

**MINISTRY OF STATISTICS AND PROGRAMME IMPLEMENTATION  
NATIONAL STATISTICAL OFFICE(NSO)**

**COMPILATION OF CONSUMER PRICE INDEX: - A TECHNICAL NOTE**

The National Statistical Office (NSO), Ministry of Statistics and Programme Implementation (MoSPI) compiles All India as well as state-wise Consumer Price Index (CPI) for Rural, Urban, Combined sectors and releases the CPI numbers on 12<sup>th</sup> day of every month for current month and Final Index of previous month. The Press Release for the Price Movements of Selected Groups/Sub-groups for May 2020 (Provisional) and April 2020 (Final) may be seen at MoSPI website. The methodology for compilation of the index is at **Annex I**.

**A. CONTEXT**

2. In view of the preventive measures and announcement of nation-wide lockdown by the Government to contain spread of COVID-19 pandemic and associated travel advisories, it was not feasible for the field enumerators to visit the specified retail outlets to collect data/price quotations. Further, there was also limited transaction of products in the market in the months of April and May, 2020 and thus price data on all the commodities were not available. In these circumstances, the NSO field enumerators collected data through telephone from the designated outlets and augmented this with visits to the neighbourhood shops for items being transacted. As the number price quotations received was less, it was not feasible to construct / compute either State level indices or the all India General CPI. The Price Movement of Sub-groups/Groups of CPI which met the principles of adequacy as follows:

- i) The price of only those items were included which have been reported from at least 25% of markets, separately for Rural & Urban sector and constituted more than 70% weight of the respective Sub-groups/Groups.
- ii) The all India indices have been compiled using Direct Approach by considering a common market at the national level in both Rural & Urban sectors separately.

Thus, of the 23 sub-groups, the indices of only 13 sub-groups in April, 2020 and 15 sub-groups in May, 2020 were compiled and disseminated.

**B. DETAILED DISCUSSIONS AT GLOBAL LEVEL**

3. During this time, a series of discussions were held with International Organisations like World Bank, ILO, ADB, IMF, Eurostat, OECD and several other countries on how they are managing and addressing the data flow challenges on account of the pandemic. While few countries estimated the missing data using

imputation methods, their scope, coverage and size was not comparable to the Indian case. Indian concern was shared with these Organisations and in parallel discussions were also going on in the Inter-Secretariat Working Group on Price Statistics (ISWG-PS) (combined forum of ILO, EuroStat, OECD, UNECE, World Bank and IMF) also brought out its **Business Continuity Guidelines** in May 2020 which goes into details of imputation for non-transacted items. The ISWG-PS recommended as follows:

*“If an entire index is missing, it is recommended to use the next level up in aggregation as the basis for making the imputation. For example, if all prices for oranges are missing, the index for citrus fruits can be used as the basis for making the imputation. If all citrus fruits are missing, the index for fruits is used as the basis for making the imputation. If all fruits are missing, the index for fruits and vegetables is used, and so on”.*

*“Missing elementary aggregates could be imputed by the all items index calculated using the price movements of those sub-indices for which there are data. First, all sub-indices which have been calculated based on a majority of observed prices are used to compile the all item index. Second, the short-term change of the all items index is then used as the basis for imputing the missing indexes. Imputation with the all items CPI corresponds to leaving the elementary aggregate out of the calculation of the CPI.”*

4. In the Indian context the above approach implied that imputation would need to be done for even those groups / sub-groups of CPI where no data was available or collected or transacted. Under normal circumstances there are disruptions in data flows for a few commodities which are generally localized in nature and imputation works well. This is for the first time that the data flow was disrupted not at the local level but at the national level for several commodities. The recommendation of the ISWG-PS for India needs to be seen in this context.

5. This Technical Note has been prepared with the objective for apprising and explaining to the CPI users and stakeholders about the methodology proposed to be followed for filling gaps in CPI Series on account of non-availability of price data during April-May, 2020, in the wake of COVID-19 pandemic.

### C. IMPUTATION METHODOLOGY FOR CPI

6. NSO has undertaken an exercise for imputing missing Sub-groups/Groups by using next level up index i.e. Groups/General CPI calculated with observed price during lockdown period in pursuance to recommendations of ISWG-PS. Under this exercise, '**Food & Beverages Group Index**', calculated on collected prices, is used for imputing the index the sub-groups of **Food & Beverages Group**. Similarly, for the **General CPI**, calculated on the basis of the sub-group indices where data was available, was used for imputing the indices for the Sub-groups/Groups other than '**Food & Beverages**' where price data was not captured. The imputation approach used the following formulae:

### Formula I

Imputed Sub-group Index (Food Group) = (Last Observed Index of Sub-group) × (Food Index of Current Month/Food Index of last Observed Month)

### Formula II

Imputed Sub-group Index (Non-Food Group) = (Last Observed Index of Sub-group) × (General CPI of Current Month/General CPI of last Observed Month).

7. Thus, for month of May, 2020, the index for sub-group of '*Prepared meals, snacks, sweets etc.*' was imputed by using the Index of '*Food & Beverages Group*', calculated on the basis of the 11 sub-groups for which price quotation data were available. Similarly, the index for the sub-groups of '*Clothing & Footwear*' group and '*Miscellaneous*' group were imputed by using **General CPI**, calculated on the basis of 15 sub-groups for which price quotation data were available. The detailed procedure is at **Annex II**.

8. Proceeding on the above lines, after imputing the indices for all sub-groups where data was not available/collected, the General CPI has been re-calculated for all 23 sub-groups. (**Table-1**)

**Table 1: Imputed Indices<sup>@</sup> of Sub-Groups/Groups of All India CPI for the month of April, 2020 and May, 2020**

(Base: 2012=100)

Group Code	Sub-group Code	Description	Rural			Urban			Combined		
			Weights	Apr. 20 Index	May 20 Index	Weights	Apr. 20 Index	May 20 Index	Weights	Apr. 20 Index	May 20 Index
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	1.1.01	Cereals and products	12.35	147.2	147.5	6.59	151.8	150.4	9.67	148.7	148.4
	1.1.02	Meat and fish	4.38	168.9 @	181.5	2.73	171.3 @	188.1	3.61	169.7 @	183.8
	1.1.03	Egg	0.49	146.9	146.4	0.36	151.9	150.0	0.43	148.8	147.8
	1.1.04	Milk and products	7.72	155.6	154.9	5.33	155.5	155.4	6.61	155.6	155.1
	1.1.05	Oils and fats	4.21	137.1	139.2	2.81	131.6	131.9	3.56	135.1	136.5
	1.1.06	Fruits	2.88	147.3	146.2	2.90	152.9	153.0	2.89	149.9	149.4
	1.1.07	Vegetables	7.46	162.7	145.1	4.41	180.0	161.8	6.04	168.6	150.8
	1.1.08	Pulses and products	2.95	150.2	151.1	1.73	150.8	151.4	2.38	150.4	151.2
	1.1.09	Sugar and Confectionery	1.70	119.8	116.2	0.97	121.2	117.2	1.36	120.3	116.5
	1.1.10	Spices	3.11	158.7	158.7	1.79	154.0	154.7	2.50	157.1	157.4
	1.2.11	Non-alcoholic beverages	1.37	139.2	141.4	1.13	133.5	134.1	1.26	136.8	138.4
	1.1.12	Prepared meals, snacks, sweets etc.	5.56	162.1 @	161.9 @	5.54	162.7 @	162.4 @	5.55	162.4 @	162.1 @
<b>1</b>		<b>Food and beverages</b>	<b>54.18</b>	<b>152.8 @</b>	<b>151.4 @</b>	<b>36.29</b>	<b>156.1 @</b>	<b>154.8 @</b>	<b>45.86</b>	<b>154.0 @</b>	<b>152.7 @</b>

Group Code	Sub-group Code	Description	Rural			Urban			Combined		
			Weights	Apr. 20 Index	May 20 Index	Weights	Apr. 20 Index	May 20 Index	Weights	Apr. 20 Index	May 20 Index
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
2		Pan, tobacco and intoxicants	3.26	171.1 @	171.2 @	1.36	179.1 @	183.4	2.38	173.2 @	174.4 @
	3.1.01	Clothing	6.32	153.9 @	154.0 @	4.72	152.6 @	153.0 @	5.58	153.4 @	153.6 @
	3.1.02	Footwear	1.04	148.1 @	148.2 @	0.85	138.3 @	138.6 @	0.95	144.0 @	144.2 @
3		Clothing and footwear	7.36	153.1 @	153.2 @	5.57	150.4 @	150.8 @	6.53	152.0 @	152.3 @
4		Housing	#	#	#	21.67	155.6	155.6	10.07	155.6	155.6
5		Fuel and light	7.94	148.4	146.4	5.58	137.1	136.2	6.84	144.1	142.5
	6.1.01	Household goods and services	3.75	152.0 @	152.1 @	3.87	145.5 @	145.9 @	3.80	148.9 @	149.2 @
	6.1.02	Health	6.83	154.3	157.0	4.81	144.8	146.1	5.89	150.7	152.9
	6.1.03	Transport and communication	7.60	136.3 @	136.3 @	9.73	128.7 @	129.1 @	8.59	132.3 @	132.5 @
	6.1.04	Recreation and amusement	1.37	151.7 @	151.8 @	2.04	142.5 @	142.9 @	1.68	146.5 @	146.8 @
	6.1.05	Education	3.46	161.7 @	161.8 @	5.62	157.6 @	158.0 @	4.46	159.3 @	159.6 @
	6.1.06	Personal care and effects	4.25	145.6 @	145.7 @	3.47	150.1 @	150.5 @	3.89	147.5 @	147.7 @
6		Miscellaneous	27.26	148.4 @	149.2 @	29.53	142.5 @	143.0 @	28.32	145.5 @	146.2 @
General Index (All Groups)			100.00	151.9 @	151.2 @	100.00	150.9 @	150.6 @	100.00	151.4 @	150.9 @
Consumer Food Price Index (CFPI)			47.25	152.1 @	150.5	29.62	155.7 @	154.2	39.06	153.4 @	151.8

Notes:

1. # : CPI (Rural) for Housing is not compiled
2. @ : Imputed Index for April and May 2020.

#### D. Impact and future implications:

9. This Technical Note has been brought out after detailed discussion with various International Organisations and countries and the recommendations of the ISWG-PS. The essential debate focused around preparing partial indices, as released for April and May, 2020 vis-à-vis preparing comprehensive indices using imputation method on a large and unprecedented scale. The ISWG-PS approach indicates that the balance of convenience, continuity and comparability (with limitations) would be in favour of preparing the complete CPI even though there is large scale imputation. A significant part of CPI price data can no longer be captured for April and May 2020. The series so compiled is only for the purpose of providing estimates as a business continuity approach. Thus the imputed CPI series are to be seen in this light especially when using it in a time series comparison within and across countries. It is proposed to continue with this approach in the following months till the data collection exercise resumes full operation.

\*\*\*

## BRIEF METHODOLOGY FOR COMPILATION OF CPI

The monthly price data are collected from selected 1114 urban markets and selected 1181 villages through personal visits by field staff of Field Operations Division (FOD) of NSO, MoSPI on a weekly roster. The collected price data are uploaded by FOD on Web Portals developed by NIC. On receiving the collected price, scrutiny and validation checks are carried out by Price Statistics Division (PSD). NSO follows well established and internationally accepted methodology and practices while handling the situation of missing prices, change in item specifications and seasonal items;

- a) **Treatment of temporarily missing prices:** The temporarily missing prices are imputed by using following formula;

$$\text{Imputed Price} = \text{Last Month's Price} \times \left[ \begin{array}{l} \text{Average of Ratio (Current Month} \\ \text{to Last Month's price, Based on} \\ \text{Quotations having prices in both} \\ \text{months)} \end{array} \right]$$

i.e.  $(\text{Imputed Price})_t = (\text{Price})_{t-1} \times \text{Avg. of } (\text{Price}_t / \text{Price}_{t-1})$

It may be noted that the imputation technique mentioned above is a standard one and recommended in *Manuals on CPI*, which is usually used at micro-level imputation in regular compilation of CPI. The need was never felt to use it for macro level imputation.

- b) **Treatment of permanently missing prices:** The allocated weights of such items are distributed (imputed) on pro-rata basis to the related items / to respective, section/sub-group/group. In such cases, the effective size of item basket will be lessened by no. of items permanently obsoleted. This practice is rarely followed in CPI, obsoleted items are, generally, substituted with items, which are serving similar purpose.

- c) **Change in item specification:** Base Price is modified, when shop or item is substituted, by using following formula;

$$\text{New Base Price} = \frac{\text{Price of new specification of last month}}{\text{Price of old specification of last month}} \times \text{Old Base Price}$$

d) **Treatment for Seasonal items:** Treatment of seasonality is applied only in case of vegetables and fruits. Whenever, in a particular month, the prices of some of the items are either not reported at all or reported in less than 25% quotations of the total allotted quotations of that State/UT then weights of such items are imputed on pro-rata basis to the items first in the respective section (i.e. root vegetables, leafy vegetables, fresh fruits, and dry fruits, etc.), as applicable.

3. The Consumer Price Index (CPI) is compiled in following stages by using *bottom-up approach*:

- (a) The elementary indices are the lowest level of aggregation where prices of each good or service are combined into price indices for which explicit expenditure weights are, generally, not available. For example, Prices of different varieties of Rice or Prices of Rice from different outlets. The Price Ratios (current month to base period) are combined by taking simple geometric mean.
- (b) Aggregation of these elementary price indices to higher level indices using relative levels of consumer expenditure as weights. At this stage, a Laspeyres-type index formula is, generally, used.
- (c) By following above two stages, Items Indices, Sub-groups/Groups indices are, first, compiled for Rural, Urban & Combined sectors of each State/UT. All India Indices are compiled by aggregating weighted State/UT indices, for Rural, Urban & Combined sectors, respectively.

IMPUTATION TECHNIQUES<sup>1</sup>

## a. Elementary (micro) Level Imputation

$$\text{Imputed Price} = \text{Last Month's Price} \times \left[ \begin{array}{l} \text{Average of Ratio (Current Month to} \\ \text{Last Month's price, Based on} \\ \text{Quotations having prices in both} \\ \text{months)} \end{array} \right]$$

$$\text{i.e. (Imputed Price)}_t = (\text{Price})_{t-1} \times \text{Avg. of } (\text{Price}_t / \text{Price}_{t-1})$$

## b. Higher Level (Macro) Imputation

- i. *Imputed Sub-group Index (Food Group)* =  
(Last Observed Index of Sub-group) × (Food Index of Current Month/Food Index of last Observed Month)
- ii. *Imputed Sub-group Index (Non-Food Group)* =  
(Last Observed Index of Sub-group) × (General CPI of Current Month/General CPI of last Observed Month).

Illustration showing compilation of indices when some group/sub-group is missing:

Sl. No.	Group/Sub-Group	Weight	Index_ current month	Modified weight	Index_ prev month
1	Sub-group 1	$W_{11}$	$C_{11}$	$W_{11}^* = (W_{11} \times 100) / (W_{11} + W_{12} + W_{22})$	$P_{11}$
2	Sub-group 2	$W_{12}$	$C_{12}$	$W_{12}^* = (W_{12} \times 100) / (W_{11} + W_{12} + W_{22})$	$P_{12}$
3	Sub-group 3	$W_{13}$	-	-	$P_{13}$
4	<b>Group 1</b>	<b><math>W_1</math></b>	$C_1$	$W_1^* = W_{11}^* + W_{12}^*$	$P_1$
5	Sub-group 4	$W_{21}$	-	-	$P_{21}$
6	Sub-group 5	$W_{22}$	$C_{22}$	$W_{22}^* = (W_{22} \times 100) / (W_{11} + W_{12} + W_{22})$	$P_{22}$
7	<b>Group 2</b>	<b><math>W_2</math></b>	-	-	$P_2$
8	<b>Group 3</b>	<b><math>W_3</math></b>	-	-	$P_3$
	Total (General)	100	C	100	P

$$W_i = \sum_j W_{ij} \text{ and } \sum_i W_i = 100, \sum_i W_i^* = 100$$

For previous month,

Group Index ( $P_i$ ) say,  $P_1 = (\sum_{j=1}^3 P_{1j} \times W_{1j}) / \sum_{j=1}^3 W_{1j}$  and so on

General Index (P) =  $(\sum_{i=1}^3 P_i \times W_i) / \sum_{i=1}^3 W_i$

<sup>1</sup>For more information on the different imputation methods, please see <https://www.imf.org/en/Data/Statistics/cpi-manual>. Chapter 6 discusses the treatment of temporarily missing prices in detail

For current month,

Group Index ( $C_i$ ) say,  $C_1 = (\sum_{j=1}^3 C_{1j} \times W_{1j}^*) / \sum_{j=1}^3 W_{1j}^*$  and so on

$$\text{General Index (C)} = \frac{(C_{11} \times W_{11}^*) + (C_{12} \times W_{12}^*) + (C_{22} \times W_{22}^*)}{W_{11}^* + W_{12}^* + W_{22}^*}$$

where  $(W_{11}^* + W_{12}^* + W_{22}^*) = 100$

To fill-in the gap of CPI series in the current month with observed index behavior from previous month:

Sl. No.	Group/Sub-Group	Modified weight	Index_current month	Index_prev month	Modified Index_Current month	Weight
1	Sub-group 1	$W_{11}^* = (W_{11}^* \times 100) / (W_{11} + W_{12} + W_{22})$	$C_{11}$	$P_{11}$	$C_{11}$	$W_{11}$
2	Sub-group 2	$W_{12}^* = (W_{12}^* \times 100) / (W_{11} + W_{12} + W_{22})$	$C_{12}$	$P_{12}$	$C_{12}$	$W_{12}$
3	Sub-group 3	-	-	$P_{13}$	$C_{13} = P_{13} \times (\frac{C_1}{P_1})$	$W_{13}$
4	<b>Group 1</b>	$W_1^* = W_{11}^* + W_{12}^*$	$C_1$	$P_1$	$C_1^*$	$W_1$
5	Sub-group 4	-	-	$P_{21}$	$C_{21}^* = P_{21} \times (\frac{C}{P})$	$W_{21}$
6	Sub-group 5	$W_{22}^* = (W_{22}^* \times 100) / (W_{11} + W_{12} + W_{22})$	$C_{22}$	$P_{22}$	$C_{22}$	$W_{22}$
7	<b>Group 2</b>	-	-	$P_2$	$C_2^*$	$W_2$
8	<b>Group 3</b>	-	-	$P_3$	$C_3^* = P_3 \times (\frac{C}{P})$	$W_3$
	Total (General)	100	C	P	$C^*$	100

Modified Group Index -

$$C_1^* = (C_{11} \times W_{11} + C_{12} \times W_{12} + C_{13}^* \times W_{13}) / W_1$$

$$C_2^* = (C_{21}^* \times W_{21} + C_{22} \times W_{22}) / W_2$$

$$\text{General Index (C}^*) = \frac{(C_1^* \times W_1 + C_2^* \times W_2 + C_3^* \times W_3)}{(W_1 + W_2 + W_3)}$$

For example, for imputing the missing index of sub-group '*Prepared meals, snacks, sweets etc.*' in April, 2020 (Rural);

- Last observed index of '*Prepared meals, snacks, sweets etc.*' is March, 2020 is **160.0**
- 'Food & Beverages' Index for the month of April, 2020 is **150.1**, which was calculated on the basis of 10 observed sub-groups of 'Food & Beverages' group
- 'Food & Beverages' Index for the month of March, 2020 is **148.2**

The imputed index of sub-group '*Prepared meals, snacks, sweets etc.*' in April, 2020 (Rural) =  $160.0 \times (150.1 / 148.2) = 162.2$



Similarly, for imputing the missing index of sub-group '**Clothing**' in April, 2020 (Rural);

- a) Last observed index of '**Clothing.**' is March, 2020 is **153.4**
- b) 'General CPI' for the month of April, 2020 is **150.3**, which was calculated on the basis of 13 observed sub-groups.
- c) 'General CPI' for the month of March, 2020 is **149.8**

The imputed index of sub-group '**Clothing**' in April, 2020 (Rural) =  $153.4 \times (150.3/149.8) = \mathbf{153.9}$

\*\*\*\*\*