EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS

2024 ESTORIL - PORTUGAL

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

IMPACT OF THE PARTIAL TIME COVERAGE OF RETAIL CHAIN DATA ON THE ACCURACY OF THE PRICE INDEX CALCULATION

Petra Mazureková
Helena Glaser-Opitzová

STATISTICAL OFFICE OF THE SLOVAK REPUBLIC

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Agenda:

- Introduction
- Current status
- Theoretical framework
- Results
- Conclusion

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly
financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Introduction

The aim

- To compare price indices calculated from 2, 3, or 4 weeks (full coverage of reference period/month)
- The best solution for the trade-off between accuracy and timeliness through partial time coverage

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Introduction

Scanner data

- defined as detailed data of sales of consumer goods obtained by scanning the bar codes for individual products at electronic points of sale in retail outlets
- detailed information about quantity sold and turnover of individual product items
- take account of the real impact of discounts

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

STATE OF PLAY

STATISTICAL OFFICE OF THE SLOVAK REPUBLIC

- from 01/2024 – scanner data for food and non-alcoholic beverages in the real production of price statistics
- data from 5 retail chains on a weekly basis
- weekly processing, validation of received data
- classification of products to ECOICOP6 (national lower level of ECOICOP classification) – 350 homogeneous groups
- dynamic approach with the Jevons formula

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

The conference is partly financed by the European Union

Theoretical framework:

Elementary price indices- bilateral:

JEVONS INDEX

$$I_{Jevons}^{m-1,m} = \prod_{i \in N} \left(\frac{p_i^m}{p_i^{m-1}} \right)^{\frac{1}{N}}$$

TÖRNQVIST INDEX

$$I_{Tornqvist}^{m-1,m} = \prod_{i \in N} \left(\frac{p_i^m}{p_i^{m-1}} \right)^{\frac{s_i^{m-1} + s_i^m}{2}}$$

- m is the current period
- m-1 is the previous period
- p_i is the price of the i-th product item
- q_i is the quantities sold of the i-th product item
- $s_i^{m-1} = p_i^{m-1} q_i^{m-1} / \sum_{i \in N} p_i^{m-1} q_i^{m-1}$ a $s_i^m = p_i^m q_i^m / \sum_{i \in N} p_i^m q_i^m$ are the shares of expenditures in given month

The conference is partly financed by the European Union

Theoretical framework:

Elementary price indices-multilateral:

GEKS-TÖRNQVIST INDEX

$$I^{0,m}_{GEKS-Törnqvist} = \prod_{l \in T} \left(I^{0,l}_{\text{Törnqvis}} \times I^{l,m}_{\text{Törnqvist}} \right)^{\frac{1}{T}}$$

- where T is the window length

  (T=13 months and calculation interval =25 months)
- bilateral Törnqvist indices are used for the compilation
- In multilateral methods, the aggregate price change between two compared periods is obtained from prices and quantities observed in multiple periods.

The conference is partly financed by the European Union

Theoretical framework:

Elementary price indices-multilateral:

- A disadvantage of multilateral methods is that they suffer from revision, i.e. each time a new index is calculated for period T+1 and so on, all previous indices are recalculated and changed for the defined time window.
- A splicing method must be used that links the latest index to previous results to avoid revisions of already published results

SPLICING METHOD- HASP

$$I_{pub}^{0,t} = I_{pub}^{0,t-1} \times \prod_{k=t-T+1}^{t-1} \left(I_{pub}^{t-1,k} \times I_{[t-T+1,t]}^{k,t} \right)^{\frac{1}{T-1}} = \prod_{k=t-T+1}^{t-1} \left(I_{pub}^{0,k} \times I_{[t-T+1,t]}^{k,t} \right)^{\frac{1}{T-1}}$$

The conference is partly financed by the European Union

Results

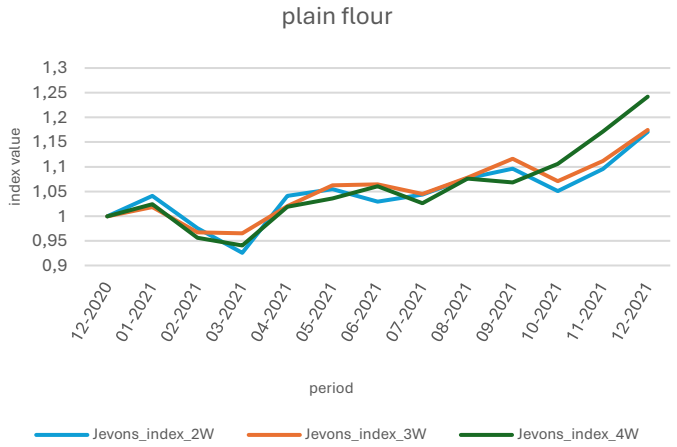
The analysis of average unit prices and time series of price indices

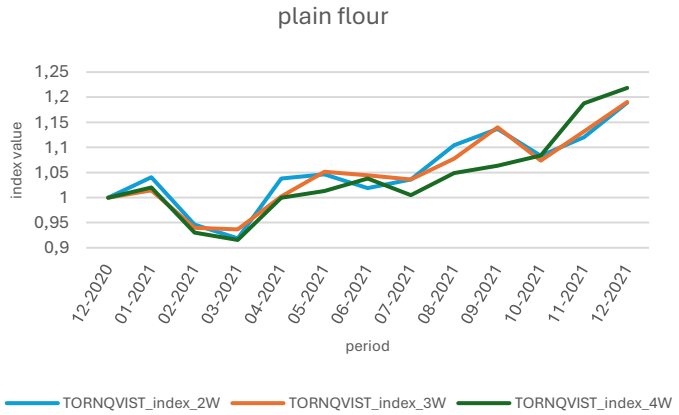
Average unit prices of goods for different time coverage in selected periods

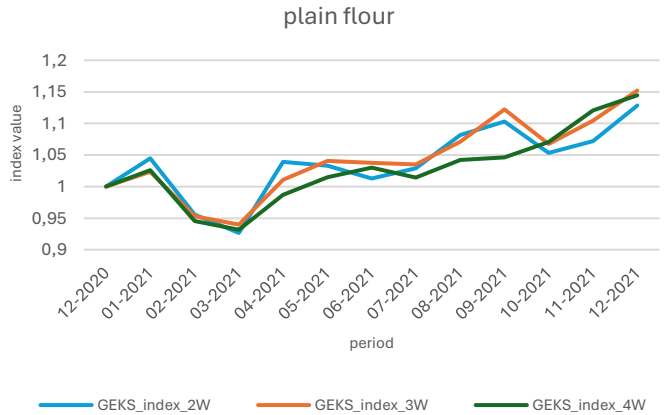
GTIN/EAN	Item description	Period	Average price in €		
			2W	3W	4W
4008671013004	Crystal sugar 1kg	04	0,9395	0,9147	0,9099
4008671013004	Crystal sugar 1kg	05	0,9150	0,9129	0,9137
4008671013004	Crystal sugar 1kg	06	0,8386	0,8902	0,9179
4008671013004	Crystal sugar 1kg	07	1,0497	1,0357	0,9961
00401301	RAJO UHT semi-skimmed milk 1,5 % 1 L	01	0,9446	0,9374	0,9390
00401301	RAJO UHT semi-skimmed milk 1,5 % 1 L	02	0,9476	0,9359	0,9387
00401301	RAJO UHT semi-skimmed milk 1,5 % 1 L	03	0,9236	0,9415	0,9305
00401301	RAJO UHT semi-skimmed milk 1,5 % 1 L	04	0,9868	1,0061	1,0169
00401301	RAJO UHT semi-skimmed milk 1,5 % 1 L	10	1,4422	0,9475	0,9553
00401301	RAJO UHT semi-skimmed milk 1,5 % 1 L	11	1,4762	1,4980	1,4832
00401301	RAJO UHT semi-skimmed milk 1,5 % 1 L	12	1,5491	1,2081	1,2057
2002006552802	Rapeseed oil RACIOL 2l	08	6,8671	6,7935	6,7628
2002006552802	Rapeseed oil RACIOL 2l	09	6,7874	5,8253	5,7718
2002006552802	Rapeseed oil RACIOL 2l	10	6,2477	6,3722	6,5170
2002006552802	Rapeseed oil RACIOL 2l	11	7,2538	7,2438	7,2702

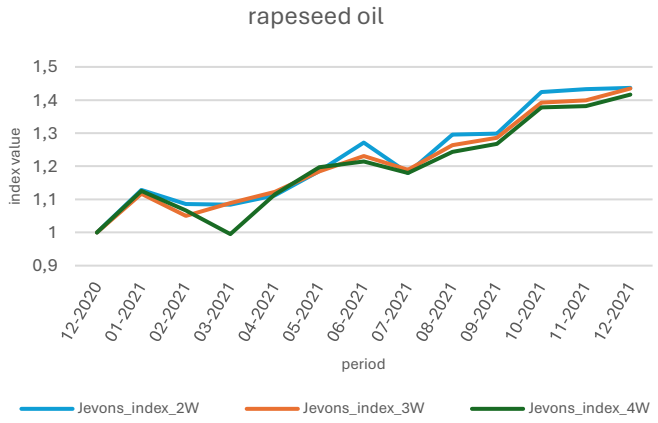
INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

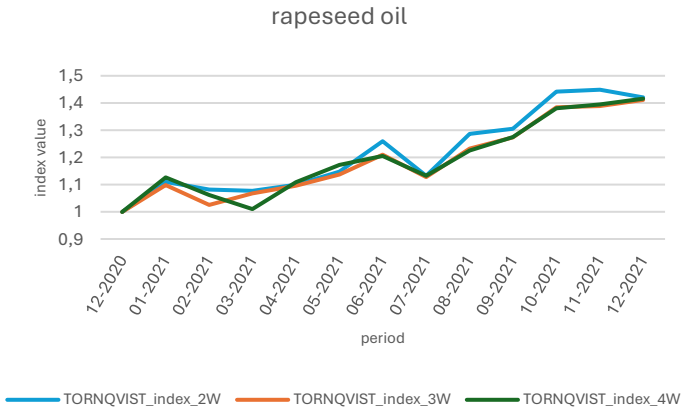
Comparison of the trend of monthly price indices on homogeneous product groups-ECOICOP6, partial coverage of 2W and 3W vs. full time coverage of 4W

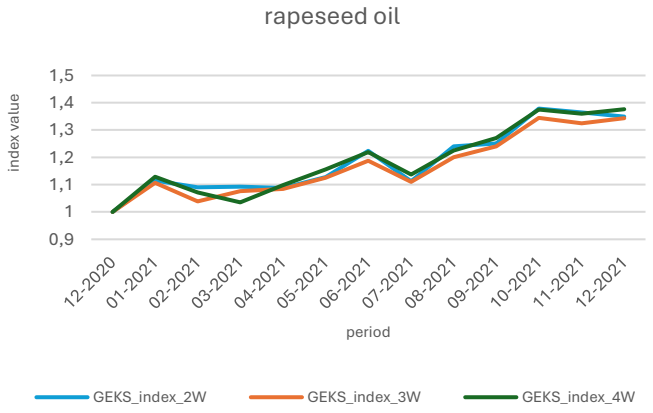






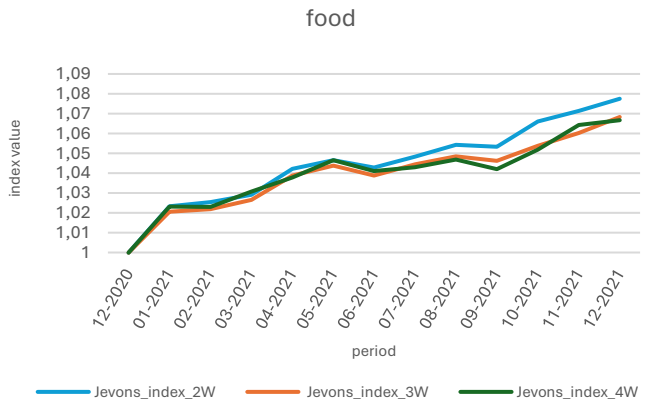


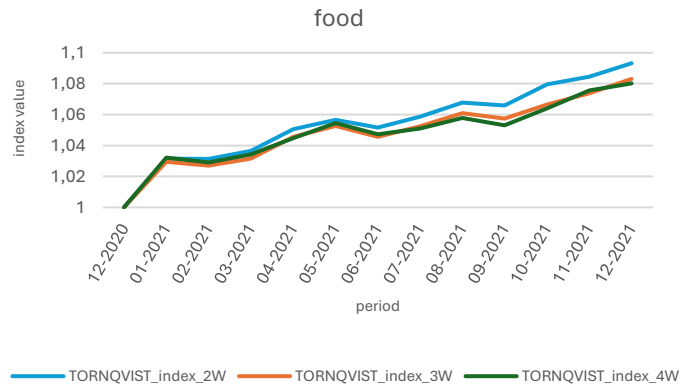


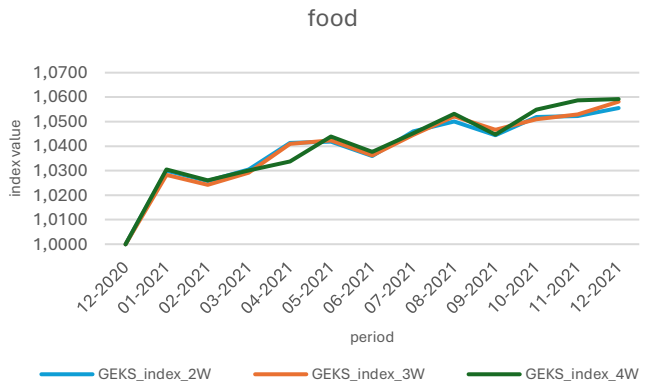


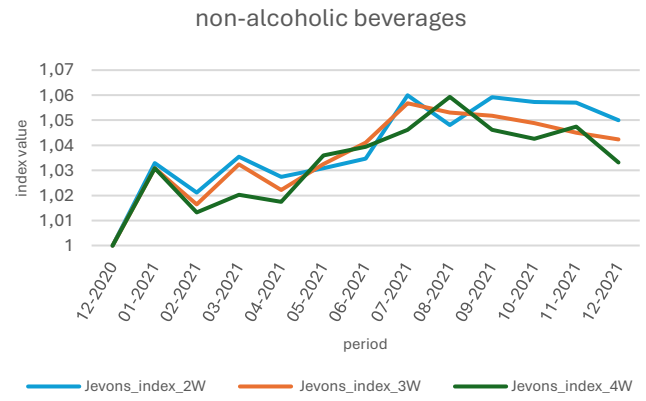
INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

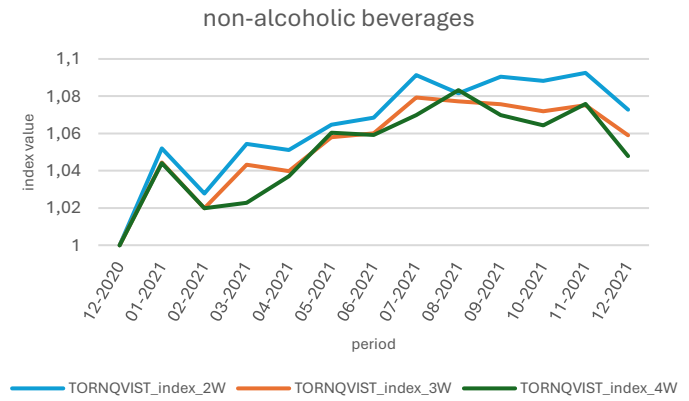
Comparison of the trend of monthly price indices on the product group – 01.1 Food and 01.2 Non-alcoholic beverages, partial 2W and 3W vs. full time coverage 4W

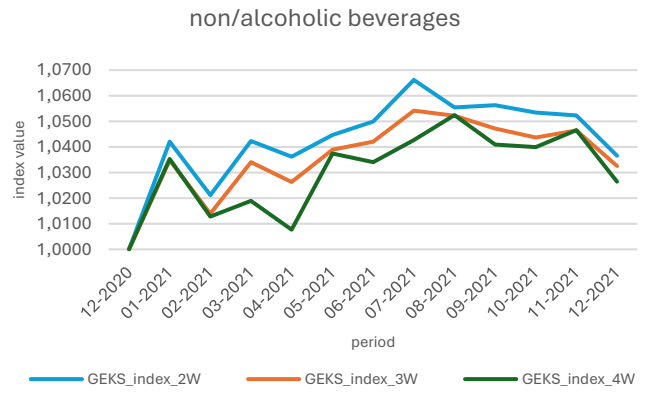












The conference is partly financed by the European Union

Mean absolute percentage error

$$MAPE = \frac{1}{12}\sum_{i=1}^{12}\frac{\left|I_i^{4w} - I_i^{*w}\right|}{I_i^{4w}} * 100$$

where I_i^{4W} is the index calculated from full month data in the i-th period and I_i^{*W} is the index calculated from data for either 2 or 3 weeks of the i-th period

	Mean Absolute Percentage Error (MAPE) caused by partial coverage					
	Jevons 2W/4W	Törnqvist 2W/4W	GEKS 2W/4W	Jevons 3W/4W	Törnqvist 3W/4W	GEKS 3W/4W
01.1 Food	0,527%	0,672%	0,238%	0,230%	0,219%	0,236%
01.2 Non-alcoholic beverages	0,998%	1,462%	1,304%	0,525%	0,534%	0,590%

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Conclusion

- the partial time coverage has not negligible impact on the value of the consumer price indices
- MAPE was reduced by more than half when partial time coverage changed from 2 to 3 weeks

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Thank you for your attention

Contact details:
Petra Mazureková
Statistical Office of the Slovak Republic,
General methodology, registers and coordination of national statistical system Directorate
Statistical Surveys and Methodology Department
Lamačská3/C, 84005 Bratislava 45, Slovak Republic
petra.mazurekova@statistics.sk

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union