UPDATES ON WEB SCRAPING OF PRICES FOR THE CONSUMER PRICE INDEX IN THE NATIONAL CAPITAL REGION

Philippine Statistics Authority

29 October 2024

OUTLINE

I. INTRODUCTION

- Collection of data for price surveys has always been done by visiting the physical stores, recording the prices using pen and paper or handheld electronic device.

- In 2019, the Philippine Statistics Authority (PSA) started exploring web scraping as an alternative price collection technique for the generation of the 2012-based Consumer Price Index (CPI). The study was timely as the online stores became more popular to consumers amidst the COVID-19 pandemic.

II. OBJECTIVE

- The purpose of this exercise is to look into the possibility of using prices from online stores as input in the generation of monthly CPI in National Capital Region (NCR).

II. OBJECTIVE

- **Specific Objective 1:** To compare the price movements and trends of official CPI with computed CPI using online prices with respect to the number of online stores. To accomplish this, the following computations were done:

- **1a (All Stores)** – The monthly average price of each commodity is computed using online prices collected from all covered online stores.

- **1b (One Store)** – The monthly average price for each commodity is the price reported by only one online store. The store is selected based on the availability and consistency of price of sample commodity.

II. OBJECTIVE

- **Specific Objective 2:** To compare the price movements and trends of official CPI with computed CPI using online prices with respect to the similarity of commodity specification. To accomplish this, the following computations were done:

- **2a (Exact Specification)** – The subclass CPI is computed using only the online prices of commodities with exact same specifications with that from the CPI market basket regardless of the number of online stores.

- **2b (Equivalent Specification)** – The subclass CPI is computed using online prices of commodities with similar specifications with that from the CPI market basket.

II. OBJECTIVE

- **Specific Objective 3:** To compare the price movements and trends of official CPI with computed CPI using a combination of offline and online prices. To accomplish this, the following computations were done:

- **3a (Hybrid1)** – The CPI is computed using the price collected through survey and web scraped price data. The online prices come from the commodities with exact or similar specifications with that of the CPI market basket, regardless the number of online stores where these were collected.

- **3b (Hybrid2)** - The CPI is computed using the price collected through survey and web scraped price data. The online prices come from the commodities with exact specifications with that of the CPI market basket, regardless the number of online stores.

III. METHODOLOGY

Geographic Coverage	NCR
Frequency and Schedule of Price Collection	Daily (except in weekends)
ample Markets	13 online stores in NCR (Abensons, Ace Hardware, Ansons, Lazada, Merry Mart, National Bookstore, PushKart, Shopee, Watsons, Western Appliance, Wilcon, Zagana, Zalora)
Number of URLs	1,354
Period Covered	February 2020 to December 2021

III. METHODOLOGY

SAMPLE MARKETS:

III. METHODOLOGY

WEB SCRAPING APPLICATION

The web scraping application was developed in-house using Python and Beautiful Soup. The user-interface of the application utilizes Mozilla Firefox, where the launcher and uniform resource locator (URL) of the target commodities are saved.

III. METHODOLOGY

WEB SCRAPING APPLICATION: FOLDERS

III. METHODOLOGY

WEB SCRAPING APPLICATION:
HTML STRUCTURE

III. METHODOLOGY

CPI WebScraper   Home   Single Product   Multiple Products

Scraping Complete

Type	Action	Message
Single	Initialized	abensons -- Scraping Initialized
Single	Completed	abensons -- Scraping Complete

Return to Homepage

WEB SCRAPING APPLICATION: COMPLETION PROMPT

III. METHODOLOGY

WEB SCRAPING APPLICATION: SAMPLE OUTPUT

	A	B	C	D	E	F	G
1	Url	Description	Sub Details	Price	Sale Price		
2	https://wv	APPLE IPAD MINI 5 WI-FI	Item is discontinued.		23,990		
3	https://wv	CONDURA CTD700MNI	Item is discontinued.	19,997			
4	https://wv	DOWELL DI 583NS	SKU 161693	798			
5	https://wv	SHARP ES-W600	SKU 112944	3,997			
6	https://wv	PANASONIC F-40DYP	SKU 56136		1,748		
7	https://wv	HANABISHI GS 600	SKU 3776		648		
8	https://wv	HANABISHI HI-89	SKU 96012		698		
9	https://wv	PANASONIC NA-S6518BSP	SKU 161243	4,799			
10	https://wv	ASAHI RB-6004	SKU 118847		2,098		
11	https://wv	STANDARD SDS 12W	SKU 135746		1,298		
12	https://wv	STANDARD SGS 235S 2B	SKU 136929		1,998		
13	https://wv	SHARP SJ DTH55BS SL	Item is discontinued.	11,697			
14	https://wv	SONY KDL 32R307F	Item is discontinued.	14,499			
15	https://wv	CANON POWERSHOT SX620HS	SKU 144585	15,198			
16	https://wv	TEFAL RK104E	SKU 163548		3,895		
17	https://wv	TEFAL RK7405	SKU 161277		8,995		
18	https://wv	TEFAL RK8145	SKU 161278		10,995		
19	https://wv	LA GERMANIA E-726 W	SKU 170556	6,798			
20	https://wv	TEKNO TKX- 180	SKU 164815	648			
21	https://wv	TEKNO TKX-780	SKU 164814	1,278			
22	https://wv	KELVINATOR WKELH010EA	SKU 147117	18,498			
23							
24							
25							
26							
27							

single-abensons-742022

Ready   Accessibility: Unavailable

III. METHODOLOGY

Table 1. Number of URLs by Store and Commodity Division

Name of Online Stores	No. of URLs	Commodity Division Code									
		01	02	03	04	05	06	07	08	09	11
Total	1,354	402	15	94	38	231	74	5	8	233	254
Abensons	16					13				3	
Ace Hardware	15				4	11					
Ansons	12					11				1	
Lazada	552	155	11	39	14	87	16	2	3	107	118
Merry Mart	3	1	1				1				
National Bookstore	23									23	
PushKart	74	65	1			1					7
Shopee	539	151	2	43	14	84	16	3	5	96	125
Watsons	45						41				4
Western Appliance	17					16				1	
Wilcon	16				6	8				2	
Zalora	12			12							
Zagana	30	30									

Legend:
01 – Food and Non-Alcoholic Beverages
02 – Alcoholic Beverages and Tobacco
03 – Clothing and Footwear
04 – Housing, Water, Electricity, Gas and Other Fuels
05 – Furnishing, Household Equipment and Routine Maintenance of the House
06 – Health
07 – Transport
08 – Communication
09 – Recreation and Culture
11 – Restaurant and Miscellaneous Goods and Services

III. METHODOLOGY

Table 2. Number of Commodities by Division

Division		CPI Market Basket	No. of Commodities Web Scraped	% to Total
01	Food and Non-Alcoholic Beverages	238	183	25
02	Alcoholic Beverages and Tobacco	12	10	1
03	Clothing and Footwear	64	41	6
04	Housing, Water, Electricity, Gas and Other Fuels	45	14	2
05	Furnishing, Household Equipment, and Routine Household Maintenance	90	85	12
06	Health	56	41	6
07	Transport	23	2	0.3
08	Communication	19	2	0.3
09	Recreation and Culture	81	64	9
10	Education	12	-	-
11	Restaurant and Hotels	80	75	10
Total		720	517	71.6

III. METHODOLOGY

Table 3. Count of Subclass and Class with Complete Commodities via Online

Division	No. of subclass			No. of Class		
	Offline	Online	With complete commodities via online	Offline	Online	With complete subclasses via online
01 – Food and Non-Alcoholic Beverages	60	57	41	11	11	3
02 – Alcoholic Beverages and Tobacco	7	7	6	4	4	3
03 – Clothing and Footwear	18	12	10	6	3	1
04 – Housing, Water, Electricity, Gas and Other Fuels	15	3	2	8	1	-
05 – Furnishing, Household Equipment, and Routine Household Maintenance	28	26	24	11	9	7
06 – Health	16	8	6	7	3	2
07 – Transport	15	2	2	10	2	1
08 – Communication	9	1	1	3	1	-
09 – Recreation and Culture	29	21	18	17	11	6
10 – Education	9	-	-	8	-	-
11 – Restaurant and Miscellaneous Goods and Services	16	11	11	7	4	3
Total	222	148	121	92	49	26

IV. RESULTS AND FINDINGS

Objective 1 – Subclass Level

Figure 1. CPI Trends for Leafy Vegetables
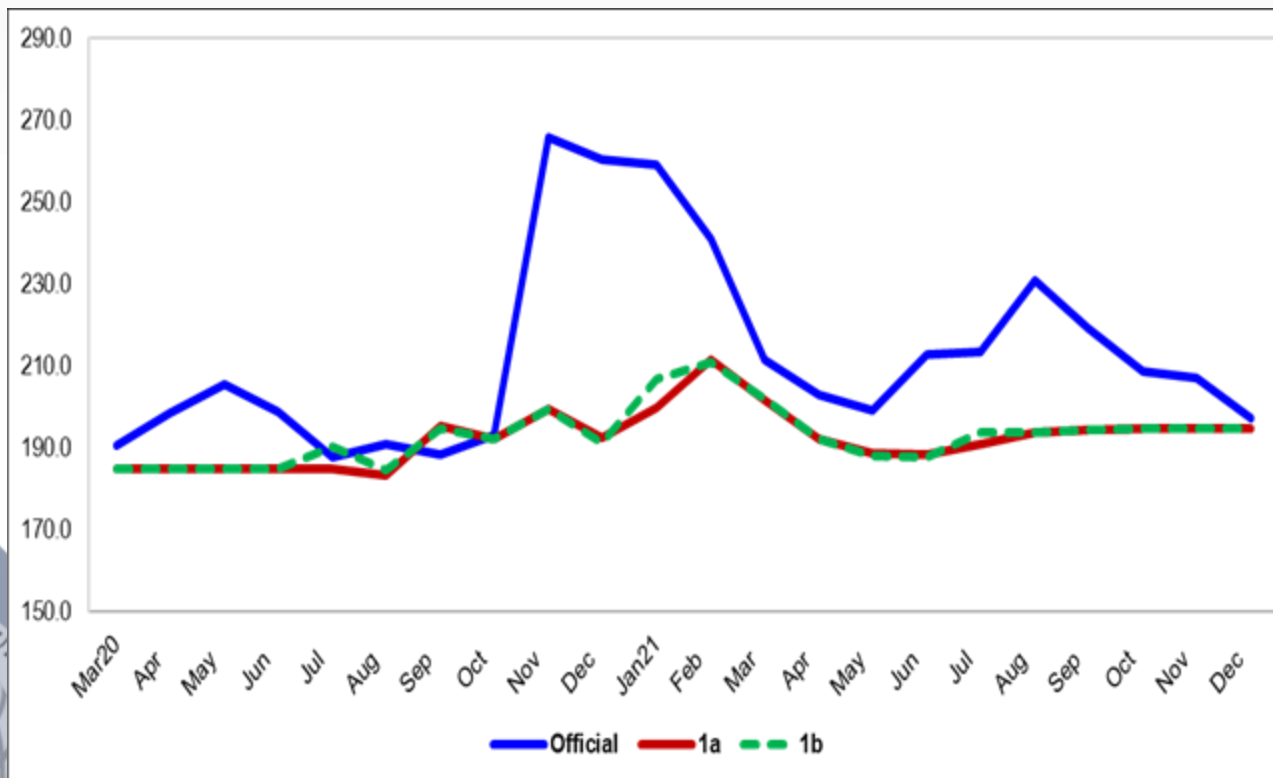
Official: Consists of 11 commodities

1a : Consists of 4 commodities collected from all available stores for each commodity.

1b : Consists of 4 commodities collected from one sample online store for each commodity.

IV. RESULTS AND FINDINGS

Objective 1 – Subclass Level

Figure 2. Month-on-month & Year-on-year Growth Rates of CPI for Leafy Vegetables

IV. RESULTS AND FINDINGS

Objective 1 – Subclass Level

Figure 3. CPI Trends for Cardiovascular Drugs

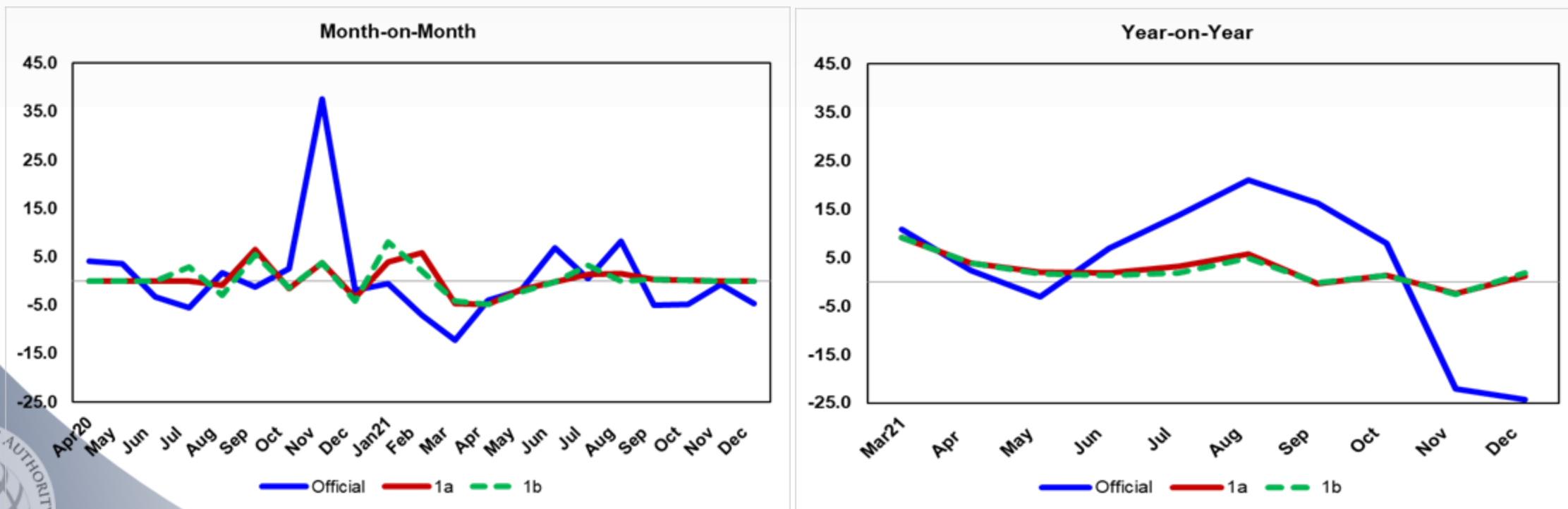**Official:** Consists of 4 commodities

1a : Consists of 3 commodities collected from all available stores for each commodity.

1b : Consists of 3 commodities collected from one sample online store for each commodity.

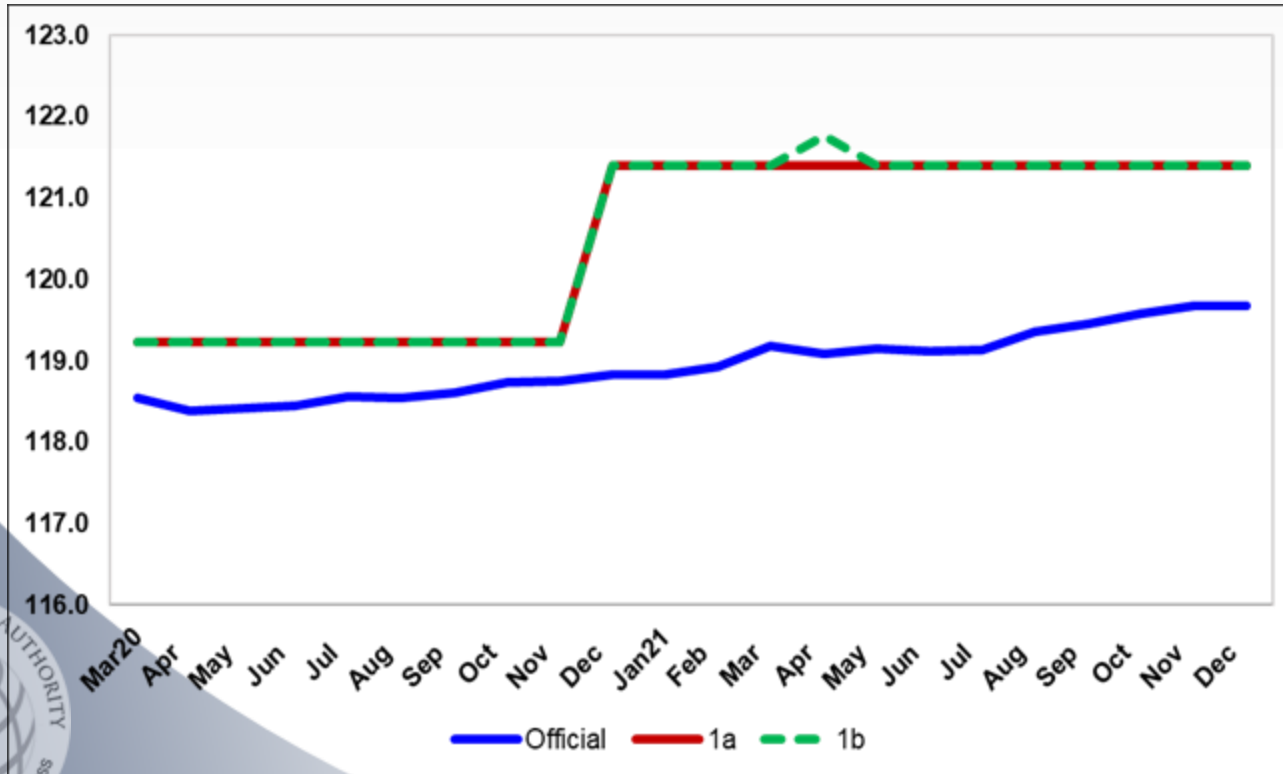## IV. RESULTS AND FINDINGS

Objective 1 – Subclass Level

Figure 4. Month-on-month & Year-on-year Growth Rates of CPI for Cardiovascular Drugs
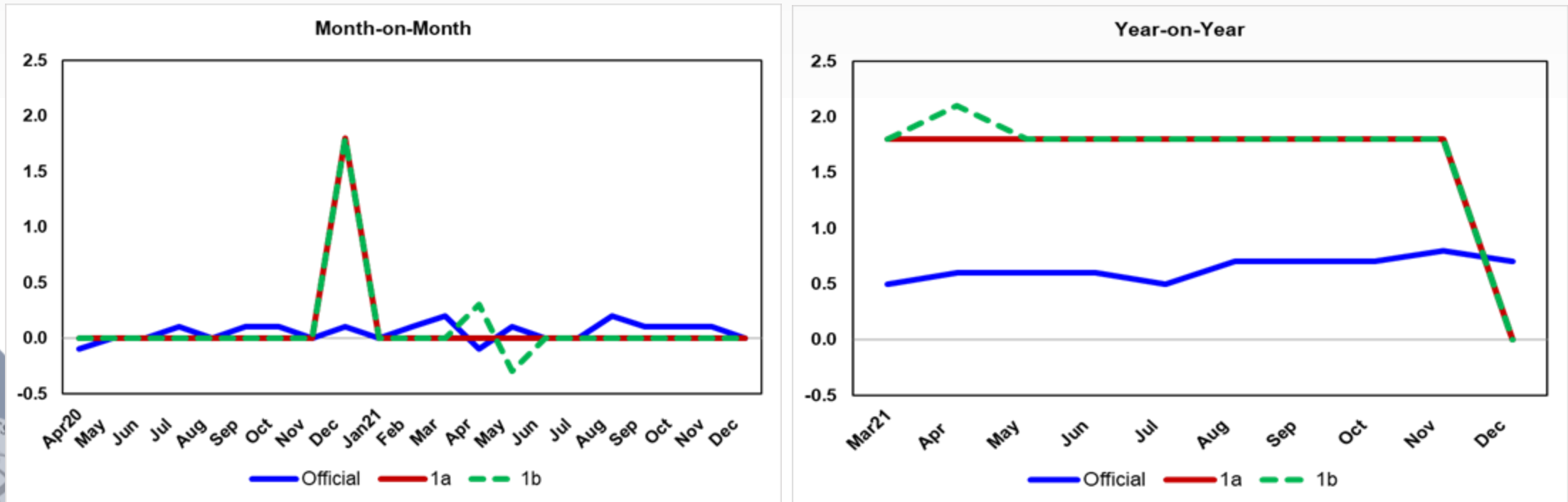
IV. RESULTS AND FINDINGS

Objective 1 – Subclass Level
Table 4. Absolute Deviations at the Subclass Level

Description	Online		Official		Absolute Deviation from the Official CPI				Period with Highest Deviation
	Mar 2020	Dec 2021	Mar 2020	Dec 2021	Lowest	Highest	Range	Average	
1a (All Stores)									
Fresh, chilled, frozen or dried vegetables cultivated for their leaves or stalks	185.0	194.6	190.7	197.3	1.02	68.16	67.14	21.05	Dec. 2020
Cardiovascular drugs	119.2	121.4	118.5	119.7	0.48	2.57	2.09	1.56	Dec. 2020
1b (One Store)									
Fresh, chilled, frozen or dried vegetables cultivated for their leaves or stalks	185.0	194.6	190.7	197.3	1.04	69.42	68.38	20.67	Dec. 2020
Cardiovascular drugs	119.2	121.4	118.5	119.7	0.48	2.69	2.21	1.58	Apr. 2021

IV. RESULTS AND FINDINGS

Objective 1 – Subclass Level

Table 5. Mean Absolute Deviations at the Subclass Level

Mean Absolute Deviation (in Percentage Point)	Number of Subclasses			
	1a (All Stores)		1b (One Store)	
	Count	%	Count	%
0.00 – 0.99	21	14	21	14
1.00 – 1.99	26	18	26	18
2.00 – 2.99	20	13	19	13
3.00 – 3.99	22	15	23	15
4.00 and higher	59	40	59	40
Total	148	100	148	100

IV. RESULTS AND FINDINGS

Objective 1 – Subclass Level

Table 6. Range of Absolute Deviations at the Subclass Level

Range of Absolute Deviations (in Percentage Point)	Number of Subclasses			
	1a (All Stores)		1b (One Store)	
	Count	%	Count	%
0.00 – 0.99	6	4	6	4
1.00 – 1.99	9	6	9	6
2.00 – 2.99	14	9	15	10
3.00 – 3.99	17	12	18	12
4.00 and higher	102	69	100	68
Total	148	100	148	100

IV. RESULTS AND FINDINGS

Objective 1 – Subclass Level

Table 7. Summary of Cointegration Test of CPI of Selected Subclass

2009 PCOICOP	Description	Pair	Trace	Critical Value (α =0.05)	ECT	Conclusion
01.1.71	Fresh, Chilled, Frozen or Dried Vegetables Cultivated for Their Leaves or Stalks	Offline vs 1a	21.40	15.41	-0.3547	Not significant cointegration
		Offline vs 1b	24.24	15.41	-0.3144	Not significant cointegration
06.1.17	Cardiovascular Drugs	Offline vs 1a	14.32	12.53	-0.0134***	Significant cointegration
		Offline vs 1b	13.84	12.53	-0.0112***	Significant cointegration

IV. RESULTS AND FINDINGS

Objective 1 – Subclass Level
Table 8. Summary of Cointegration Test of CPI at the Subclass Level

	Total[/b]	Cointegrated	Not cointegrated
Offline vs 1a (All Stores)	148	32	94
Offline vs 1b (One Store)	148	34	92

[/b] Components do not add up to total because some of the series have no variation, thus, not appropriate for cointegration test.

IV. RESULTS AND FINDINGS

Objective 1 – Subclass Level

Table 9. Summary of Cointegration Test of Subclass CPI by Division

Division	Total Subclass with Online Price	No. of Subclasses with Cointegration	
		1a (All Stores)	1b (One Store)
01 - Food and Non-alcoholic Beverages	57	12	13
02 - Alcoholic Beverages and Tobacco	7	1	2
03 - Clothing and Footwear	12	5	5
04 - Housing, Water, Electricity, Gas and Other Fuels	3	0	0
05 - Furnishings, Household Equipment and Routine Household Maintenance	26	6	7
06 - Health	8	4	4
07 - Transport	2	0	0
08 - Communication	1	0	0
09 - Recreation and Culture	21	3	2
10 - Education	-	-	-
11 - Restaurant and Miscellaneous Goods and Services	11	1	1

IV. RESULTS AND FINDINGS

Objective 1 – Subclass Level

- Results from 1a (all stores) and 1b (one store) showed that there are subclasses from online CPI that have less than one percentage point mean absolute deviations (MAD) from the offline CPI.

- These are notable in food and non-alcoholic beverages (particularly, processed food); clothing and footwear; housing, water, electricity, gas, and other fuels; furnishings, household equipment and routine household maintenance; health; and recreation and culture.

IV. RESULTS AND FINDINGS

Objective 1 – Subclass Level

- Volatile commodities like meat, fruit, and fresh vegetables are more stable in online outlets compared with offline stores. Thus, higher MAD are observed in these commodity groups.

- Cointegration test showed that using 1a (all stores) there are 32 subclasses have significant cointegration with the official CPI and 34 using 1b (one store). Thirty of these subclasses are common to the two explorations.

IV. RESULTS AND FINDINGS

Objective 2 – Subclass Level

Figure 5. CPI Trend for Other Cereal Preparations
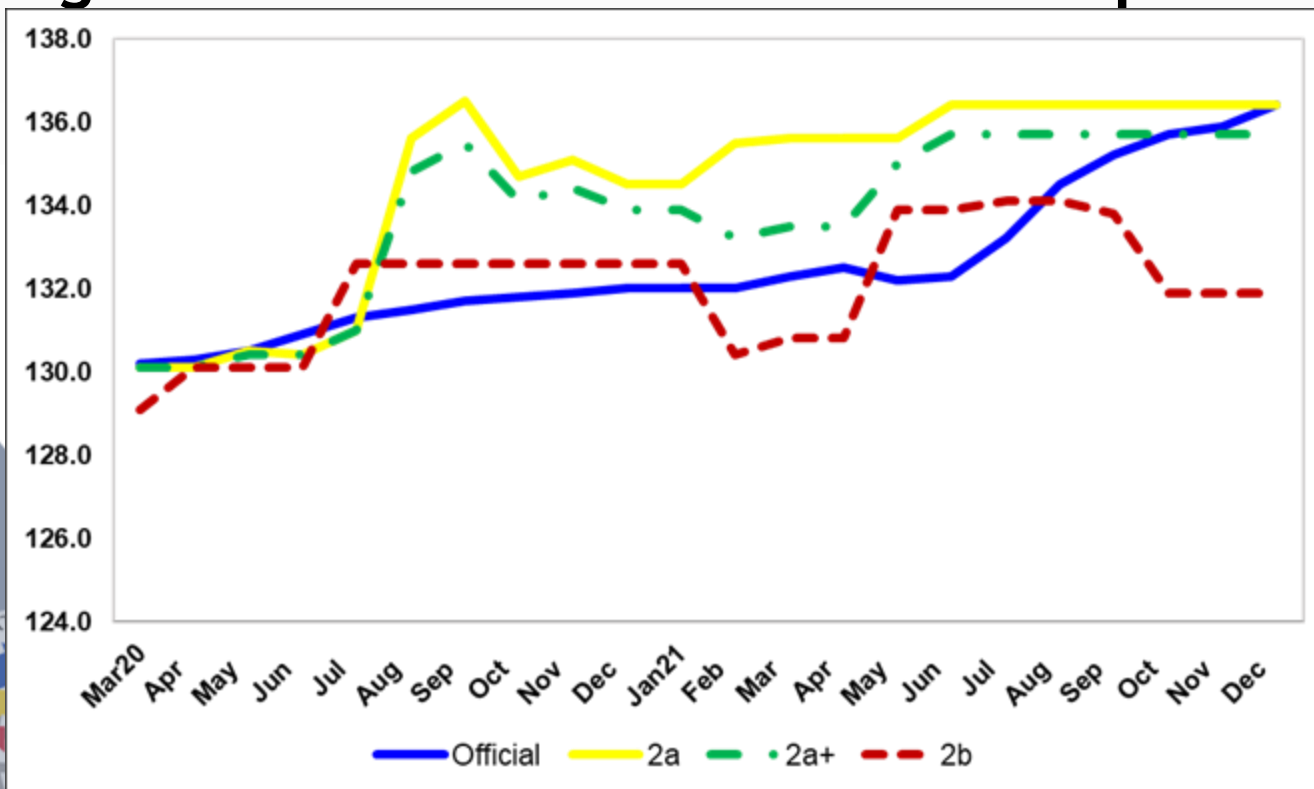
Official: Consists of 11 commodities

2a : Consists of 5 commodities with exact specifications with the official CPI market basket
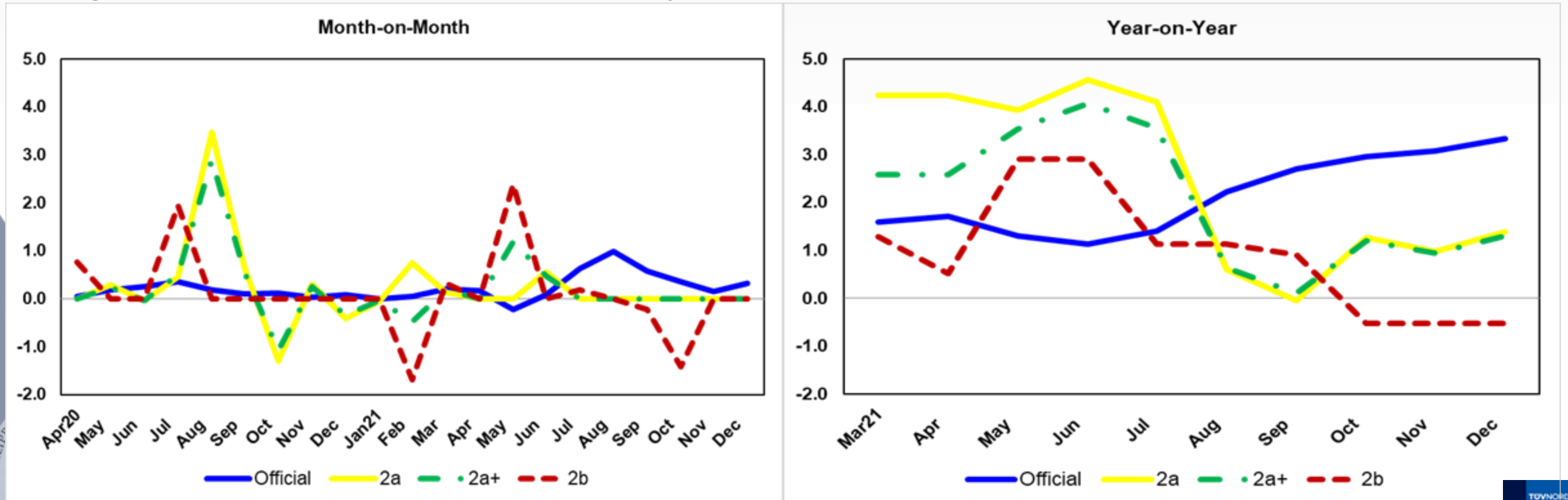
2a+ : Consists of 5 commodities with exact specifications and 1 commodity with similar specification.

2b : Consists of 3 commodities with equivalent specifications with the commodities in the official CPI market basket.

IV. RESULTS AND FINDINGS

Objective 2 – Subclass Level

Figure 6. Month-on-month & Year-on-year Growth Rates of CPI for Other Cereal Preparations

IV. RESULTS AND FINDINGS

Objective 2 – Subclass Level

Figure 7. CPI Trend for Other Pharmaceutical Products

Official: Consists of 17 commodities

2a : Consists of 14 commodities with exact specifications with the official CPI market basket

2a+ : Consists of 2 commodities with similar specifications with the commodities in the official CPI market basket

2b : Consists of 3 commodities with equivalent specifications with the commodities in the official CPI market basket.

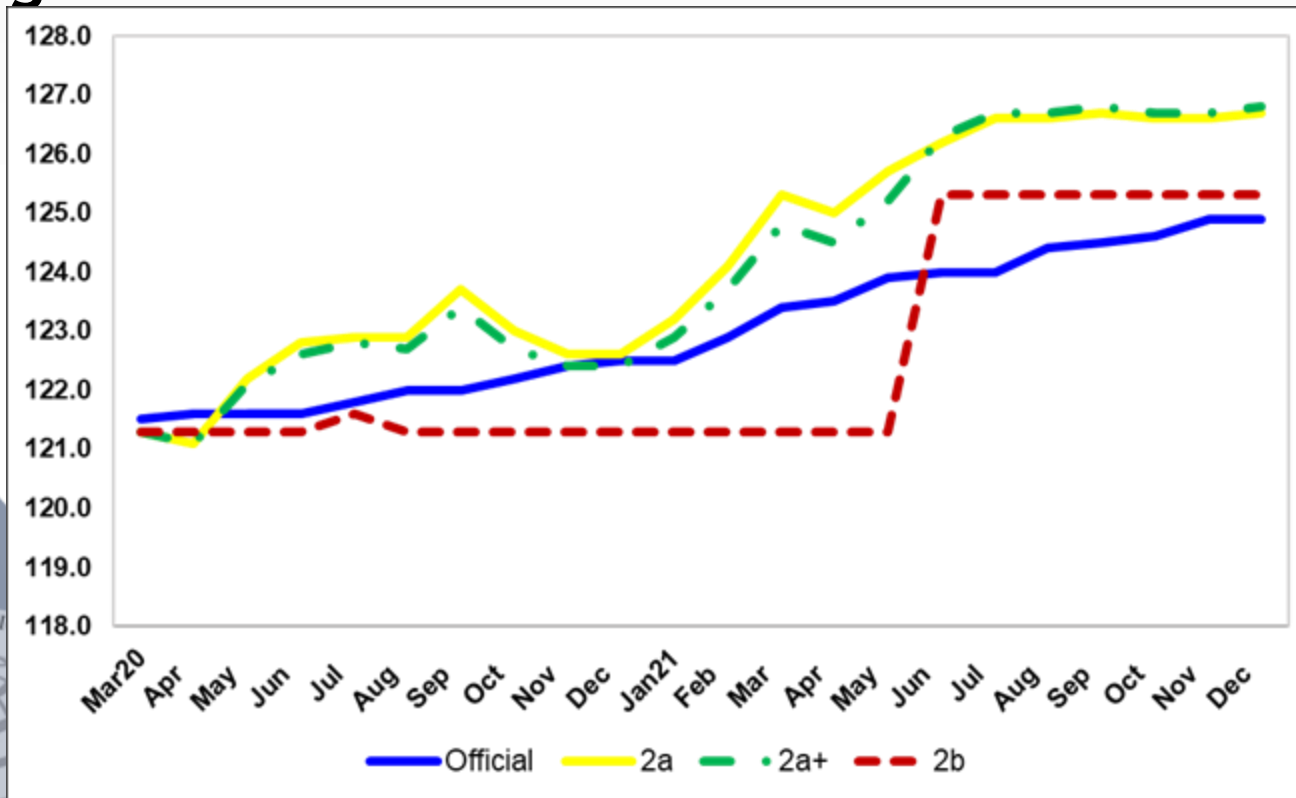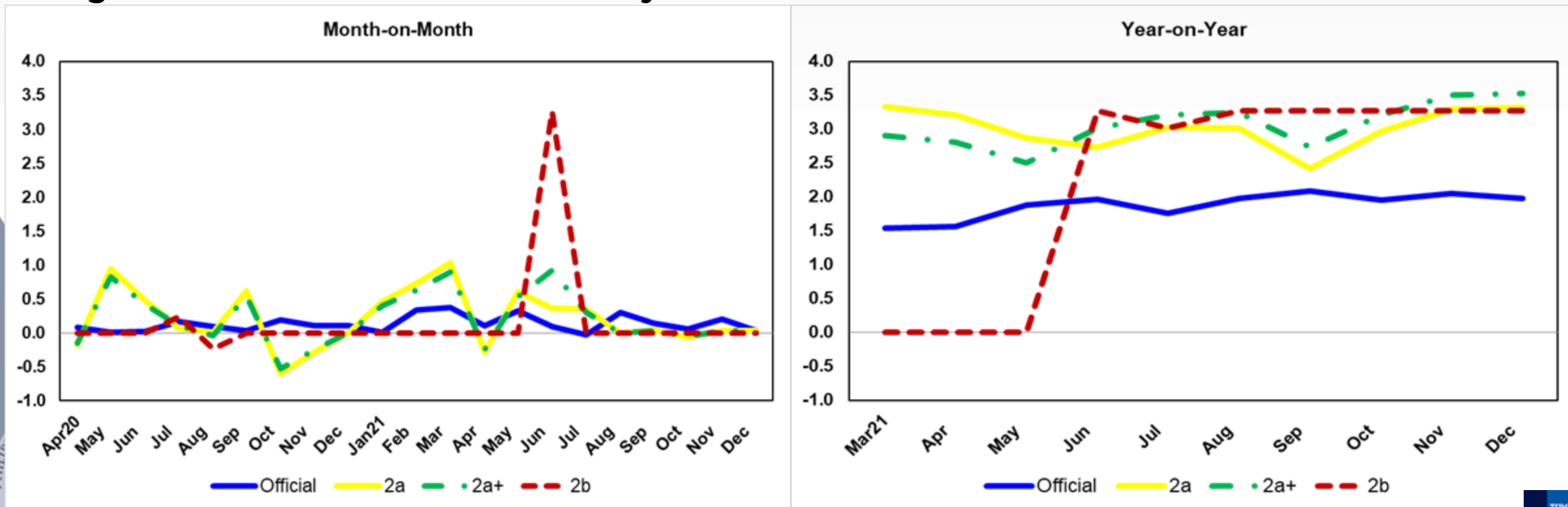## IV. RESULTS AND FINDINGS

Objective 2 – Subclass Level

Figure 8. Month-on-month & Year-on-year Growth Rates of CPI for Other Pharmaceutical Products

IV. RESULTS AND FINDINGS

Objective 2 – Subclass Level

Table 10. Absolute Deviations at the Subclass Level

Description	Online		Official		Absolute Deviation from the Official CPI				Period with Highest Deviation
	Mar 2020	Dec 2021	Mar 2020	Dec 2021	Lowest	Highest	Range	Average	
2a (Exact)									
Other cereal preparations	130.1	136.4	130.2	136.4	0.03	4.79	4.76	2.09	Sep. 2020
Other pharmaceutical products	120.3	126.7	121.5	124.9	0.11	2.64	2.53	1.32	Jul.2021
2a+ (Exact + Similar commodities)									
Other cereal preparations	130.1	135.7	130.2	136.4	0.03	3.87	3.84	1.44	Sep. 2020
Other pharmaceutical products	120.3	126.8	121.5	124.9	0.05	2.71	2.67	1.19	Jul. 2021
2b (Equivalent)									
Other cereal preparations	129.1	131.9	130.2	136.4	0.17	4.43	4.26	1.44	Dec. 2021
Other pharmaceutical products	121.3	125.2	121.5	124.9	0.20	2.60	2.40	0.96	May 2021

IV. RESULTS AND FINDINGS

Objective 2 – Subclass Level

Table 11. Mean Absolute Deviations of CPI Subclass using 2a, 2a+, and 2b

Mean Absolute Deviation (in Percentage Point)	Number of Subclasses					
	2a (Exact)		2a+ (Exact + Similar)		2b (Equivalent)	
	Count	%	Count	%	Count	%
0.00 – 0.99	10	17	6	10	13	22
1.00 – 1.99	10	17	14	24	15	25
2.00 – 2.99	12	20	12	20	7	12
3.00 – 3.99	7	12	9	15	3	5
4.00 and higher	20	34	18	31	21	36
Total	59	100	59	100	59	100

IV. RESULTS AND FINDINGS

Objective 2 – Subclass Level

Table 12. Range of Absolute Deviations of CPI Subclass using 2a, 2a+, and 2b

Range of Absolute Deviations (in Percentage Point)	Number of Subclasses					
	2a (Exact)		2a+ (Exact + Similar)		2b (Equivalent)	
	Count	%	Count	%	Count	%
0.00 – 0.99	2	3	2	3	3	5
1.00 – 1.99	4	7	3	6	7	12
2.00 – 2.99	10	17	12	20	10	17
3.00 – 3.99	6	10	5	8	5	8
4.00 and higher	37	63	37	63	34	58
Total	59	100	59	100	59	100

IV. RESULTS AND FINDINGS

Objective 2 – Subclass Level

Table 13. Cointegration Test of CPI Subclass using 2a, 2a+, & 2b vs Offline

Description	Pair	Trace	Critical Value (α =0.05)	ECT	Conclusion
Other cereal preparations	Offline vs 2a	13.02	18.17	-	No cointegration
	Offline vs 2a+	12.12	18.17	-	No cointegration
	Offline vs 2b	11.66	18.17	-	No cointegration
Other pharmaceutical products	Offline vs 2a	10.36	15.41	-	No cointegration
	Offline vs 2a+	10.36	15.41	-	No cointegration
	Offline vs 2b	17.38	15.41	0.0235	No cointegration

IV. RESULTS AND FINDINGS

Objective 2 – Subclass Level

Table 14. Summary of Cointegration Test of CPI Subclass using 2a, 2a+, & 2b vs Offline

Comparison	Total	Cointegrated	Not cointegrated
Offline vs 2a (Exact)[/]	59	9	43
Offline vs 2a+ (Exact+)[/b]	59	9	44
Offline vs 2b (Equivalent)	59	9	37

[/b] Components do not add up to total because some of the series have no variation, thus, not appropriate for cointegration test.

IV. RESULTS AND FINDINGS

Objective 2 – Subclass Level

Table 15. Summary of Cointegration Test of Subclass CPI using 2a, 2a+, & 2b by Division

Division	Total Subclass with Exact and Equivalent Specifications	No. of Subclasses with Cointegration		
		2a	2a+	2b
01 - Food and Non-alcoholic Beverages	25	1	2	2
02 - Alcoholic Beverages and Tobacco	1	0	0	0
03 - Clothing and Footwear	5	1	1	1
04 - Housing, Water, Electricity, Gas and Other Fuels	2	0	0	0
05 - Furnishings, Household Equipment and Routine Household Maintenance	10	2	2	3
06 - Health	3	3	3	1
07 - Transport	0	0	0	0
08 - Communication	0	0	0	0
09 - Recreation and Culture	9	1	0	1
10 - Education	-	-		-
11 - Restaurant and Miscellaneous Goods and Services	4	1	1	1

IV. RESULTS AND FINDINGS

Objective 2 – Subclass Level

- Comparing 2a (exact) with the official CPI, there are 10 subclasses with small MAD (less than one percentage point).

- Using 2a+ (exact + similar commodities), six out of 59 subclasses showed less than 1 percentage point in MAD.

- Using 2b (equivalent), there are 13 subclasses with small MAD (less than one percentage point) from the official CPI.

- Results of cointegration test showed the number of subclasses with cointegrating relation with the official CPI for each exploration.

| 2a (exact)	2a+ (exact+)	2b (equivalent)
9 subclasses	9 subclasses	9 subclasses

IV. RESULTS AND FINDINGS

Objective 2 – Subclass Level

- There are three subclasses common to all explorations. These are small electric household appliances, cleaning and maintenance products, and analgesics and muscle relaxants.

- Selected subclasses with higher percentage of commodities with exact specifications such as analgesic/antipyretics and muscle relaxants; cardiovascular drugs; and other pharmaceutical products showed small MADs and have significant cointegrating relations to the offline CPI.

IV. RESULTS AND FINDINGS

Objective 3 – Subclass Level

Figure 9. CPI Trend for Other Cereal Preparations
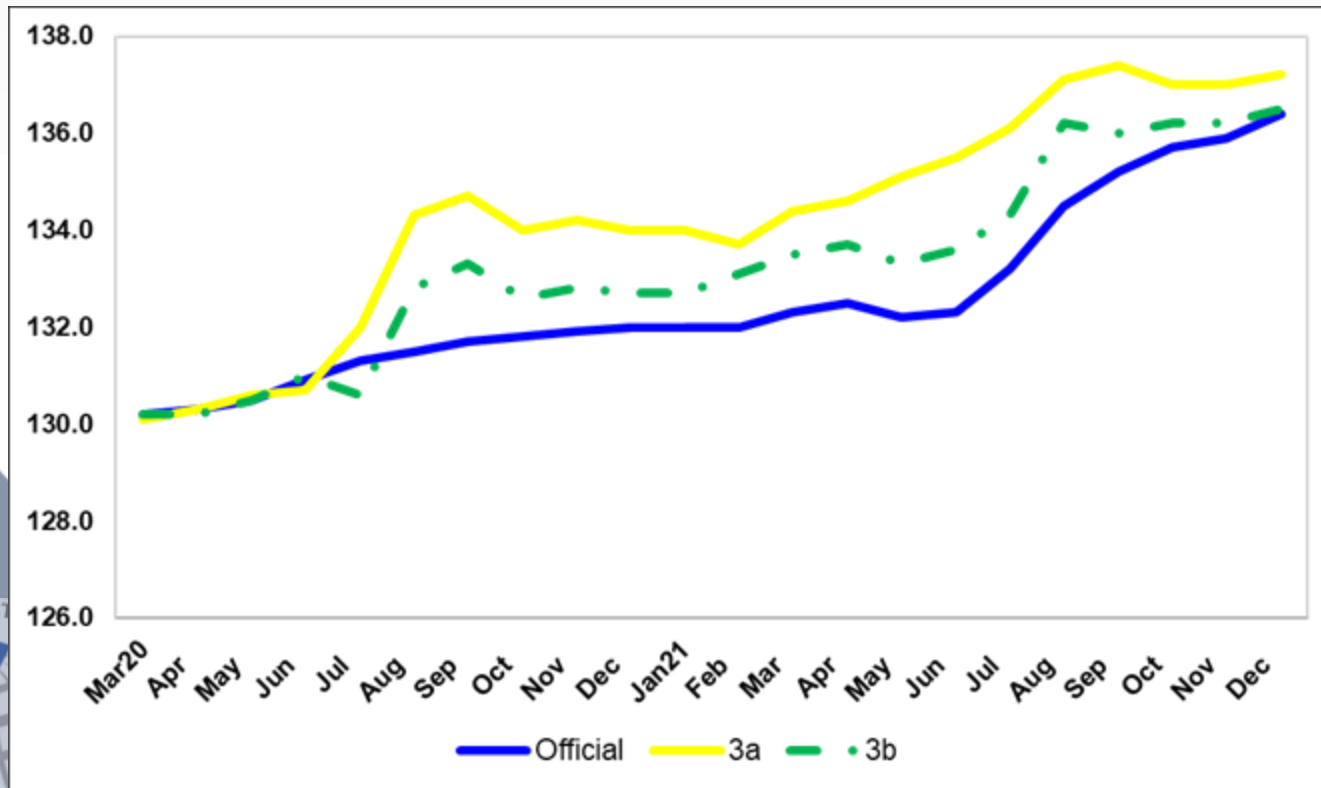
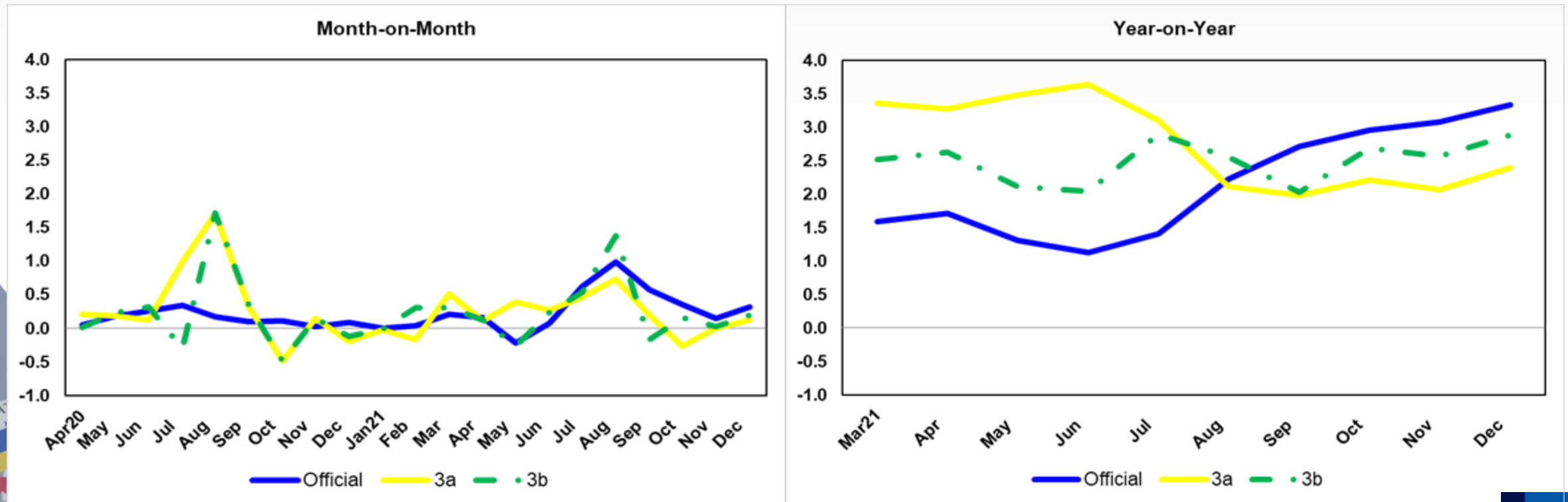Official: Consists of 11 commodities

3a : Consists of 8 commodities from online and 3 commodities from the official CPI market basket

3b : Consists of 5 online commodities with exact specifications and 6 commodities from the official CPI market basket

IV. RESULTS AND FINDINGS

Objective 3 – Subclass Level

Figure 10. Month-on-month & Year-on-year Growth Rates of CPI for Other Cereal Preparations

IV. RESULTS AND FINDINGS

Objective 3 – Subclass Level

Figure 11. CPI Trend for Cardiovascular Drugs

Official: Consists of 4 commodities

3a : Consists of 3 online commodities and 1 commodity from the official CPI market basket

3b : Consists of 2 online commodities with exact specifications 2 commodities in the official CPI market basket.

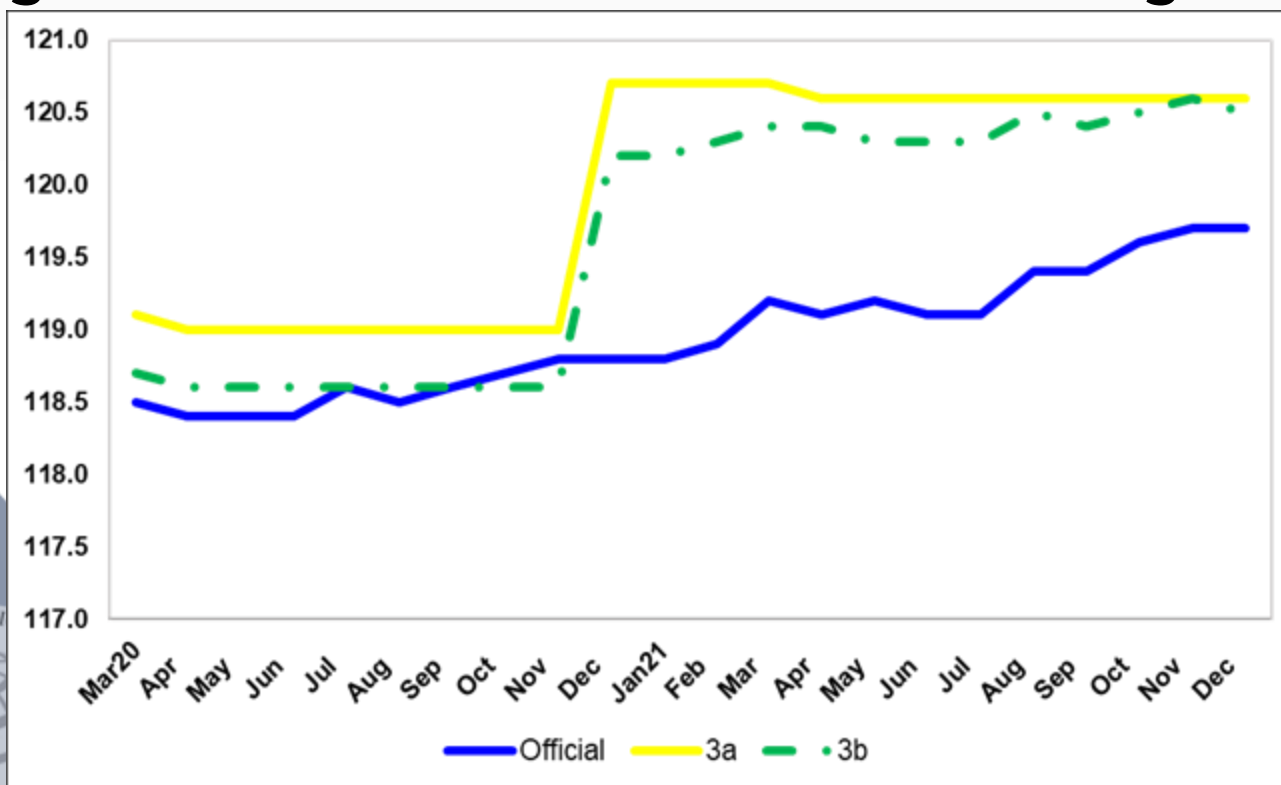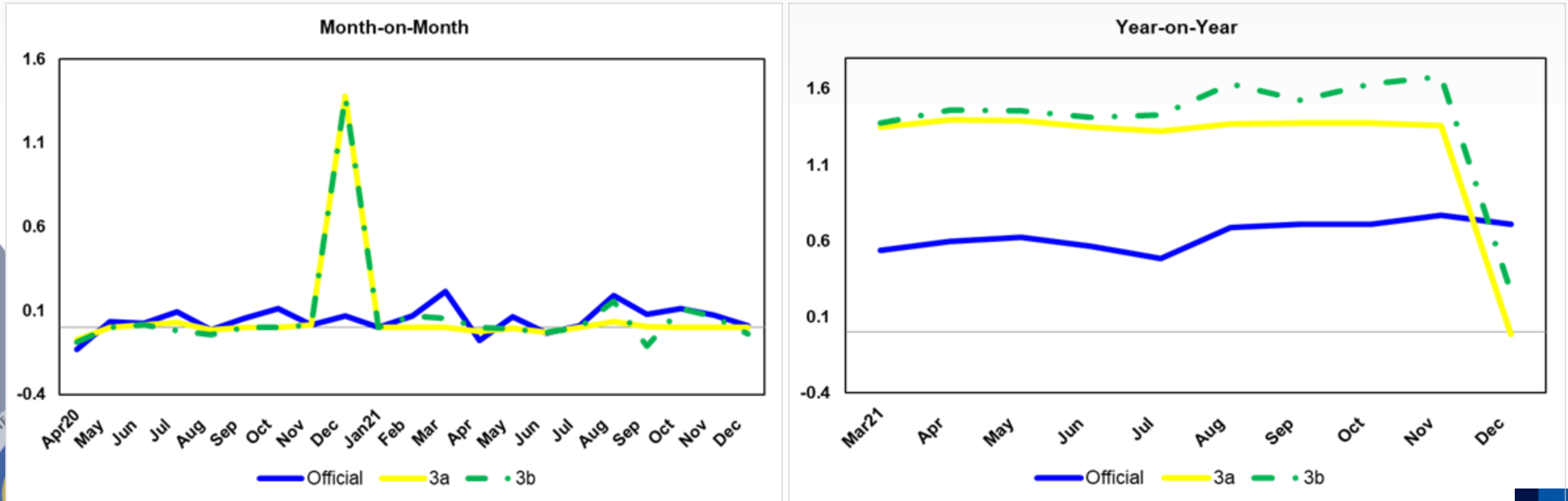## IV. RESULTS AND FINDINGS

Objective 3 – Subclass Level

Figure 12. Month-on-month & Year-on-year Growth Rates of CPI for Cardiovascular Drugs

IV. RESULTS AND FINDINGS

Objective 3 – Subclass Level

Table 16. Absolute Deviations at the Subclass Level

Description	Online		Official		Absolute Deviation from the Official CPI				Period with Highest Deviation
	Mar 2020	Dec 2021	Mar 2020	Dec 2021	Lowest	Highest	Range	Average	
3a (Hybrid1 = 1a + Offline) *(combination of online & offline regardless of stores and specs)*									
Other cereal preparations	130.1	137.2	130.2	136.4	0.04	3.15	3.11	1.75	Jun. 2021
Cardiovascular drugs	119.1	120.6	118.5	119.7	0.28	1.83	1.55	1.02	Dec. 2020
3b (Hybrid2 = 2a + Offline) *(combination of online and offline regardless of stores, but considering exact specs)*									
Other cereal preparations	130.2	136.5	130.2	136.4	0.01	1.72	1.71	0.78	Aug. 2021
Cardiovascular drugs	118.7	120.5	118.5	119.7	0.02	1.38	1.36	0.74	Dec. 2020

IV. RESULTS AND FINDINGS

Objective 3 – Subclass Level

Table 17. Mean Absolute Deviations of CPI Subclass using 3a and 3b from the official CPI

Mean Absolute Deviation (in Percentage Point)	Number of Subclasses			
	3a (Hybrid1)		3b (Hybrid2)	
	Count	%	Count	%
0.00 – 0.99	26	18	25	43
1.00 – 1.99	25	17	13	22
2.00 – 2.99	21	14	9	15
3.00 – 3.99	23	15	3	5
4.00 and higher	53	36	9	15
Total	148	100	59	100

IV. RESULTS AND FINDINGS

Objective 3 – Subclass Level

Table 18. Range of Absolute Deviations of CPI Subclass using 3a and 3b from the official CPI

Range of Absolute Deviations (in Percentage Point)	Number of Subclasses			
	3a (Hybrid1)		3b (Hybrid2)	
	Count	%	Count	%
0.00 – 0.99	11	7	12	20
1.00 – 1.99	10	7	11	19
2.00 – 2.99	15	10	8	14
3.00 – 3.99	15	10	10	17
4.00 and higher	97	66	18	30
Total	148	100	59	100

IV. RESULTS AND FINDINGS

Objective 3 – Subclass Level

Table 19. Cointegration Test of CPI Subclass using 3a & 3b vs Offline

Description	Pair	Trace	Critical Value (α =0.05)	ECT	Conclusion
Other cereal preparations	Offline vs 3a	6.22	15.41	-	No cointegration
	Offline vs 3b	25.73	25.32	-0.0741	No cointegration
Cardiovascular drugs	Offline vs 3a	14.55	12.53	-0.0206***	Significant cointegration
	Offline vs 3b	13.77	12.53	-0.0171***	Significant cointegration

IV. RESULTS AND FINDINGS

Objective 3 – Subclass Level

Table 20. Summary of Cointegration Test of CPI Subclass using 3a & 3b vs Offline

Comparison	Total	Cointegrated	Not cointegrated
Offline vs 3a (Hybrid1)[b]	148	32	94
Offline vs 3b (Hybrid2)	59	8	51

[b] Components do not add up to total because some of the series have no variation, thus, not appropriate for cointegration test.

IV. RESULTS AND FINDINGS

Objective 3 – Subclass Level

Table 21. Summary of Cointegration Test of Subclass CPI using 3a & 3b by Division

Division	No. of subclass conformed to Hybrid 1	3a (Hybrid1)	No. of subclass conformed to Hybrid 2	3b (Hybrid2)
01 - Food and Non-alcoholic Beverages	57	12	25	1
02 - Alcoholic Beverages and Tobacco	7	1	1	0
03 - Clothing and Footwear	12	5	5	1
04 - Housing, Water, Electricity, Gas and Other Fuels	3	1	2	0
05 - Furnishings, Household Equipment and Routine Household Maintenance	26	6	10	2
06 - Health	8	4	3	3
07 - Transport	2	0	0	0
08 - Communication	1	0	0	0
09 - Recreation and Culture	21	2	9	1
10 - Education	-	-	-	-
11 - Restaurant and Miscellaneous Goods and Services	11	1	4	0

IV. RESULTS AND FINDINGS

Objective 3 – Subclass Level

- Comparing 3a (Hybrid1, combination of 1a and offline) with the official CPI, there are 26 out of 148 subclasses have less than 1 percentage point MAD.

- Using 3b (Hybrid2, combination of 2a and offline), 25 of 59 subclasses showed less than 1 percentage point in MAD from the official CPI.

- Results from cointegration test, showed 32 of 148 subclasses are cointegrated using 3a (Hybrid1) and 8 of 59 subclasses are cointegrated using 3b (Hybrid2).

- Eight cointegrated subclasses are common for the two explorations.

V. RECOMMENDATIONS

1. To have further studies exploring one-to-one correspondence of offline and online sample stores and one-to-one correspondence of commodity specifications (2a) to address the issue on data source for computed CPI.

2. To follow the NCR's schedule of price collection for web scraping.

3. To use longer series of web scraped data for a more reliable cointegration test.

4. To increase the number of items covered for each commodity groups for online price collection, if possible.

Thank you

psa.gov.ph

openstat.psa.gov.ph

@PSAgovph