

Detekcja „ukrytej” inflacji oraz deflacji na bazie danych skanowanych

Jacek Białek ^{1,2}, Anna Bobel ², Dagmara Oprych-Franków ²

1. Uniwersytet Łódzki, 2. Główny Urząd Statystyczny

V Kongres Statystyki Polskiej, Warszawa, 2025

Agenda

- 1 Dane skanowane jako nowoczesne źródło danych w pomiarze CPI
- 2 "Klasyczna" shrinkflacja
- 3 Nowa, proponowana klasyfikacja *downsizingu* i *upsizingu*
- 4 Propozycja procedury detekcji dołączonej do R pakietu *PriceIndices*
- 5 Wyniki badania empirycznego
- 6 Wnioski

Podstawowe informacje na temat danych skanowanych

Dane skanowane

Dane skanowane są to „dane pochodzące z elektronicznych terminali w punktach sprzedaży sieci handlowych, które dostarczają informacji o sprzedawanych produktach na poziomie kodu kreskowego GTIN (ang. Global Trade Item Number) lub EAN (ang. European Article Number), w których zawarte są dane o cenach i poziomie konsumpcji produktów”).

	data transakcji	outlet	segment	kategoria	nr produktu	EAN	etykieta	cena	ilosc
1	2024-03-21	00-199	RYBY SAMOBSLUGA	PROD. RYBNE PRZETWORZONE	32994	7.311170e+12	PASTA Z TUŃCZYKA 145G ABBA	9.98	23.00
2	2024-03-21	00-199	WARZYWA	WARZYWA	34021	2.200031e+11	POMIDOR UKŁADANY LUZ	12.97	179.98
3	2024-03-21	00-199	OWOCE	OWOCE PODSTAWOWE	34041	2.202055e+11	BANAN LUZ	3.72	2212.75
4	2024-03-21	00-199	ZDROWA ZYWNOSC	ZYWNOSC EKOLOGICZNA	81189	5.905699e+12	BIO SYROP MALINOWY 500ML Z DOMU REMBOWSKICH	25.89	4.00
5	2024-03-21	00-199	KONSERWY ZUPY DANIA GOTOW	PASZTETY	103793	3.596710e+12	) PASZTET Z ŻOŁĄDK. DROBIOWY. 180G AUCHAN MP	5.49	4.00
6	2024-03-21	00-199	WARZYWA	WARZYWA GOTOWE	143537	5.900449e+12	SURÓWKA Z MARCHEWKI 300G MAGA	3.59	46.00
7	2024-03-21	00-199	DESERY I DODATKI	BUDYN	170498	5.900983e+12	BUDYŃ MLECZNA CZEKOLADA KLEKS 42G DELECTA	1.34	5.00
8	2024-03-21	00-199	HIGIENA TOALETOWA	MYDŁA W KOSTCE	188510	5.900536e+12	MYDŁO W KOSTCE COTTON 90G LUKSJA	2.18	17.00
9	2024-03-21	00-199	TLUSZCZE, MLEKO, JAJA	TLUSZCZE	188827	4.001954e+12	KERRYGOLD MASŁO 200G	6.59	96.00
10	2024-03-21	00-199	PRODUKTY MACZNO ZBOZOWE	KASZA	284681	5.906827e+12	KASZA PĘCZAK KUJAWSKI 0.9KG MELVIT	5.09	13.00

Zalety i korzyści:

- ❶ dane skanowane mają ogromną objętość i dostarczają wielu informacji na temat zakupionych produktów
- ❷ wykorzystywanie danych skanowanych jest relatywnie tanie i z reguły niemal w pełni zautomatyzowane
- ❸ zbiory danych są kompletne na najniższym poziomie agregacji
- ❹ dostarczają informacje o rzeczywistych wydatkach dla sprzedanych pozycji

	okres	rozmiar	liczba_rekordów	segmenty	kategorie	kody_EAN
1	marzec 2024	738 MB	5458988	125	626	154821

Wyzwania:

- 1 automatyczna klasyfikacja produktów do grup COICOP 5 lub niżej (*machine-learning, text-mining*)
- 2 parowanie produktów w czasie (*data matching*)
- 3 filtrowanie i agregowanie danych (*data filtering, data aggregating*)
- 4 wybór wskaźnika cenowego (*multilateral indices*)

Ostatnio coraz więcej uwagi poświęca się problemowi obciążania pomiaru wynikającego z występowaniem zjawiska **shrinkflacji** (*shrinkflation, downsizing*). My rozszerzyliśmy te pojęcia wprowadzając też pojęcie *upsizing* i zdefiniowaliśmy jego różne rodzaje. A zatem wyzwaniem nad którym skupimy się w prezentacji będzie **automatyczna detekcja downsizingu i upsizingu**.

Chcemy m.in. dokonywać detekcji takich przypadków:



"Klasyczna" shrinkflacja

Ekonomiczny termin **shrinkflacja** został użyty po raz pierwszy w roku 2009 przez brytyjską ekonomistkę Pippę Malmgren. Jest to proces zmniejszania rozmiaru lub ilości produktów przy zachowaniu tych samych cen.

Bardziej precyzyjna definicja ekonomiczna: shrinkflacja to wzrost ogólnego poziomu cen towarów na jednostkę wagi lub objętości spowodowany zmniejszeniem wagi lub rozmiaru sprzedawanego przedmiotu.

Proponowana definicja shrinkflacji: terminu shrinkflacja będziemy używać, gdy zmniejszeniu rozmiaru produktu towarzyszy brak zmiany jego ceny (wąskie, istniejące znaczenie) lub (pozorny) spadek jego ceny (szersze znaczenie), co w konsekwencji oznacza, że cena jednostkowa takiego produktu wzrosła.

"Klasyczna" shrinkflacja

Shrinkflation

STATBEL

- Shrinkflation or product relaunches where the price changes proportionally less than the packing size are problematic in scanner data
- Products tend to get a new unique product identifier
- Typically matched-methods are used with scanner data, these product relaunches are missed

Cat-astrophe? Whiskas pet food criticised over shrinking servings

Cat food brand caught in shrinkflation storm as owners say new pouches and recipe leaves pets hungry



(The Guardian, 2023)

Shrinkflacja w tradycyjnej kolekcji danych - przykłady międzynarodowe:

Table: Przykłady downsizingu

Firma	Produkt	Data	Redukcja rozmiaru	Redukcja rozmiaru [%]
Nestlé	After Eight	2010	From 200 g to 170 g	15
Burton' Biscuit	Cadbury Fingers	2015	From 125 g to 111 g	11.2
Milka	Milka Nuts & Raisins	2017	From 300 g to 270 g	10
Unilever	Ben & Jerry ice-cream	2020	From 500 ml to 465 ml	7
General Mills	Boxes of cereal	2021	From 19.3 oz to 18.1 oz	6.2
Procter & Gamble	Roll of toilet paper	2022	From 264 to 244 sheets	7.6
Mars, Incorporated	Whiskas cat food	2023	From 100 g to 85 g	15

Shrinkflacja w tradycyjnej kolekcji danych - przykłady z Polski (2023):

	product	size change
1	Makaron bezajeczny	Lubella z 500 na 400 g
2	Jogurt naturalny	Bakoma z 390 na 370 g, ZOTT 370 na 400 i odwrotnie
3	Jogurt pitny	Bakoma z 380 na 370 g, Danone z 290 na 270 g
4	Margaryna (tłuszcz roślinny) do smarowania pieczywa	Rama z 450 na 400 g, Delma z 500 na 450 g
5	Żelki	Haribo 100 na 85 g; 175 na 160 g
6	Musztarda	Kamis sarepska z 200 na 185 g
7	Mleko w proszku dla niemowląt	Nan z 800 na 650 g
8	Woda smakowa, niegazowana	Żywiec 1,5 na 1,2 l
9	Cydr o zawartości 2%-7% alkoholu, w butelce	Cydr lubelski z 1000 na 750 ml
10	Wódka czysta	z 50 na 40 ml
11	Koniak, brandy lub whiskey	z 50 na 40 ml
12	Płyn lub koncentrat do płukania tkanin	różne firmy np. Silan z 1800 na 1450 ml, lenor 1360 na 1230 ml, 1800 na 1600 ml
13	Środek do udrożniania rur odpływowych	Kret 400 na 560 g - zwiększenie ale dyn. w górę
14	Płyn do prania tkanin delikatnych	Perwoll 900 na 960 ml - zwiększenie ale dyn. w górę; w maju 960 na 990 ml
15	Kapsułki do prania	różne firmy np. Ariel 33 na 30 szt., 26 na 24 szt., Persil 28/30 na 33 szt.
16	Kefir	Bakoma z 390 na 370 g
17	Ogórki konserwowe całe, w słoiku	poj. przypadki z 900 na 850, 860 g
18	Koncentrat pomidorowy 25-30%, w słoiku	Dawtona z 200 na 190 g, Pudliszko ze 195 na 200 g i odwrotnie
19	Czekolada gorzka, min.60% kakao, bez dodatków	Wawel ze 100 na 90 g
20	Baton wieloskładnikowy	poj. przypadki np. z 51 na 50 g

Zauważmy, że *shrinkflacja* i *skimpflacja* to nie to samo zjawisko!

Inflation



Zutaten: Tomatensauce, **ital. Mozzarella**, Champignons, Paprika, Salami

~~1,00 €~~
1,50 €

Shrinkflation



Zutaten: Tomatensauce, **ital. Mozzarella**, Champignons, Paprika, Salami

1,00 €

Skimpflation



Zutaten: Tomatensauce, **Gouda und Tilsiter**, Champignons, Paprika, Salami

1,00 €

Nowa, proponowana klasyfikacja *downsizingu* i *upsizingu*

Table: Kategorie *downsizingu* i *upsizingu* z okresu $\tau - 1$ na okres τ

<i>Downsizing</i>			
zmiana rozmiaru	zmiana ceny (unit value)	zmiana ceny jednostkowej	kategoria zjawiska
$r_i^w(\tau - 1, \tau) < 1$	$r_i^p(\tau - 1, \tau) \leq 1$	$r_i^u(\tau - 1, \tau) > 1$	shrinkflation
$r_i^w(\tau - 1, \tau) < 1$	$r_i^p(\tau - 1, \tau) < 1$	$r_i^u(\tau - 1, \tau) < 1$	shrinkdeflation
$r_i^w(\tau - 1, \tau) < 1$	$r_i^p(\tau - 1, \tau) > 1$	$r_i^u(\tau - 1, \tau) > 1$	sharkflation
<i>Upsizing</i>			
zmiana rozmiaru	zmiana ceny (unit value)	zmiana ceny jednostkowej	kategoria zjawiska
$r_i^w(\tau - 1, \tau) > 1$	$r_i^p(\tau - 1, \tau) \geq 1$	$r_i^u(\tau - 1, \tau) < 1$	unshrinkdeflation
$r_i^w(\tau - 1, \tau) > 1$	$r_i^p(\tau - 1, \tau) > 1$	$r_i^u(\tau - 1, \tau) > 1$	unshrinkflation
$r_i^w(\tau - 1, \tau) > 1$	$r_i^p(\tau - 1, \tau) < 1$	$r_i^u(\tau - 1, \tau) < 1$	sharkdekflation

Propozycja procedury detekcji dołączona do R pakietu *PriceIndices*

❶ Dopasowanie produktów w czasie:

```
df <- data_matching(data = dataDOWNSIZED,  
start="2024-01", end="2024-02",  
codeIN = TRUE, codeOUT = TRUE, description = TRUE,  
onlydescription = FALSE, precision = 0.9, interval=FALSE)
```

❷ Ekstrakcja gramatury i jednostki sprzedaży:

```
df <- data_unit(df, units = c("g|ml|kg|l"), multiplication = "x")
```

❸ Normalizacja cen i ilości:

```
df<-data_norm(df, rules = list(c("ml","l",1000),  
c("g","kg",1000)))
```

❹ Detekcja downsizingu i upsizingu oraz ich charakterystyka:

```
result <- shrinkflation(data = df, start = "2024-01",end =  
"2024-02", prec = 3, interval = FALSE, type="sharkflation)
```

Dane od sieci handlowej działającej w Polsce (ponad 510 outletów).
Rozważany okres czasu: grudzień 2021 - grudzień 2023. Uwzględniono następujące grupy elementarne produktów:

- a) kategoria spożywcza: **jogurt** (COICOP 5: 011441), **ryż** (COICOP 5: 011111), **kasze** (COICOP 5: 011123), **pieczywo** (COICOP 5: 011131), **kawa** (COICOP 5: 012111);
- b) kategoria niespożywcza: **kosmetyki i środki higieny osobistej** (COICOP 5: 121321), **akcesoria (karma) dla zwierząt** (COICOP 5: 093421).

Table: Automatyczna detekcja: *downsizing* vs *upsizing*

<i>Downsizing</i>							
typ: dostępne produkty	jogurt 524	ryż 68	kasze 65	pieczywo 83	kawa 704	kosmetyki 9045	karma 915
shrinkflation	7 (2.765)	-	-	-	-	11 (0.128)	1 (0.080)
shrinkdeflation	4 (1.692)	-	-	-	-	-	-
sharkflation	10 (4.544)	1 (1.169)	1 (1.164)	-	-	1 (0.001)	-
<i>Upsizing</i>							
typ:	jogurt	ryż	kasze	pieczywo	kawa	kosmetyki	karma
unshrinkdeflation	-	-	-	-	-	9 (0.096)	1 (0.070)
unshrinkflation	1 (0.023)	-	-	-	-	-	-
sharkdeflation	1 (0.695)	1 (0.967)	1 (6.851)	-	1 (0.013)	2 (0.009)	-

(★) Liczba przypadków (udział w sprzedaży [%])

Table: Przykładowe wykryte przypadki: *downsizing* vs *upsizing*

<i>Downsizing</i>						
typ	grupa	opisy przed/po (★)	data zmiany	zmiana rozmiaru (p.p.)	zmiana oryginalnej ceny (p.p.)	zmiana ceny jednostkowej (p.p.)
shrinkflation	cosmetics	szampon do włosów 400ml zielarki, szampon do włosów 350 ml zielarki	2023-08, 2023-09	-12.50	-6.97	6.32
shrinkdeflation	yoghurt	jogurt grecki 180G Bakoma, jogurt grecki 170G Bakoma	2022-09, 2022-10	-5.55	-12.18	-6.99
sharkflation	yoghurt	Jogurt naturalny 2% 175G LOKAL, Jogurt naturalny 2% 170G LOKAL	2022-05, 2022-06	-2.86	7.03	10.20
<i>Upsizing</i>						
typ	grupa	opisy przed/po (★)	data zmiany	zmiana rozmiaru (p.p.)	zmiana oryginalnej ceny (p.p.)	zmiana ceny jednostkowej (p.p.)
unshrinkdeflation	pet food	karma z kurcz/wątrób w sos 85g, karma z kurcz/wątrób w sos 100g	2022-09, 2022-10	17.64	10.00	-6.38
unshrinkflation	yoghurt	Jogurt naturalny 2% 170G, Jogurt naturalny 2% 175G LOKAL	2022-04, 2022-05	2.94	4.22	1.20
sharkdeflation	coffee	kawa rozp zbożowa 100g AUCH, kawa rozp. zbożowa 200g AUCH	2022-01, 2022-02	100	-51.52	-75.76

Table: Charakterystyka zmian cen jednostkowych i rozmiaru produktów (yogurt)

Downsizing								
Statystyki:		Statystyki dla zmian cen jednostkowych				Statystyki dla zmian rozmiaru produktu		
typ	średnia (p.p.)	zmiana	SD zmian (p.p.)	VC zmian (%)	średnia (p.p.)	zmiana	SD zmian (p.p.)	VC zmian (%)
shrinkflation	3.755		1.996	53.1	-5.733		1.655	28.9
shrinkdeflation	-3.08		2.699	87.6	-5.556		0.786	14.1
sharkflation	6.066		3.186	52.5	-3.787		1.458	38.5
Upsizing								
Statystyki:		Statystyki dla zmian cen jednostkowych				Statystyki dla zmian rozmiaru produktu		
typ	średnia (p.p.)	zmiana	SD zmian (p.p.)	VC zmian (%)	średnia (p.p.)	zmiana	SD zmian (p.p.)	VC zmian (%)
unshrinkflation	1.200	-	-	-	2.941	-	-	-
sharkdeflation	-7.146	-	-	-	5.882	-	-	-

Table: Charakterystyka zmian cen jednostkowych i rozmiaru produktów (kosmetyki i środki higieny osobistej)

Downsizing								
Statystyki:	Statystyki dla zmian cen jednostkowych				Statystyki dla zmian rozmiaru produktu			
typ	średnia (p.p.)	zmiana	SD zmian (p.p.)	VC zmian (%)	średnia (p.p.)	zmiana	SD zmian (p.p.)	VC zmian (%)
shrinkflation	78.973		28.866	36.5	-26.751		29.129	108.9
sharkflation	77.994		-	-	-62.5		-	-
Upsizing								
Statystyki:	Statystyki dla zmian cen jednostkowych				Statystyki dla zmian rozmiaru produktu			
typ	średnia (p.p.)	zmiana	SD zmian (p.p.)	VC zmian (%)	średnia (p.p.)	zmiana	SD zmian (p.p.)	VC zmian (%)
unshrinkdeflation	-44.587		28.509	63.9	42.341		33.350	78.7
sharkdeflation	-21.658		10.040	46.4	6.786		4.546	67

Table: Średnia absolutna różnica (p.p.) pomiędzy indeksem policzonym na oryginalnym zbiorze danych a indeksem wyznaczonym na zbiorze zredukowanym o wykryte produkty (jogurt)

<i>Downsizing</i>										
typ / indeks cen	Jevons	Chain Jevons	Fisher	Chain Fisher	GEKS	CCDI	GEKS-L	GEKS-AQU	Geary-Khamis	TPD
shrinkflation	0.325	0.389	0.217	0.267	0.218	0.222	0.231	0.225	0.204	0.285
unshrinkflation	0.038	0.023	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
sharkflation	0.169	0.306	0.210	0.333	0.220	0.224	0.229	0.229	0.222	0.344
<i>Upsizing</i>										
typ/indeks cen	Jevons	Chain Jevons	Fisher	Chain Fisher	GEKS	CCDI	GEKS-L	GEKS-AQU	Geary-Khamis	TPD
shrinkdeflation	0.167	0.209	0.111	0.146	0.111	0.114	0.121	0.122	0.103	0.154
sharkdeflation	0.056	0.039	0.060	0.059	0.060	0.055	0.093	0.073	0.041	0.038

Table: Średnia absolutna różnica (p.p.) pomiędzy indeksem policzonym na oryginalnym zbiorze danych a indeksem wyznaczonym na zbiorze zredukowanym o wykryte produkty (kosmetyki i środki higieny osobistej)

<i>Downsizing</i>										
typ / indeks cen	Jevons	Chain Jevons	Fisher	Chain Fisher	GEKS	CCDI	GEKS-L	GEKS- AQU	Geary- Khamis	TPD
shrinkflation	0.044	0.051	0.011	0.033	0.020	0.034	0.017	0.019	0.039	0.040
sharkflation	0.021	0.020	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000
<i>Upsizing</i>										
typ/indeks cen	Jevons	Chain Jevons	Fisher	Chain Fisher	GEKS	CCDI	GEKS-L	GEKS- AQU	Geary- Khamis	TPD
unshrinkdeflation	0.076	0.047	0.021	0.012	0.020	0.039	0.014	0.018	0.039	0.039
sharkdeflation	0.001	0.002	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.001

Podsumowując uzyskane wyniki można stwierdzić, że zasadniczo zaobserwowaliśmy **trzy grupy produktów**:

(1) produkty bez lub z marginalnym występowaniem *downsizingu* i *upsizingu* (np. **ryż, kasza, pieczywo, kawa i karma dla zwierząt domowych**);

(2) produkty o stosunkowo niewielkiej ilości produktów których dotyczy *downsizing* i *upsizing*, ale charakteryzujące się dużymi zmianami rozmiaru i zauważalnymi dla konsumenta zmianami cen jednostkowych (np. **kosmetyki i produkty higieniczne**);

(3) produkty ze stosunkowo dużą liczbą produktów gdzie występuje *downsizing* i/lub *upsizing*, ale charakteryzujące się niewielkimi zmianami rozmiaru i nieznacznymi zmianami cen jednostkowych (np. **jogurty**).

(II) Wstępny wniosek jest taki, że wpływ na wskaźnik cen produktów o nawet dużych zmianach rozmiaru i dużych zmianach cen nie jest znaczący, jeśli produkty te mają stosunkowo niski udział w sprzedaży (jak w przypadku **kosmetyków**). Ta uwaga jest zrozumiała w przypadku indeksów ważonych, ale w dużej mierze odnosi się również do formuł nieważonych.

(II) Wpływ produktów typu *downsized* i *upsized* na błąd pomiaru inflacji może być jednak duży, jeśli ich udział w sprzedaży wynosi przynajmniej kilka procent, nawet jeśli zmiany wielkości i ceny jednostkowej są dość niewielkie (jak w przypadku **jogurtów**).

(III) Biorąc pod uwagę fakt, że pokrycie koszyka CPI danymi ze skanera koszyka CPI danymi ze skanerów sięga w krajach UE 25-30%, a nawet więcej, można łatwo wyobrazić sobie zagrożenia dla dokładności pomiaru inflacji wynikające z *downsizingu* i *upsizingu*.

Wstępne wnioski wymagające dalszych badań:

- **Indeksy multilateralne** generują mniejsze obciążenie z powodu *downsizingu* i *upsizingu* niż indeksy bilateralne.
- Wśród indeksów multilateralnych indeks **TPD** wydaje się być najbardziej wrażliwy na występowanie produktów których dotyczy *downsizing* i *upsizing*.
- **Indeksy łańcuchowe** mają tendencję do eskalacji obciążenia wynikającego ze zjawiska shrinkflacji.

Dziękujemy za uwagę!