Immeasurability of shrinkflation in the CPI? Automatic downsizing detection using scanner data.

Jacek Białek [1 2], Anna Bobel [1], Dagmara Oprych – Franków [1]

[1]Statistics Poland [2]University of Lodz

Definition of shrinkflation

- In economics, **shrinkflation**, also known as **package downsizing** or **weight-out**, is the process of items shrinking in size or quantity while the prices remain the same. **Skimpflation** involves a reformulation or other reduction in quality. The term **shitflation**, which refers to the case of maintaining the price of a product while reducing quality, can also be found in the literature (Giner, 2022). This term has already been used by online communities such as Reddit or Twitter since 2020.
- A more precise economic definition: **shrinkflation** is a rise in the general price level of goods per unit of weight or volume brought about by a reduction in the weight or size of the item sold.
- **Our definition of shrinkflation:** We will use the term **shrinkflation** when a reduction in the size of a product is accompanied by either no change in its price (narrow, existing meaning) or an apparent decrease in its price (broader meaning), which consequently means that the unit price of such a product has increased.

Proposal for an automatic procedure to detect downsizing when working on scanner data

Input: **Data** - a typical scanner data set, containing columns: **time**, **prices**, **quantities**, product codes (like **codeIN** - the retailer code or **codeOUT** - the external code, like GTIN or SKU) and **description**, which (optionally) contains information about the product's size (grammage) and sales unit.
Our automatic procedure:

1. **Matching products over time** by using all available product codes and their descriptions. In our procedure, we use the product retailer code, the EAN code, and product description including size and sales unit information. Since, as a rule, reducing the size of the product does not change the product code, but obviously affects the change of its description, it is important that the product matching is based on the text distance with an appropriately selected threshold. We used the **Jaro-Winkler text distance** measure.
2. **Extraction of size and sales unit** information using regular expressions (**regex**) based on the product description (unless the retail chain provides such information in separate columns).
3. **Price and quantity normalization** so that they relate to the unit of sale.
4. **Detection** of codes for matched products that, in the face of a **decrease in size**, have recorded an **increase in unit price**.

An example that is reproducible in the *PriceIndices* package

- **dataDOWNSIZED** - a sample coffee products collection included in the *PriceIndices* package (artificial).
- The automatic procedure of downsized product detection in the *PriceIndices* package:

1. **Data matching** over time:
   df <- **data_matching**(data = dataDOWNSIZED,
   start="2024-01", end="2024-02",
   codeIN = TRUE, codeOUT = TRUE, description = TRUE,
   onlydescription = FALSE, precision = 0.9, interval=FALSE)
2. **Extraction** of information about product grammage:
   df <- **data_unit**(df, units = c("g|ml|kg|l"), multiplication = "x")
3. **Price and quantity normalization**:
   df<-**data_norm**(df, rules = list(c("ml","l",1000), c("g","kg",1000)))
4. **Downsized products detection**:
   result <- **shrinkflation**(data = df,
   start = "2024-01",end = "2024-02",
   prec = 3, interval = FALSE)
   result$changes
   result$products_downsized
   result$df_downsized
   result$df_reduced
   result$df_summary

Input data (dataDOWNSIZED)

Original prices and quantities are not standardized. Information about the size and unit of sale is included in the product description.

	time	prices	quantities	codeIN	codeOUT	description
1	2024-01-01	20	100	1	10000000000100	coffee ABC 200g
2	2024-01-02	20	200	1	10000000000100	coffee ABC 200g
3	2024-01-03	30	150	2	10000000000200	coffee DEF 300 g
4	2024-01-04	35	150	2	10000000000200	coffee DEF 300 g
5	2024-01-05	40	300	3	10000000000300	coffee GHI 2 x 400g
6	2024-01-06	42	400	3	10000000000300	coffee GHI 2 x 400g
7	2024-01-07	50	200	4	10000000000400	coffee JKL 250 ml
8	2024-01-08	50	200	4	10000000000400	coffee JKL 250 ml
9	2024-01-09	52	100	4	10000000000400	coffee JKL 250 ml
10	2024-01-10	70	50	5	10000000000500	coffee super 0,5 l
11	2024-01-11	50	100	7	10000000000900	cofee extra
12	2024-02-01	18	100	1	10000000000100	coffee ABC 180g
13	2024-02-02	18	200	1	10000000000100	coffee ABC 170g
14	2024-02-03	31	150	2	10000000000200	coffee DEF 300 g
15	2024-02-04	37	150	2	10000000000200	coffee DEF 300 g
16	2024-02-05	41	300	3	10000000000300	coffee GHI 2 x 390g
17	2024-02-06	42	400	3	10000000000300	coffee GHI 2 x 390g
18	2024-02-07	50	200	4	10000000000400	coffee JKL 250 ml
19	2024-02-08	50	200	4	10000000000400	coffee JKL 240 ml
20	2024-02-09	50	100	4	10000000000400	coffee JKL 240 ml
21	2024-02-10	69	50	5	10000000000500	coffee super 0,4 l
22	2024-02-11	50	120	7	10000000000900	cofee extra

Potential downsized products

Results: *results$changes*

	prodID	grammage	unit	mean_price	size_decrease	price_increase	downsizing	description	time
1	8	0.20	kg	100.000	-	-	FALSE	coffee ABC 200g	2024-01
2	8	0.18	kg	100.000	10 %	0 %	FALSE	coffee ABC 180g	2024-02
3	8	0.17	kg	105.882	5.556 %	5.882 %	TRUE	coffee ABC 170g	2024-02
4	9	0.80	kg	51.429	-	-	FALSE	coffee GHI 2 x 400g	2024-01
5	9	0.78	kg	53.297	2.5 %	3.632 %	TRUE	coffee GHI 2 x 390g	2024-02
6	10	0.25	l	201.143	-	-	FALSE	coffee JKL 250 ml	2024-01 ; 2024-02
7	10	0.24	l	208.333	4 %	3.575 %	TRUE	coffee JKL 240 ml	2024-02
8	11	0.50	l	140.000	-	-	FALSE	coffee super 0,5 l	2024-01
9	11	0.40	l	172.500	20 %	23.214 %	TRUE	coffee super 0,4 l	2024-02

Summary of downsized products: *results$df_summary*

	stats	value
1	Downsized product shares:	----------
2	number of all products	6
3	number of downsized products	4
4	share of downsized products	66.667 %
5	turnover of all products	157400
6	turnover of downsized products	126450
7	turnover share of downsized products	80.337 %
8	Average measures:	----------
9	mean size decreases of downsized products	8.014 %
10	mean price increase of downsized products	9.076 %
11	median size decreases of downsized products	4.778 %
12	median price increase of downsized products	4.757 %
13	Volatility measures:	----------
14	standard deviation of size decreases	8.087 %
15	standard deviation of price increases	9.487 %
16	volatility coefficient of size decreases	1.009
17	volatility coefficient of price increases	1.045

Empirical study

- Data from a big retail chain in Poland (over 510 outlets) concerning the following elementary product groups: a) food category: **yoghurt** (COICOP 5: 011441), **rice** (COICOP 5: 011111), **groats** (COICOP 5: 011123), **baked goods** (COICOP 5: 011131), **coffee** (COICOP 5: 012111);
  b) non-food category: **cosmetics and hygiene products** (COICOP 5: 121321)
- Time interval: December, 2020 - February, 2024
- Product groups where downsizing was detected: **yoghurt**, **cosmetics and hygiene products**

Sample TRUE detections of downsized products

	product_group	description_before_change	description_after_change
1	yoghurt	7 ZBOZ TRUSKAWKA 150G	7 ZBOZ TRUSKAWKA 140G
2	yoghurt	JOGURT TWIST 400G OWOCE LESNE	JOGURT TWIST 380G OWOCE LESNE
3	yoghurt	JOGURT NATURALNY 2% 175G LOKAL	JOGURT NATURALNY 2% 170G
4	cosmetics and hygiene products	czarne mydło 37 polnych ziół 340ml receptury zielarki	czarne mydło 37 polnych ziół 300ml receptury zielarki
5	cosmetics and hygiene products	szampon d/włosów propolisowy 400ml receptury zielarki	szampon d/włosów propolisowy 350ml receptury zielarki
6	cosmetics and hygiene products	żel p/prysz. drzewo gwajakowe 400ml yope	żel p/prysz. drzewo gwajakowe 300ml yope

Sample FALSE detections of downsized products

	product_group	description_before_change	description_after_change
1	rice	ryz brazowy 500kg britta	ryz britta brazowy 0,5kg
2	groats	kasza gryczana 4x100g	kasza gryczana 4*100g

Empirical results

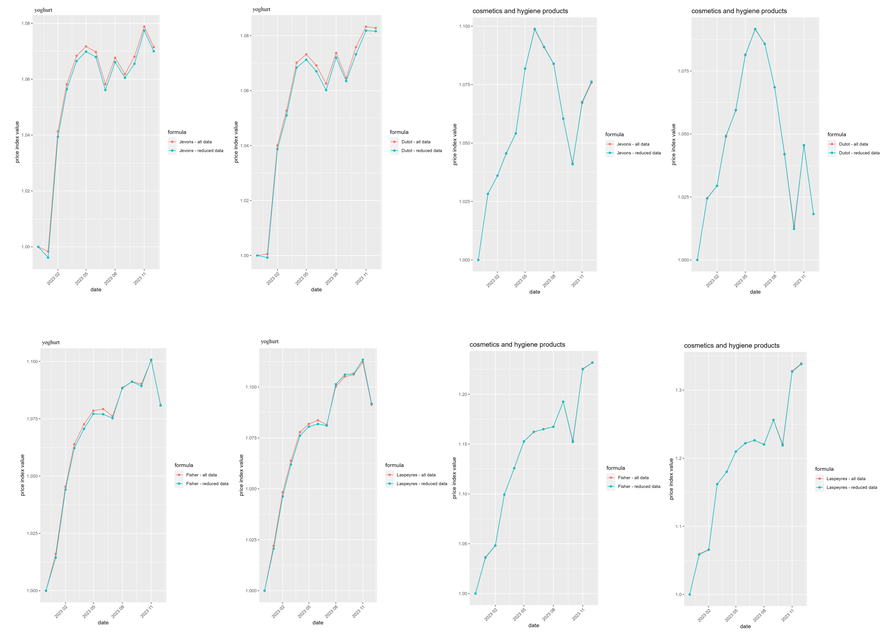
yoghurt

	stats	value
1	Downsized product shares:	----------
2	number of all products	507
3	number of downsized products	16
4	share of downsized products	3.16 %
5	turnover of all products	332057271.77
6	turnover of downsized products	22483853.51
7	turnover share of downsized products	6.77 %
8	Average measures:	----------
9	mean size decreases of downsized products	4.69 %
10	mean price increase of downsized products	26.77 %
11	median size decreases of downsized products	5 %
12	median price increase of downsized products	26.92 %
13	Volatility measures:	----------
14	standard deviation of size decreases	1.68 %
15	standard deviation of price increases	15.28 %
16	volatility coefficient of size decreases	0.36
17	volatility coefficient of price increases	0.57

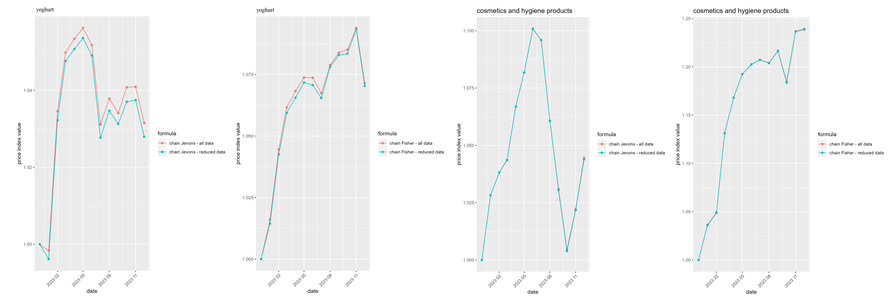
cosmetics and hygiene products

	stats	value
1	Downsized product shares:	----------
2	number of all products	3730
3	number of downsized products	17
4	share of downsized products	0.46 %
5	turnover of all products	560558784.56
6	turnover of downsized products	1485508.39
7	turnover share of downsized products	0.27 %
8	Average measures:	----------
9	mean size decreases of downsized products	20.06 %
10	mean price increase of downsized products	28.58 %
11	median size decreases of downsized products	12.5 %
12	median price increase of downsized products	18.41 %
13	Volatility measures:	----------
14	standard deviation of size decreases	15.99 %
15	standard deviation of price increases	34.47 %
16	volatility coefficient of size decreases	0.8
17	volatility coefficient of price increases	1.21

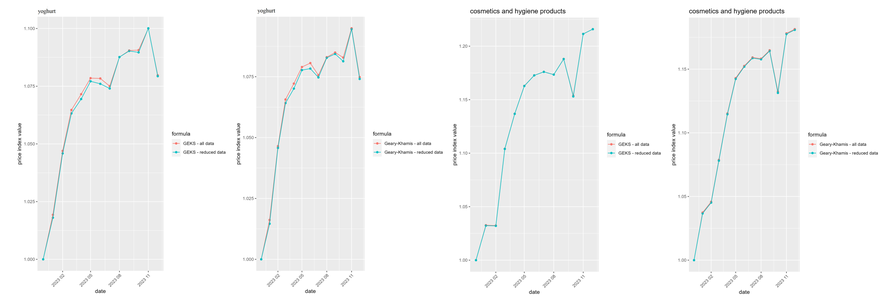
Bilateral price index effect



Chain price index effect



Multilateral price index effect



Conclusions

- The automatic downsizing detection procedure can be effective as long as correct product matching based on both product codes and an appropriate measure of text distance is carried out. The procedure also requires normalization of prices, so information on the product's weight and selling unit is needed.
- When the share of sales of downsized products is small (less than one percent) then their impact on the value of the price index will be marginal. This observation holds true in the context of cosmetics and hygiene products.
- In the case of a few percent (or more) of sales of downsized products, we may see a measurable overestimation of the price index. We observed this phenomenon in the case of **yoghurt**, where this overestimation could exceed 0.35 p.p. (and please note that this is only one elementary group). Chain indices seem to be the most sensitive to downsized products, while unweighted bilateral formulas seem to be the least sensitive.

References

[1] Jacek Białek.
PriceIndices – a new R package for bilateral and multilateral price index calculations.
Statistika – Statistics and Economy Journal, 36(2):122–141, 2021.

[2] Jacek Białek.
Scanner data processing in a newest version of the PriceIndices package.
Statistical Journal of the IAOS, 4(38):1369–1397, 2022.

[3] Eurostat.
Guide on Multilateral Methods in the Harmonised Index of Consumer Prices.
Luxembourg: Publications Office of the European Union, 2022.

[4] Dario Giner.
Hi shrinkflation, meet your cousin shitflation.
THE BLIND SPOT, *https://the-blindspot.com/hi-shrinkflation-meet-your-cousin-shitflation/*, 2022.

[5] International Monetary Fund | International Labour Organization Statistical Office of the European Union (Eurostat) United Nations Economic Commission for Europe Organisation for Economic Co-operation and Development | The World Bank.
Consumer Price Index manual: Concepts and methods.
Washington, D.C, 2020.

J.Bialek@stat.gov.pl, A.Bobel@stat.gov.pl, D.Oprych-frankow@stat.gov.pl