

**CODE FOR SELF REGULATION FOR USE OF IRS DATA**  
**(updated w.e.f. 5<sup>th</sup> June, 2020)**

**A. Objective**

The IRS is a collaborative study serving the research needs of the members of the MRUCI. As such, it is desirable that the IRS data is used responsibly by all the member-subscribers and hence the need for a self-regulating code for use of the IRS data. The Code is to be read along with the IRS End Users License Agreement signed and accepted by all subscribers along with their respective IRS 2019 subscriptions, and which is binding on all Subscribers. The Code for self-regulation is to provide for:

- a) Common understanding of research terms, data definition and data analysis norms as provided by MRUCI, including the statistical rules governing the IRS (Appendix I);
- b) Adherence to data analysis and reporting norms for truthful representation of data to clients, readers of respective publications and the public at large (Appendix I);
- c) Redressal mechanism for complaints received in respect of data misrepresentation, abusing/disregarding set statistical norms and/or any other data usage provided by this Code.
- d) Any claim or comparison for publicity/promotion, hoardings or communication in any other form which is in violation of the protocols laid down in para B below will be subject to adjudication as per the redressal mechanism specified in this document.

**Disclaimer**

The IRS 2019 data is licensed to the IRS Subscriber for a period of two years and as such the IRS data can only be used by registered subscribers of the data. MRUCI retains the right to initiate legal action against any access or unauthorized use of the data by non-subscribers and any unauthorized sharing of the data by any subscriber.

## B. Code of Self-Regulation

### a) Comparisons with previous rounds

IRS 2019 data comprises 4 quarters data of 2019 (Q1 '19 + Q2'19 + Q3'19 + Q4'19).

Note that the 2019 Q4 release contains the most recent data collected and across 4 quarters essentially from 2019. Unlike earlier Quarterly reports (IRS 2019 Q1 ,IRS 2019 Q2, and IRS 2019 Q3) there is no break in the data collection period. Please consider the field work duration of previous report/s while comparing the data.

### b) Use of IRS 2019 data in publicity and promotions

In all cases of publicity/promotion, hoardings or communication in any other form, where Readership / Listenership / Viewership data may be compared with two or more publications / radio stations / TV channels / any other media, users should abide by the following protocol:

- (i) Any claim for “leading”, “No.1” or to establish top position by any parameter/s should **ONLY** be based on a like to like comparison, i.e. the same set of readership / listenership / viewership data to be compared amongst publications / radio stations / TV channels / any other media, sourced only from the same quarterly report of IRS 2019.
- (ii) Comparison should be restricted to publications of similar publishing frequencies, i.e., Daily newspapers to Daily newspapers, weekly newspapers to weekly newspapers and magazines to magazines.
- (iii) Comparison should be between similar readership frequencies, and restricted to any one of the pre-defined metrics – i.e. AIR to AIR, Upto 3 days to Upto 3 days, Upto 7 days to Upto 7 days and TR to TR.

(iv) For any specific demographic, geographic, product profile category the comparison should be done with the same set of comparable readership / listenership / viewership numbers for all publications / radio stations / TV channels / any other media.

**The applicable parameter/s along with the specific geographical area (City, State, Country) must be clearly mentioned (in the same font size as the headline) in the publicity material.**

**c) Variant readership**

- (i) The “main” edition and the “variant” edition should be treated as two different newspapers
- (ii) Publishers should refrain from issuing misleading communication wherein readership numbers of Main + Variant are reported as a single readership number for that publication.
- (iii) Publishers can use the Main + Variant readership numbers together, provided they report the Main & Variant numbers separately and not only as one single consolidated number in their communication. For eg, if for a given publication, Main Issue Readership = 100 readers; Variant Issue Readership = 20 readers, then the publication can mention in their communication that their readership numbers are: Main Issue =100; Variant Issue =20; Main + Variant Total = 120.
- (iv) Media planners should verify readership being used for comparison between publications, i.e., Main or Variant, and ensure like-to-like comparisons are used for selection of publications
- (v) Readership numbers of Main issue of one publication should be compared only with the Main issue readership numbers for other publications.

(vi) For any cost comparisons, it is essential to verify whether the right rates, from the Main issue or Variant issue rate cards, are being used for all selected publications to ensure no incorrect comparisons are made.

**d) Comparisons between quarterly IRS reports**

The IRS 2019 quarterly reports are based on the principle of Moving Average Totals, wherein the data over the immediate 3 preceding quarters of fieldwork are aggregated along with the latest quarter of fieldwork, to provide an annual estimate over 4 successive quarters. For illustrative purposes, IRS Q4 2019 Report will comprise of data collected over the fieldwork period of **Q1 2019 (Nov2018 – Apr 2019) + Q2 2019 (Apr 2019 – Jul 2019) + Q3 2019 (Aug – Nov 2019) + Q4 2019(Dec 2019 – Mar 2020)**.

The IRS Q3 2019 Report will comprise of data collected over the fieldwork period of **Q4 2017 (Aug 2017 – Dec 2017) + Q1 2019 (Nov2018 – Mar 2019) + Q2 2019 (Apr 2019 – Jul 2019) + Q3 2019 (Aug – Nov 2019)**.

As such, subsequent Quarterly reports of the IRS study cannot be used to arrive at a combined average estimate. For illustrative purposes, estimates published respectively in IRS Q1 2019, IRS Q2 2019, IRS Q3 2019 and IRS Q4 2019 reports cannot be averaged to represent a derived estimate for full year 2019. The Subscriber should not present data which is averaged over subsequent IRS reports, but treat each individual report as an annualized estimate of the immediate previous 4 quarters of fieldwork.

**e. Any derivative work done by subscribers using IRS data**

Any estimates derived by subscribers on their own, using published IRS data and used for public communication, should clearly highlight the fact that the said estimates are derived independently by the subscriber, and MRUCI will not be held accountable for the accuracy of the same. . It is strongly recommended that before publishing any such derived estimates, subscribers should verify the same with MRUCI.

MRUCI shall be obliged to inform all their members / subscribers if they are not in agreement with the numbers independently derived by the subscribers, and this may be further publicised by any member / subscriber, as they desire.

### C. Redressal Mechanism

Any violation under this Code by any subscriber/member/user, will be adjudicated by a Disciplinary Committee (**DisCom**) constituted by the MRUCI.

- a) Any member and/or subscriber affected by the release of any publicity material violating the code provided shall write to the MRUCI **DisCom** within seven days of the date of publicity with the following:
  - (i) Covering letter from the Owner/CEO of the Complainant giving
    - 1) Name and Address of the Complainant
    - 2) Name and Address of the Opposite party
    - 3) Details of the nature of violation along with corresponding claim of the Complainant
  - (ii) Actual copy in the case of printed publicity (newspaper, magazine or flyer), colour photograph in case of any Outdoor medium and a recording in case of Audio-visual
  - (iii) A DD of Rs. 10,000/- issued in favour of MRUCI
- b) The **DisCom** shall deliberate on the Complaint within seven days of receipt of the complaint and convey its decision to the Complainant and the Opposite party within seven days of its decision
- c) If the Opposite Party is found guilty of contravening the Code, **DisCom** will propose following corrective action against the offender, which may include:

- The said publication to print a corrigendum on the same page and size of the publicity in case of publicity/campaign in a newspaper and/or magazine (text of corrigendum as advised by MRUCI)
  - In case of any other form of publicity/campaign, then the corrigendum to be published on page 3 in size 20 x 3 cc as per draft of the corrigendum advised by MRUCI.
  - Corrigendum to be published within a period of 15 days from the date of decision of the **DisCom**
  - MRUCI may advise all affected competing publications to publish the same corrigendum on no cost basis.
  - MRUCI may also inform all its members about the said violation.
- d) MRUCI also retains the right to take suo moto cognisance of any violation of the guidelines under this code.
- e) The decision of the MRUCI Disciplinary Committee will be final.

## **APPENDIX I**

### **Guidelines for Responsible Use of IRS Data**

#### **1) Statistical rules governing IRS data**

- a) The estimations made and provided in the IRS are based on random sampling methodology and all such estimates operate within 20% (approximate) sampling / non-sampling error level for any reporting breaks with 90% confidence level for 10% incidence.

- b) Estimates for a title are suppressed if the unweighted all-adult readership for the title for the period of the survey, in the respective reporting unit of IRS, falls below 30.
- c) For the internal use of subscribers, a separate database referred as “**Reference Only Publication DB**” has been provided in the IRS report, which provides limited geographic/demographic variables along with print readership variables such as TR and AIR. This database includes readership estimates for such publications/titles whose sample counts are less than 30, and thereby not reported in the Main IRS database.
- a. The **Reference Only Publication DB** database provides only the ‘Reporting Units’ variable under Demographics Information. This data provides publications with their penetration number but does not allow any further detailed analysis.
  - b. The difference in reporting units as provided under IRS Main database vs **Reference Only Publication DB** database are as below:

<b>List of Reporting Units</b>	
<b>Main Database</b>  (Does not include readership estimates based on low sample count i.e below 30)	<b>Reference Only Publication DB</b>  (Includes readership estimates based on low sample count i.e below 30)
State	State
Urban/Rural	Urban/Rural
Zone	-----
Pop-strata	Pop-strata
SCR	SCR
District/Dist. Cluster	District/Dist. Cluster
Reporting Towns	Reporting Towns
Top 4 Metros by Zone	-----

iii. The **Reference Only Publication DB** database has been released by MRUC only for internal reference / planning by the subscribers, and readership estimates reported in this database cannot be published/used for any commercial purposes.

Any estimate based on a respondent count of less than 30, does not meet Globally accepted statistical thresholds and carries a very high Margin of Error.

Below is an illustrative example from the **Reference Only Publication DB**, wherein it is observed that while Publication A's AIR estimate in Mumbai can be read as 52000, (derived basis sample count of 16), the relative error associated with it is 43%, which is way above the acceptable norm of 20%.

AIR of Publication A	Chennai UA	Mumbai UA
Sample Size	3860	5872
<b>Unweighted Count</b>	<b>114</b>	<b>16</b>
Weighted Count ('000)	263	52
<b>Margin of Error</b>	<b>16.9%</b>	<b>43%</b>

## 2) Data classification

### a) Household vs. Individual data

- a. The IRS reports Household data and Individual data separately
- b. Penetration of Durables ownership and purchase of consumer goods is captured at household level and should not be used for universe sizing in the Individual database.
- c. Individual data should be considered only for consumer profiling. It should be read as the targeted Individuals who own those durables or purchase the respective consumer goods, in their household.



b) Readership metrics

Readership for Daily Newspapers is reported by the following metrics-

- AIR - Yesterday Readership
- 1 to 3 days Readership
- 1 to 7 days Readership
- Total Readership – last 4 weeks

c) Variant Readership

- IRS 2019 captures readership for both - main paper as well as variants of the main paper for select markets.
- Variants have been identified basis differences in mastheads / presence or absence of supplements and/or price point differences (as reported in ABC or by details as provided by respective publication houses).

### 3) Causes for Variations in Sample Surveys

All estimates based on a sample survey are subject to 'sample variation'. However tightly controlled, the results from one sample of people will differ somewhat from another sample of people drawn in exactly the same way.

Any characteristic (e.g. % owning a cell phone, % reading a newspaper etc.) observed in any sample survey or sample surveys conducted at two different time periods, could show different results. These observed differences can be of two types:

- a) Real change has occurred in the characteristic being measured. Such as cell phone ownership may have gone up.

- b) No real change has actually happened, but the survey shows some differences.

Obviously, there is no issue in the first case. Let us now focus on the second case. The observed differences can be due to many reasons. These can be classified into three broad groups:

- i. First: Sampling Error - Sampling error represents the uncertainty in survey estimates that occurs because we observe data on a sample in the population rather than on every unit of population. Any sample survey, including the best and the largest in the world, will have variations in estimates, simply because it is a sample survey. One can only minimize sampling error by designing samples to provide the most precise estimates at available resources. There is no way to avoid this error other than to conduct a Census. Sampling error is often expressed as standard errors or in simple terms 'Margin of Error' (at a design confidence level, for estimates). The magnitude of Margin of Error depends on the incidence of observed characteristic and the sample size. (Refer below the paragraph on Sampling Variations)
- ii. Second: The survey design (the theoretical parameters such as representativeness and accuracy of the household selection frame, in our case the electoral rolls, etc.) will have a role to play in this.
- iii. Third: Non-Sampling Errors - Lastly, errors creeping in due to non-response; respondent's understanding of questions; wrong or incomplete responses from respondents; respondent selection processes not followed accurately; interviewer mistakes and errors in data punching or processing; etc.

**Our objective is to**

- a) Operate within the ranges defined by globally accepted random sampling methods (i.e. within 'First point' above)
- b) Create a robust theoretical design to minimize errors referenced under 'Second'.

- c) Control processes as much as possible such that errors occurring due to 'Third' are kept to a minimum

**Sampling Variations:**

The level of sampling variation (Margin of Error) is what the survey designers have agreed to accept (indicated by the survey's reporting standard) for any survey estimate. The IRS reporting standards define that the estimates be reported within **20% Margin of Error at 90% confidence level for an incidence of 10%**.

Let's explain a few related and important points on this:

- (i) 90% confidence means that if a survey were to be conducted 100 times, on 90 occasions the variation would be within range defined by the reporting standard. Please note that this means that in 10% cases the estimate may well be beyond the defined range.
- (ii) IRS estimates are not point estimates, but a range estimate and the range depends on the 'Margin of Error' associated with each estimate. And as per IRS design statistics, there are 90% chances that the actual estimate lies within this range (lower and upper confidence limit).
- (iii) IRS reporting standards define that the estimates be reported within 20% Margin of Error at 90% confidence level for an incidence of 10%. Hence, a higher incidence will have a lower Margin of Error and a lower incidence will have a higher Margin of Error.

Please refer below examples for better clarity:

**To understand Margin of Error in sampling and associated estimate range**

- The table below, Table A, shows the upper and lower confidence limits that are associated with Brand A usership estimate at 90% confidence level
- This means that there are 90% chances that the actual user of brand A fall somewhere within these limits: in this example the reported usership estimate is 100,000 users, and there is a 90% chance that the actual number of users lie somewhere between 88,000 and 112,000 users.

**Table A**

<b>Survey Sample Size</b>	<b>Brand A usership Incidence (Vert%)</b>	<b>Associated margin of error</b>	<b>Reported Estimate</b>	<b>Lower Confidence Limit Basis margin of error)</b>	<b>Upper Confidence Limit Basis margin of error)</b>
1000	20%	11.6%	100000	88000	112000

#### **Recommended minimum sample size for analysis**

- Any estimate based on a respondent count of less than 30, does not meet Globally accepted statistical thresholds and hence should not be analysed.
- To avoid any inadvertent misinterpretation of data and for the benefit of IRS data users, IRS technical committee has decided to not report any estimate based on a sample count of less than 30. Hence, any estimate (cell estimate in IRS Clear Decisions) based on a sample count of less than 30 will not get reported/displayed.
- Please note that while conducting any cross tab in Clear Decisions, if the numbers are not visible in any cell, it does not necessarily mean that the cell has zero cell value. Cell estimates which are based on sample count of more than 30 gets reported/ are visible and cells which are based on less than 30 sample count do not get reported/are not visible. All sample counts are getting counted in the overall total.

- For instance, please refer the below Table B; blank cell for Brand E and F do not mean Zero owners of Brand E and F, the data for these cells have not been reported as the sample count is less than 30, however they contribute to total estimate value.

**Table B**

		<b>All India</b>	<b>Market X</b>
Totals	(000s)	257993	782
	Sample Count	59018	541
Brand A	(000s)	36768	344
	Sample Count	9210	253
Brand B	(000s)	132830	271
	Sample Count	27723	193
Brand C	(000s)	20605	76
	Sample Count	5017	40
Brand D	(000s)	31449	45
	Sample Count	8500	30
Brand E	(000s)	3609	-
	Sample Count	972	9
Brand F	(000s)	32732	-
	Sample Count	7596	16