Treatment of Scanner (Transaction) Data in the European Comparison Programme (ECP)

International Comparison Program (ICP)
Technical Advisory Group (TAG)
March 8 2025

SESSION V: Current Research - EUROSTAT

Introduction

- Scanner data becomes relatively important data source for consumer price statistics (CPIs)
  - 16 Member States (MSs) and 2 EFTA countries use transaction data
  - 4 MSs test and explore the use of TD
- Scanner data is generated by point-of-sales terminals in shops and provides information at the level of the barcode or, more correctly, Global Trade Item Number (GTIN)
- Eurostat supports the modernisation of price statistics
  - PPP statistics – two task forces in the last 7 years
  - financial assistance that Eurostat provides to Member States and EFTA countries
    - Recently DE, DK, HR, LU, NL, FI, IS and NO have been awarded grants to work on this topic

Guiding principles

- Transparency:
  - Openness is needed about the methods of collection and compilation of the input data
    - PPP regulation 1445/2007 Annex 1- paragraph 5.1: Minimum standards for basic information
  - particularly important in the validation process
    - it must be possible to verify whether products for which prices are reported satisfy the agreed item description
    - PPP regulation 1445/2007 Annex 1- paragraph 5.2: Minimum standards for the validation of price survey results
- The use of transaction data should not bring any major changes in the data to be reported and/or the details provided by the countries using it

Guidelines for price reporting

- Selection of GTINs at item level
- Reporting of data
- Weighting of prices
- Using transaction data (TD) in combination with traditional price collection (TPC)
- National average prices vs capital city prices
- Representativity allocation in transaction data

Selection of GTINs at item level

What criteria for selecting GTIN to PPP item?

- Select GTINs that fit the PPP item specification
- Select the most important GTINs per item (according to turnover)
- Follow current rules as in traditional price collection

Reporting of data

Country will report :

- Option A: Average price for each GTIN per shop chain per item (incl. weights within PPP item)
  - this is preferred option
- Option B: Weighted average price for each PPP item

Price collection period should be aligned with the HICP collection period within the month

Reporting of data – example option A

VAT (%) 21 | Is it representative? ○ Yes ○ No | Average Prices? ○ Yes ◉ No | Transaction Data? ◉ Yes ○ No

+ | - | Clear All | Eliminate/Restore | Eliminate All | Finalise

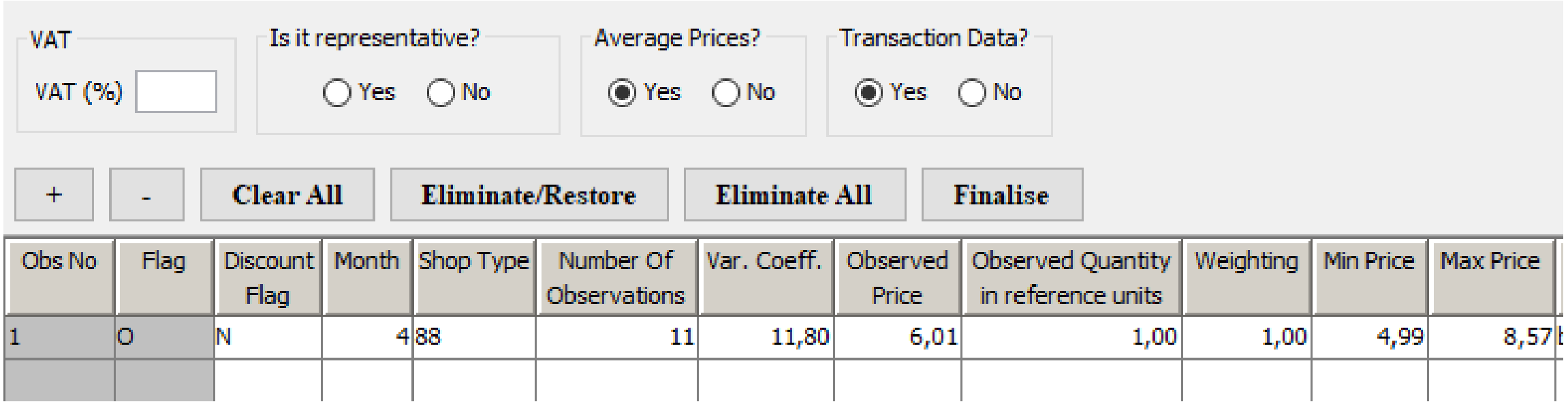
Obs No	Flag	Discount Flag	Month	Shop Type	Shop Identifier	Observed Price	Observed Quantity in reference units	SOB	Weighting	GTIN Description	Comments	Price Collector	Other	Price	Price Ratio
1	O	N	4	88	AA	8,57	1,00	N	2578,00	bertolli olive oil ev				8,57	1,43
2	O	N	4	88	AA	3,31	0,50	Y	11167,00	SOB olive oil ev				6,62	1,10
3	O	N	4	88	AA	6,25	1,00	Y	14599,00	SOB olive oil ev				6,25	1,04
4	O	N	4	88	AA	6,35	1,00	Y	29530,00	SOB olive oil traditional				6,35	1,06
5	O	N	4	88	AA Express	6,25	1,00	Y	3883,00	SOB olive oil ev				6,25	1,04
6	O	N	4	88	AA Express	6,35	1,00	Y	8861,00	SOB olive oil traditional				6,35	1,06
7	O	N	4	88	BudgetMark...	5,15	0,75	N	1978,50	CARB olive oil E.V. 750ML				6,87	1,14
8	O	N	4	88	BudgetMark...	4,99	1,00	Y	14395,00	SOB olive oil EV 1L				4,99	0,83
9	O	N	4	88	CCC	3,15	0,50	Y	4380,00	SOB olive oil EV 500ML				6,30	1,05
10	O	N	4	88	CCC	5,19	1,00	Y	25675,00	SOB olive oil CLASS 1L				5,19	0,86
11	O	N	4	88	CCC	3,21	0,50	Y	5013,50	SOB olive oil MILD 500				6,42	1,07

Reporting of data – example option A (cont.)

Item Statistics	
Average Price:	6,01
Unweighted Average Price:	6,38
Min Price:	4,99
Max Price:	8,57
Var. Coef.:	11,80
Net Price:	4,97
No Of Obs. (N,Q,T):	11
No Of Eliminated Obs.:	0
No Of Obs. with discounts:	0
No Of Obs. with SOB:	9
Add. Check:	

- At PPP item level the following item statistics will automatically be derived in the DET:
- * Weighted average price
- * Unweighted average price
- * Minimum price
- * Maximum price
- * Variation coefficient
- * Number of observations (count of reported GTINs per shop chain)

Reporting of data – example option B



Obs No	Flag	Discount Flag	Month	Shop Type	Number Of Observations	Var. Coeff.	Observed Price	Observed Quantity in reference units	Weighting	Min Price	Max Price
1	O	N	4	88	11	11,80	6,01	1,00	1,00	4,99	8,57

European Commission

Weighting of prices

Two weighting methods:

- Weights are available on GTIN level - dynamic weighting
  - Dynamic weighting is based on quantities
- Weights are available on the level of the shop chain - fixed weighting
  - Fixed weighting - all GTINs have same weight (market share of shop chain)

Weighting of prices – example dynamic weighting

Example for dynamic weighting: 7 GTINs at 4 retailers reported for PPP item

GTIN	SHOP	GTIN quantity	Reference quantity	turnover	units sold	Weight based on Reference unit	average GTIN price (AV)	turnover share	Quantity share based on reference unit	quantity weighted average price (A)
		(1)	(2)	(3)	(4)	(5) =((1)/(2)) * (4)	(6) =((3)/(4)) * ((2)/(1))	(7)	(8)	(9) = (6)*(8)
GTIN A	AA	1.00	1	22093	2578	2578	8.570	0.030	0.021	0.181
GTIN B	AA	0.50	1	73926	22334	11167	6.620	0.101	0.091	0.606
GTIN C	AA	1.00	1	91244	14599	14599	6.250	0.124	0.120	0.748
GTIN D	AA	1.00	1	187516	29530	29530	6.350	0.256	0.242	1.536
GTIN B	AA Express	1.00	1	24269	3883	3883	6.250	0.033	0.032	0.199
GTIN D	AA Express	1.00	1	56267	8861	8861	6.350	0.077	0.073	0.461
GTIN E	BudgetMarket B	0.75	1	13586	2638	1979	6.867	0.019	0.016	0.111
GTIN C	BudgetMarket B	1.00	1	71831	14395	14395	4.990	0.098	0.118	0.588
GTIN B	CCC	0.50	1	27594	8760	4380	6.300	0.038	0.036	0.226
GTIN F	CCC	1.00	1	133253	25675	25675	5.190	0.182	0.210	1.092
GTIN G	CCC	0.50	1	32187	10027	5014	6.420	0.044	0.041	0.264
total				733765	143280	122060		1.000	1.000	6.012

(AV) average GTIN price = total turnover/ total units sold * reference quantity/GTIN quantity

(A) quantitity weighted average price = average GTIN price * Quantity share

Weighting of prices – example fixed weighting

VAT | Is it representative? | Average Prices? | Transaction Data?

VAT (%) | ○ Yes ○ No | ○ Yes ◉ No | ◉ Yes ○ No

+ | - | Clear All | Eliminate/Restore | Eliminate All | Finalise

Obs No	Flag	Discount Flag	Month	Shop Type	Shop Identifier	Observed Price	Observed Quantity in reference units	SOB	Weighting	GTIN Description	Comments	Price Collector	Other	Price	Price Ratio
1	O	N	4	88	Shop A	6,25	1,00	N	36,00	GTIN 1				6,25	1,01
2	O	N	4	88	Shop A	3,15	0,50	N	18,00	GTIN 2				6,30	1,02
3	O	N	4	88	Shop A	4,99	1,00	N	6,00	GTIN 3				4,99	0,80
4	O	N	4	88	Shop B	6,35	1,00	N	12,50	GTIN 1				6,35	1,02
5	O	N	4	88	Shop B	3,31	0,50	N	8,75	GTIN 2				6,62	1,07
6	O	N	4	88	Shop B	5,15	0,75	N	3,75	GTIN 4				6,87	1,11
7	O	N	4	88	Shop C	6,29	1,00	N	10,50	GTIN 1				6,29	1,01
8	O	N	4	88	Shop C	5,15	1,00	N	4,50	GTIN 3				5,15	0,83

Using transaction data in combination with traditional price collection

On the level of item, two data sources for different outlets can be used in calculation of average price.

- In that case, both type of prices (observations) could be reported; dynamic weighting for TD and fixed weighting for TPC
- Try to avoid combination of two data sources on item level

Example combination transaction data (option A) and traditional price collection

+	-	Clear All	Eliminate/Restore	Eliminate All	Finalise

Obs No	Flag	Discount Flag	Month	Shop Type	Shop Identifier	Observed Price	Observed Quantity in reference units	SOB	Weighting	GTIN Description	Comments	Price Collector	Other	Price	Price Ratio
1	O	N	4	88	Shop D	2,49	1,00	N	24,30	GTIN 1				2,49	0,93
2	O	N	4	88	Shop D	2,72	1,00	N	16,20	GTIN 2				2,72	1,01
3	O	N	4	88	Shop D	2,61	1,00	N	13,50	GTIN 3				2,61	0,97
4	O	N	4	2	Shop E	2,65	1,00	N	13,00	Item 1				2,65	0,99
5	O	N	4	2	Shop E	2,95	1,00	N	13,00	Item 2				2,95	1,10
6	O	N	4	2	Shop E	2,70	1,00	N	13,00	Item 3				2,70	1,01
7	O	N	4	6	Bakery 1	2,85	1,00	N	1,17	Item 1				2,85	1,06
8	O	N	4	6	Bakery 2	3,15	1,00	N	1,17	Item 2				3,15	1,17
9	O	N	4	6	Bakery 3	2,95	1,00	N	1,17	Item 3				2,95	1,10
10	O	N	4	6	Bakery 4	2,85	1,00	N	1,17	Item 4				2,85	1,06
11	O	N	4	6	Bakery 5	2,95	1,00	N	1,17	Item 5				2,95	1,10
12	O	N	4	6	Bakery 6	3,05	1,00	N	1,17	Item 6				3,05	1,14

Example combination transaction data (option B) and traditional price collection

Is it representative? ○ Yes ○ No | Average Prices? ◉ Yes ○ No | Transaction Data? ◉ Yes ○ No

Clear All | Eliminate/Restore | Eliminate All | Finalise

ag	Discount Flag	Month	Shop Type	Number Of Observations	Var. Coeff.	Observed Price	Observed Quantity in reference units	Weighting	Min Price	Max Price	GTIN Description	Comments	Price Collector	Other	Price
	N	4	88	3	3,80	2,59	1,00	54,00	2,49	2,72	GTIN 1 GTIN 2 GTIN 3				2,59
	N	4	2	1	0,00	2,65	1,00	13,00			Item 1				2,65
	N	4	2	1	0,00	2,95	1,00	13,00			Item 2				2,95
	N	4	2	1	0,00	2,70	1,00	13,00			Item 3				2,70
	N	4	6	1	0,00	2,85	1,00	1,17			Item 1				2,85
	N	4	6	1	0,00	3,15	1,00	1,17			Item 2				3,15
	N	4	6	1	0,00	2,95	1,00	1,17			Item 3				2,95
	N	4	6	1	0,00	2,85	1,00	1,17			Item 4				2,85
	N	4	6	1	0,00	2,95	1,00	1,17			Item 5				2,95
	N	4	6	1	0,00	3,05	1,00	1,17			Item 6				3,05

National average prices vs capital city prices

Combination of national and capital city prices is possible on level of BH and/or item

- Spatial adjustment factors (SAFs) reported only on BH level
- If SAF <> 1; then reporting combination of both prices can be problematic

Representativity allocation in transaction data

- Turnover share from transaction data can be an important starting point
- Quantities sold can also be useful
- If turnover and quantity do not lead to clear indication of what is representative
  - other information

PPP tools - implementation

The main changes for the tools were:

- Country is able to report weights (e.g. expressed as shares),
- Calculation of a weighted average price and an unweighted average price,
- Possibility to report weights, for TD and for TPC at observation level (multiplication issue),
- Enable to import data in Data Entry Tool (DET) directly from Excel file.

Use of transaction data – Example Food survey

Country	TD	TPC	
AT		460	9
BE		306	128
CH		88	355
CZ		132	331
DE		392	74
DK		419	22
ES		428	463
FI		421	10
IS		308	71
IT		363	107
LU		452	432
NL		372	9
NO		437	
SE		407	40
SI		462	429
SK		439	25

Country	Option A	Option B
AT	X	
BE	X	
CH		X
CZ	X	
DE		X
DK		X
ES		X
FI	X	
IS	X	
IT	X	
LU	X	
NL		X
NO	X	
SE	X	
SI	X	
SK	X	

Impact of use of transaction data

- Not easy to do at the level of PLIs - any idea is welcome
- TD: covers all discounted prices
  - Weighted prices are usually lower than unweighted prices
  - Expected impact on the representativity allocation
- TPC: price collectors visit a store once in the collection period.
- Some countries (CH, DK and FI) tested impact of TD implementation, no quantified impact but discounts have significant impact

Challenges for the use of TD

- Obtaining Transaction data
  - Receive data from suppliers, and timely delivery of data
  - already the default data source in consumer price statistics (CPI/HICP)
  - that Regulation (EC) No 223/2009 has been recently amended to enhance the access of NSIs and Eurostat to privately held data
- Processing Transaction data
  - Initially, the use of TD does not lead to less workload
  - Different suppliers may have different structure in data
  - Linking and identification of GTINs to PPP items
  - Combine TPC/TD, apply fixed and dynamic weights
  - Price comparison between countries with TD and countries with TPC
- Reporting Transaction data
  - Price difference related to outlets, still TPC is needed next to TD
  - No detailed information on discounts
  - No items specifications available

Thank you

European Commission