the same in the periods being compared), we can construct bilateral price and quantity indexes in a number of ways. One approach to constructing a proper price index is the axiomatic approach, in which the index should satisfy a series of required postulates (so-called *tests*).

Let p_i^τ and q_i^τ denote the price and quantity of the i -th product at the time $\tau \in \{0, t\}$. Let p_τ and q_τ denote vectors of prices and quantities observed at the time $\tau \in \{0, t\}$. We can now signify the price and quantity indices that compare the current period t with the base period 0 as respectively: $P_{0,t} \equiv P_{0,t}(p_0, q_0, p_t, q_t)$ and $Q_{0,t} \equiv Q_{0,t}(p_0, q_0, p_t, q_t)$. Let us also denote by $IP_{0,t} \equiv IP_{0,t}(p_0, q_0, p_t, q_t)$ and $IQ_{0,t} \equiv IQ_{0,t}(p_0, q_0, p_t, q_t)$ the price and quantity indicators that operate on price and quantity differences. The value of products consumed at the time $\tau \in \{0, t\}$ can be computed as $V_\tau = \sum_i p_i^\tau q_i^\tau$. Price index theory considers a restrictive *factor reversal test* (von der Lippe, 2007) that expects an index to satisfy an equation: $V_t / V_0 = P_{0,t} Q_{0,t}$. The main idea of creating indicators is such their design that ensures that: $V_t - V_0 = IP_{0,t} + IQ_{0,t}$ (Balk *et al.*, 2004; Diewert, 2005).

Due to the high turnover of scanner products, a frequently used approach in determining price indices is the matched sample approach. Let $G_{0,t}$ denote a set of matched products in both considered periods, i.e. $G_{0,t} = G_0 \cap G_t$. Let also G_0^t denote a set of all available products in both considered periods, i.e. $G_0^t = G_0 \cup G_t$.

In the paper, we will only consider the case of all available products, which is technically more difficult than the case of matched products due to the nature of the Montgomery indicators (see Section *The Montgomery indicator*). In the case of the Bennet indicator, the situation of product unavailability in period τ is technically simple, i.e. we just assume that $p_i^\tau = q_i^\tau = 0$. We refer the reader interested in the case of matched products to the paper by Bialek and Pawelec (2024).

3.1 The Bennet Indicator

Under the above signification, the Bennet price and quantity indicators can be written respectively as follows (Bennet, 1920):

$$IP_{0,t}^B = \sum_{i \in G_0^t} \frac{q_i^0 + q_i^t}{2} (p_i^t - p_i^0) \quad (1)$$

and

$$IQ_{0,t}^B = \sum_{i \in G_0^t} \frac{p_i^0 + p_i^t}{2} (q_i^t - q_i^0) \quad (2)$$

It can be shown (Bennet, 1920) that

$$V_t - V_0 = IP_{0,t}^B + IQ_{0,t}^B, \quad (3)$$

which means that the Bennet indicators meet the *sum test* (or the *value change test*).

3.2 The Montgomery Indicator

Please note that in the case of scanner data, the set of products available in the base and current periods, i.e., $G_0^t = G_0 \cup G_t$, is generally an oversampling of the set of matched products $G_{0,t} = G_0 \cap G_t$. In the case of the Bennet indicators, this does not generate technical problems in their calculation since it is simply assumed that a product that is available in the period t_1 , but is not available in the period t_2 , has zero price and quantity in the period t_2 . Such a procedure does not apply to the Montgomery indicators since they use logarithms from prices and quantities in their

syntax. However, it can be shown (Białek and Pawelec, 2024) that the determination of prices and quantities of mismatched products in the periods being compared can effectively follow the following procedure:

$$p_i^t = q_i^t = \varepsilon, \text{ for } i \in G_0 \setminus G_t, \text{ and } p_i^0 = q_i^0 = \varepsilon, \text{ for } i \in G_t \setminus G_0, \quad (4)$$

which must hold for a very small, positive and real number ε .

With the above-presented procedure for treating mismatched products we can define the Montgomery indicators for all available products (including mismatched products) as follows:

$$IP_{0,t}^M = \sum_{i \in G_0^t} L(p_i^t q_i^t, p_i^0 q_i^0) \text{Ln}\left(\frac{p_i^t}{p_i^0}\right), \quad (5)$$

and

$$IQ_{0,t}^M = \sum_{i \in G_0^t} L(p_i^t q_i^t, p_i^0 q_i^0) \text{Ln}\left(\frac{q_i^t}{q_i^0}\right), \quad (6)$$

where $L(x, y)$ denotes the *Vartia* (1976) *mean* of two positive real numbers x and y , which is known in the mathematical literature as the *logarithmic mean* (von der Lippe, 2007) and is defined as follows:

$$L(x, y) = \begin{cases} \frac{x - y}{\text{Ln}(x) - \text{Ln}(y)} : x \neq y \\ x : x = y \end{cases} \quad (7)$$

Please note that in general, when $G_0^t \neq G_{0,t}$, the Montgomery indicators (5) and (6) are functions of the parameter ε (under the assumption (5)), and thus they can be signified by $IP_{0,t}^M(\varepsilon)$ and $IQ_{0,t}^M(\varepsilon)$. Nevertheless, the Montgomery indicators satisfy the *sum test* (otherwise known as the *value change test*) with a good approximation if the procedure (4) is taken into consideration (Białek & Pawelec, 2024). In other words it holds that

$$V_{G_0^t}^t - V_{G_0^t}^0 \equiv \sum_{i \in G_t \setminus G_0} p_i^t q_i^t - \sum_{i \in G_0 \setminus G_t} p_i^0 q_i^0 + \sum_{i \in G_{0,t}} p_i^t q_i^t - \sum_{i \in G_{0,t}} p_i^0 q_i^0 = \lim_{\varepsilon \rightarrow 0^+} (IP_{0,t}^M(\varepsilon) + IQ_{0,t}^M(\varepsilon)). \quad (8)$$

In the empirical study conducted, we will use the value $\varepsilon = 0.000001$, which provides a very accurate approximation (4).

4. Discussion: Indicators Applied for Scanner Data

Zero prices and quantities are common in scanner data sets due to the high product churn observed in supermarkets. Scanner data, which support CPI calculations in many countries, mean transaction data that specify turnover and numbers of items sold by barcodes, e.g., Global Trade Article Number (GTIN), European Article Number (EAN), or Stock Keeping Unit (SKU) codes.

Scanner data have numerous advantages compared to traditional survey data collection because such data sets are much bigger than traditional ones and they contain complete transaction information, i.e., information about prices and quantities at the barcode level and also many other product characteristics (i.e., product grammage, sales unit, product label, VAT level, etc.).

Nevertheless, at the barcode level, the occurrence of zero-scanner prices is quite common because in supermarkets we often see a variety of discounts and price reductions (even down to zero), as well as a significant turnover of products. It causes analytical problems for statistical offices as well as for supermarket owners wishing to compare the sales performance of different product segments in two time periods.

Thus, an approach based on differences in price values, quantities and expenditures can also be very useful in analysing scanner data. However, according to the best of the authors' knowledge, there is a lack of papers in the literature that apply the indicators to the analysis of scanner data, which is the main objective of this paper. One of the few papers that does this is that of Bialek and Pawelec (2024).

In the case of scanner data and the multilateral index (or indicator) approach in which we observe price and quantity processes over a fixed time window, the construction of a price and quantity indicators is more complicated. In addition, it should be taken into consideration that the world of scanner products is not static, i.e., dynamic scanner data sets related to product turnover (so-called *product churn*) are observed. For this reason, in the remainder of the paper, the consideration of price and quantity indicators will be transferred to scanner data case.

5. New Approach: Transitivity and Multilateral Indicators for Scanner Data Analysis

5.1 Transitivity

For international or inter-regional comparisons, *transitivity* means that estimates of price dynamics and quantities of selected attributes do not depend on the choice of underlying country or region. Similarly, for comparisons across firms, computing transitive price and quantity indices (indicators) do not depend on the choice of the firm-benchmark. A lack of index transitivity is a well-known problem in the

literature on international comparisons or scanner data (Gini, 1931; Eltetö and Köves, 1964; Szulc, 1964; Ivancic *et al.*, 2011; Chessa, 2016). By definition, *transitivity* eliminates the chain drift problem. To omit this problem one can expect that when all prices and quantities in a current period revert back to their values from the base period, then the index should indicate no price change and it should equal one. In the same situation, the price indicator should equal zero.

In the case of any price indicator IP and quantity indicator IQ , *transitivity* means that the following relationships occur for any $0 < s < t$:

$$IP_{0,s} + IP_{s,t} = IP_{0,t} \text{ and } IQ_{0,s} + IQ_{s,t} = IQ_{0,t} . \quad (9)$$

It is easy to verify that bilateral Bennet and Montgomery indicators are not transitive. In fact, using the *coffee* data set from the *PriceIndices* R package (Bialek, 2021; 2022), taking *March 2019* as the base period 0 , *April 2019* as an internal period s , and *May 2019* as the current period t , and running the *bennet()* and *montgomery()* package functions, we obtain:

$$IP_{0,s}^B + IP_{s,t}^B = 55529.61 \neq IP_{0,t}^B = 62461.82 , \quad (10)$$

and

$$IP_{0,s}^M + IP_{s,t}^M = 44803.48 \neq IP_{0,t}^M = 57397.08 . \quad (11)$$

5.2 The Transitive Bennet and Montgomery Indicators

Let us denote by $[0, T]$ the considered time interval, which typically (while computing multilateral indices) consists of 13 or 25 months (Eurostat, 2022). Let $G_{[0,T]}$ denote the set of available products in the whole interval $[0, T]$, i.e.,

$G_{[0,T]} = \bigcup_{\tau=0}^T G_{\tau}$. The following design of the multilateral Bennet and Montgomery indicators is an adaptation of the Fox's (2006) transformation for the scanner data case (Bialek and Pawelec (2024) for details). For comparison any periods τ and t let us generally denote by $IP_{\tau,t}$ the price (Bennet or Montgomery) indicator and by $IQ_{\tau,t}$ the quantity (Bennet or Montgomery) indicator (both calculated on the basis of products from a set $G_{[0,T]}$). Let us use the following, additional significations:

$$\overline{IP}_{\tau_0} = \frac{1}{T+1} \sum_{\tau=0}^T IP_{\tau,\tau_0} , \quad (12)$$

$$\overline{IQ}_{\tau_0} = \frac{1}{T+1} \sum_{\tau=0}^T IQ_{\tau,\tau_0} . \quad (13)$$

Now, having significations (12) and (13), the price and quantity multilateral indicators can be defined respectively:

$$MIP_{0,t} = \bar{IP}_t - \bar{IP}_0, \quad (14)$$

$$MIQ_{0,t} = \bar{IQ}_t - \bar{IQ}_0. \quad (15)$$

Please note that the multilateral indicators (14) and (15) are *transitive*, i.e.

$$MIP_{0,s} + MIP_{s,t} = \bar{IP}_s - \bar{IP}_0 + \bar{IP}_t - \bar{IP}_s = \bar{IP}_t - \bar{IP}_0 = MIP_{0,t}, \quad (16)$$

$$MIQ_{0,s} + MIQ_{s,t} = \bar{IQ}_s - \bar{IQ}_0 + \bar{IQ}_t - \bar{IQ}_s = \bar{IQ}_t - \bar{IQ}_0 = MIQ_{0,t}.$$

In the rest of the paper, we will distinguish multilateral Bennet indicators from multilateral Montgomery indicators using symbols, respectively: $MIP_{0,t}^B, MIQ_{0,t}^B$ and $MIP_{0,t}^M, MIQ_{0,t}^M$. It can be shown that indicators $MIP_{0,t}^B, MIQ_{0,t}^B$ satisfy the *sum test*, i.e.,

$$V_{G_{[0,T]}}^t - V_{G_{[0,T]}}^0 = MIP_{0,t}^B + MIQ_{0,t}^B, \quad (17)$$

and indicators $MIP_{0,t}^M, MIQ_{0,t}^M$ satisfy asymptotically the *sum test*, i.e.,

$$V_{G_{[0,T]}}^t - V_{G_{[0,T]}}^0 = \lim_{\varepsilon \rightarrow 0^+} (MIP_{0,t}^B(\varepsilon) + MIQ_{0,t}^B(\varepsilon)). \quad (18)$$

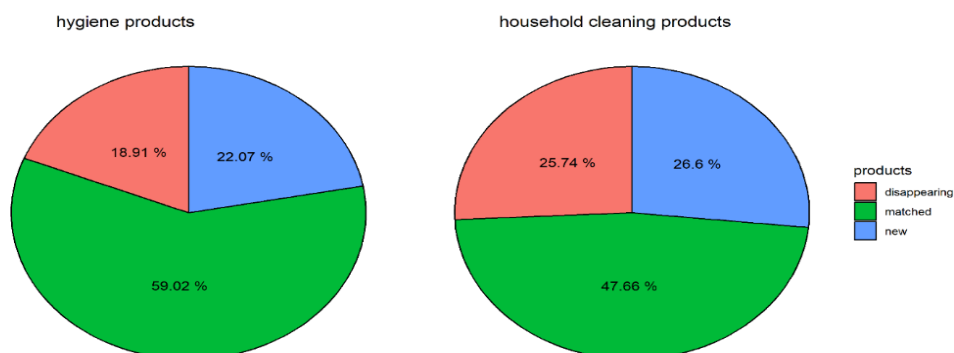
6. Empirical Study

In this empirical study we use scanner data obtained from single retail chain operating in Poland. We analyze monthly data regarding *hygiene products* and *household cleaning products* (COICOP 5 levels: 121322 and 056111 respectively). This data is from December 2022 to December 2023, respectively 390277 and 1132189 records.

The columnar variables, which are in the *hygiene products* and *household cleaning products* data set, are as follows: **time** - the date of the transaction, **prices** - prices of the sold product (PLN), **quantities** - quantities of the sold product; **prodID** - unique product codes assigned after the matching of products over time, **description** - a description of the sold products (i.e., the product label).

Figure 1 shows the new, matched and disappearing products detected for both scanner data sets during a 13 month time window.

Figure 1. A comparison of the structure of considered data sets due to new, disappearing and matched products.



Source: <https://cran.r-project.org/web/packages/PriceIndices/index.html>

As one can see, the numbers of disappearing and new goods are bigger in the case of household cleaning products and as a consequence, the number of matched products (over these 13 months) is bigger in the case of hygiene products (59.02% vs 47.66%). Thus, there is a need to use indicators which are not sensitive on such a large product turnover and resulting zero prices.

For compare our data we main use functions from the *PriceIndices* R Package. (Białek, 2022). One of the initial tasks was to develop the *data_selecting()* function, which involved compiling dictionaries with key words and phrases essential for distinguishing various product categories. Following this, the *data_classification()* function was employed to address products that posed classification challenges, requiring the manual creation of training samples from historical data.

Later on, product matching was executed by utilizing GTIN (Global Trade Item Number) barcodes, internal retail chain codes and product labels. This matching process was facilitated, specifically through the implementation of the *data_matching()* package function. Multilateral Bennet and Montgomery indicators were computed by using the *mbennet()* and *mmontgomery()* package functions.

6.1 Disaggregated vs Aggregated Data Description

The hygiene products consists 7 products subgroups: *toilet paper* (128 IDs), *wet wipes* (95 IDs), *sanitary wipes* (56 IDs), *baby diapers* (210 IDs), *tampons* (25 IDs), *sanitary pads* (80 IDs) and *hygiene pads* (20 IDs). The household cleaning products

consists 18 subgroups: *dishwasher tablets* (183 IDs), *pipe cleaner* (37 IDs), *stain remover* (85 IDs), *fabric softener powder* (2 IDs), *insecticide* (17 IDs), *toilet bowl cleaner* (203 IDs), *laundry detergent* (214 IDs), *laundry powder* (116 IDs), *laundry pods* (254 IDs), *dishwashing liquid* (188 IDs), *fabric softener* (362 IDs), *glass cleaner* (93 IDs), *furniture polish* (39 IDs), *shoe polish* (50 IDs), *floor cleaner* (88 IDs), *toilet cleaner* (113 IDs), *all-purpose cleaner* (37 IDs), *solvent* (18 IDs).

6.2 Data Filters Used

The datasets were filtered by using three typical filters from the *PriceIndices* package: the extreme price filter, the dump prices filter, the low sales filter. The extreme price filter rejects 1% of the lowest and 1% of the highest price changes comparing December 2022 to December 2023. The dump price filter rejects products with the price ratio being less than 0.7 and with the expenditure ratio being less than 0.7 in the same time and the low sales filter rejects products with relatively *low sale* in compared months.

Table 1 shows how data counts change depending on the use of a particular filter or all filters at once.

Table 1. *Change in count depending on the filter used*

<i>hygiene products</i>					
	without filters	extreme price filter	dump price filter	low sales filter	all filters
number of products	390 277	190 706	192 073	113 733	112 549
<i>household cleaning products</i>					
	without filters	extreme price filter	dump price filter	low sales filter	all filters
number of products	1 132 189	427 194	428 877	297 461	295 959

Source: Own elaboration in the *PriceIndices R* package.

For hygiene products, the use of all three filters resulted in a reduction of products by 277846 rows. If we are talking about household cleaning products, this change is much bigger, as we use of three filters, because it reduces nearly 80% of products.

6.3 Research Results

The effect of data filtering and aggregation on the differences between transitive Benet and Montgomery indicators is shown in Table 2 and Table 3. These tables show the detailed values of the transitive Benet and Montgomery indicators in five cases: (1) without use of any data filters; (2) after the use of extreme prices filter; (3) after the use of dump prices filter; (4) after the use of low sales filter; and (5) after the application of all data filters. These data are presented in two variants: the first variant concerns the lowest data aggregation level, i.e. the GTIN code level, and the

second variant, considers a homogeneous product more broadly, i.e. on the basis of the COICOP 6 category.

Table 2. Comparison of the multilateral Bennet and Montgomery indicators across data aggregation level and data filtering for hygiene products (all available products are considered, the normalized values are in brackets)

Multilateral Bennet indicator					
GTIN level					
characteristics	without filters	extreme price filter	dump price filter	low sales filter	all filters
sales value	-5807345 (100)	-5024354 (100)	-4774773 (100)	-4231513.4 (100)	-4501442.05 (100)
price indicator	1108525 (-19.09)	323699 (-6.44)	187465.9 (-3.93)	-1115400.55 (26.36)	-973784.79 (21.63)
quantity indicator	-6915870 (119.09)	-5348053 (106.44)	-4962239 (103.93)	-3116112.85 (73.64)	-3527657.26 (78.37)
COICOP 6 level					
characteristics	without filters	extreme price filter	dump price filter	low sales filter	all filters
sales value	-5807345 (100)	-4169073 (100)	-4774773 (100)	-2930963 (100)	9893952.63 (100)
price indicator	244962.3 (-4.22)	-224275.2 (5.38)	-91943.86 (1.93)	-303933.7 (10.37)	4781424.23 (48.33)
quantity indicator	-6052307 (104.22)	-3944798 (94.62)	-4682829 (98.07)	-2627030 (89.63)	5112528.4 (51.67)
Multilateral Montgomery indicator					
GTIN level					
characteristics	without filters	extreme price filter	dump price filter	low sales filter	all filters
sales value	-5807345 (100)	-5024354.28 (100)	-4774772.71 (100)	-4231513.4 (100)	-4501442.05 (100)
price indicator	966705.56 (-16.65)	258197.17 (-5.14)	148426.46 (-3.11)	-973717.05 (23.01)	-861018.08 (19.13)
quantity indicator	-6774050.29 (116.65)	-5282551.45 (105.14)	-4923199.17 (103.11)	-3257796.35 (76.99)	-3640423.97 (80.87)
COICOP 6 level					
characteristics	without filters	extreme price filter	dump price filter	low sales filter	all filters
sales value	-5807345 (100)	-5024354 (100)	-4774773 (100)	-2930963 (100)	9893952.63 (100)
price indicator	232706.28 (-4.01)	258197.2 (-5.14)	-97453.59 (2.04)	-304561.2 (10.39)	3586784.89 (36.25)
quantity indicator	-6040051 (104.01)	-5282551 (105.14)	-4677319 (97.96)	-2626402 (89.61)	6307167.74 (63.75)

Source: Own elaboration in the PriceIndices R package.

Table 3. Comparison of the Bennet indicators and Montgomery indicators across data aggregation level and data filtering for household cleaning products (the normalized values are in brackets)

Multilateral Bennet indicator					
GTIN level					
characteristics	without filters	extreme price filter	dump price filter	low sales filter	all filters
sales value	4050926 (100)	1846788 (100)	1682243 (100)	1531750 (100)	1695315 (100)
price indicator	7156013 (176.65)	3365233 (182.22)	3445489 (204.82)	1213404 (79.22)	1208138 (71.26)
quantity indicator	-3105088 (-76.65)	-1518445 (-82.22)	-1763247 (-104.82)	318346.3 (20.78)	487177.1 (28.74)
COICOP 6 level					
characteristics	without filters	extreme price filter	dump price filter	low sales filter	all filters
sales value	4050926 (100)	1729723 (100)	1682243 (100)	1173459 (100)	-443792.4 (100)
price indicator	5200080 (128.37)	2824730 (163.31)	3108084 (184.76)	2389376 (203.62)	1424843 (-321.06)
quantity indicator	-1149155 (-28.37)	-1095007 (-63.31)	-1425842 (-84.76)	-1215917 (-103.62)	-1868636 (421.06)
Multilateral Montgomery indicator					
GTIN level					
characteristics	without filters	extreme price filter	dump price filter	low sales filter	all filters
sales value	4050926 (100)	1846788 (100)	1682243 (100)	1531750 (100)	1695315 (100)
price indicator	6069437 (149.83)	2859077 (154.81)	2914184 (173.23)	1167059 (76.19)	1167627 (68.87)
quantity indicator	-2018511 (-49.83)	-1012289 (-54.81)	-1231941 (-73.23)	364691.1 (23.81)	527688.2 (31.13)
COICOP 6 level					
characteristics	without filters	extreme price filter	dump price filter	low sales filter	all filters
sales value	4050926 (100)	1729723 (100)	1682243 (100)	1173459 (100)	-443792.4 (100)
price indicator	5144017 (126.98)	2778356 (160.62)	3057751 (181.77)	2343151 (199.68)	1516226 (-341.65)
quantity indicator	-1093091 (-26.98)	-1048633 (-60.62)	-1375508 (-81.77)	-1169692 (-99.68)	-1960018 (441.65)

Source: Own elaboration in the PriceIndices R package.

7. Conclusions

The use of data filters, regardless of the data set used, level of data aggregation considered and indicator formula adopted, shallows the difference between the value of price and quantity indicators. Sometimes, data filtering even leads to a change in

the sign of the indicator (more often price than quantity), which can be observed in particular for the low sales filter and hygiene products (Table 2). Observing the results obtained for both scanner data sets, we can draw the recurring conclusion that the transitive Bennet formula leads to greater differences in values between the price and quantity indicators than the transitive Montgomery formula.

This conclusion is identical for each type of data filter and for each data aggregation level (Table 2 and Table 3). However, it can be seen that for a higher level of data aggregation (COICOP level 6), the differences between the corresponding multilateral Bennet and Montgomery indicators are much smaller (regardless of the scanner dataset) than for the barcode level (GTIN level).

In principle, it can be said that at a higher level of aggregation, the Bennet and Montgomery multilateral indicators lead to fairly similar values and any differences between their values are substantial at the disaggregated level. Please note that the above conclusions should be regarded as preliminary, since they concern only a specific group of products, one year of observation and one retail chain. In the future, the authors of the paper plan to expand the study to other product groups by obtaining data from more retail chains.

References:

- Balk, B.M., Färe, R., Grosskopf, S. 2004. The theory of economic price and quantity indicators. *Economic Theory*, 23, 149-164.
- Bennet, T.L. 1920. The theory of measurement of changes in cost of living. *Journal of the Royal Statistical Society*, 83(3), 455-462.
- Białek, J. 2021. Priceindices – a new R package for bilateral and multilateral price index calculations. *Statistika – Statistics and Economy Journal*, 36(2), 122-141.
- Białek, J. 2022. Scanner data processing in a newest version of the Price Indices package. *Statistical Journal of the IAOS*, 38(4), 1369-1397.
- Białek, J., Pawelec, N. 2024. The use of the transitive Montgomery indicators for scanner data analysis. *Argumenta Oeconomica*, 2(53), (in press).
- Chessa, A.G. 2016. A new methodology for processing scanner data in the Dutch CPI. *Eurostat review of National Accounts and Macroeconomic Indicators*, 1(2016), 49-69.
- Cross, R.M., Färe, R. 2009. Value data and the Bennet price and quantity indicators. *Economics Letters*, 102(1), 19-21.
- De Boer, P., Rodrigues, J.F. 2020. Decomposition analysis: when to use which method? *Economic Systems Research*, 32(1), 1-28.
- Diewert, W.E. 2005. Index number theory using differences rather than ratios. *American Journal of Economics and Sociology*, 64(1), 311-360.
- Eltető, O., Köves, P. 1964. On a problem of index number computation relating to international comparison. *Statisztikai Szemle*, 42(10), 507-518.
- Eurostat. 2022. Guide on Multilateral Methods in the Harmonised Index of Consumer Prices. Luxembourg: Publications Office of the European Union.
- Fox, K.J. 2006. A method for transitive and additive multilateral comparisons: A transitive Bennet indicator. *Journal of Economics*, 87, 73-87.

- Gini, C. 1931. On the circular test of index numbers. *Metron*, 9(9), 3-24.
- International Labour Office. 2004. Consumer price index manual: Theory and practice. Geneva.
- International Monetary Fund. 2020. Consumer Price Index manual: Concepts and methods. Washington, D.C.
- 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.
- Montgomery, J.K. 1929. Is There a Theoretically Correct Price Index of a Group of Commodities? *L' Universale Tipogr, Poliglotta*.
- Szulc, B. 1964. Indices for multiregional comparisons. *Przegląd statystyczny*, 3, 239-254.
- Vartia, Y.O. 1976. Ideal Log-Change Index Numbers. *Scandinavian Journal of Statistics* 3, 121-126.
- Von der Lippe, P.M. 2007. Index theory and price statistics (Vol. 817). Frankfurt am Main: Peter Lang.

Załączniki – Oświadczenia współautorów

Jacek Białek

imię i nazwisko

Uniwersytet Łódzki,

Katedra Metod Statystycznych

afiliacja

Łódź, 22.05.2026.

miejsowość i data

OŚWIADCZENIE

Oświadczam, że w pracy Białek J., Pawelec N. (2024), *Proposition of the Quality Adjusted Geks-Type Price Index*, „Statistika: Statistics & Economy Journal”, vol. 105(3).

.....
(autorzy, rok wydania, tytuł, czasopismo lub wydawca, tom, strony)

mój udział polegał na:

- dostarczeniu danych skanowanych do badania empirycznego;
- współpracy w zakresie opracowania warstwy metodologicznej artykułu;
- współpracy w zakresie interpretacji wyników i wnioskowania;

Swój udział szacuję na 20%.

(osoba ubiegająca się o nadanie stopnia doktora, powinna opisać szczegółowo swój udział w powstaniu pracy).

.....

podpis

Załącznik nr 2

Sylwia Roszkowska

imię i nazwisko

Uniwersytet Łódzki

afiliacja

Łódź, dnia 28 maja 2026 r.

miejsowość i data

OŚWIADCZENIE

Oświadczam, że w pracy Białek J., Pawelec N., Roszkowska S. (2021), *Estimating the elasticity of substitution when compiling the CES cost of living index on scanner data*, „Quality & Quantity”, vol. 56(6), s. 5997–6021.

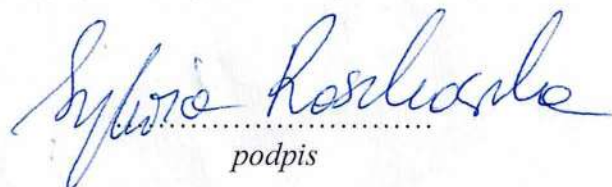
(autorzy, rok wydania, tytuł, czasopismo lub wydawca, tom, strony)

mój udział polegał na:

- opracowaniu zagadnień dotyczących funkcji CES w ujęciu teoretycznym;
- analizę i opis zastosowania funkcji CES w części empirycznej artykułu;
- udział w interpretacji wyników dotyczących elastyczności substytucji oraz indeksu CES;
- współtworzenie części wniosków końcowych wynikających z przeprowadzonej analizy.

Mój udział szacuję na 33,3%.

(osoba ubiegająca się o nadanie stopnia doktora, powinna opisać szczegółowo swój udział w powstaniu pracy).


.....
podpis

Jacek Białek

Łódź, 22.05.2026

imię i nazwisko

miejsowość i data

Katedra Metod Statystycznych

afiliacja

OŚWIADCZENIE

Oświadczam, że w pracy Białek J., Pawelec N., Roszkowska S. (2021), *Estimating the elasticity of substitution when compiling the CES cost of living index on scanner data*, „Quality & Quantity”, vol. 56(6), s. 5997–6021.

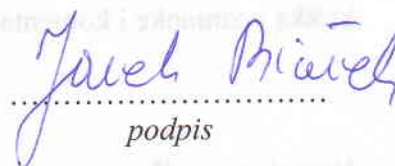
.....
(autorzy, rok wydania, tytuł, czasopismo lub wydawca, tom, strony)

mój udział polegał na:

- dostarczeniu danych skanowanych do badania empirycznego;
- współpracy w zakresie opracowania warstwy metodologicznej artykułu;
- współpracy w zakresie interpretacji wyników i wnioskowania;

Swój udział szacuję na 33%

(osoba ubiegająca się o nadanie stopnia doktora, powinna opisać szczegółowo swój udział w powstaniu pracy).


.....
podpis

Jacek Bialek

imię i nazwisko

Uniwersytet Łódzki,
Katedra Metod Statystycznych

afiliacja

Łódź, 22.05.2026.

miejsowość i data

OŚWIADCZENIE

Oświadczam, że w pracy *Bialek J., Pawelec N. (2024), The use of transitive Montgomery indicators for scanner data analysis, „Argumenta Oeconomica”, vol. 53(2), s. 1–13.*

.....
(autorzy, rok wydania, tytuł, czasopismo lub wydawca, tom, strony)

mój udział polegał na:

- dostarczeniu danych skanowanych do badania empirycznego;
- współpracy w zakresie opracowania warstwy metodologicznej artykułu;
- współpracy w zakresie interpretacji wyników i wnioskowania;

Swój udział szacuję na 20%.

(osoba ubiegająca się o nadanie stopnia doktora, powinna opisać szczegółowo swój udział w powstaniu pracy).


.....
podpis

Jacek Białek
imię i nazwisko
Uniwersytet Łódzki,
Katedra Metod Statystycznych
afiliacja

Łódź, 22.05.2026.
miejsowość i data

OŚWIADCZENIE

Oświadczam, że w pracy Białek J., Pawelec N. (2024), *Comparison of transitive Bennet and Montgomery indicators on the basis of scanner data sets*, „European Research Studies Journal”, vol. XXVII(4), s. 472–485.

.....
(autorzy, rok wydania, tytuł, czasopismo lub wydawca, tom, strony)

mój udział polegał na:

- dostarczeniu danych skanowanych do badania empirycznego;
- współpracy w zakresie opracowania warstwy metodologicznej artykułu;
- współpracy w zakresie interpretacji wyników i wnioskowania;

Swój udział szacuję na 20%.

(osoba ubiegająca się o nadanie stopnia doktora, powinna opisać szczegółowo swój udział w powstaniu pracy).


.....
podpis