

DEAR LORD, GIVE US THIS DAY OUR DAILY NEWS.... EFFECTS OF RETAIL DISTRIBUTION ON READERSHIP IN RURAL INDIA

Katy Merchant, Indian Market Research Bureau (IMRB)
Ramesh Thadani, Indian Market Research Bureau (IMRB)

Introduction

Rural India comprises nearly half a million villages inhabited by a population of over 728 million - 72% of the country's total population, as at July 1999.

Until recently, this vast consumer market was largely unrepresented in the Media Measurement Studies. This is because any research design for Rural India needs to take into account the wide heterogeneity of the market, which, in research terms, translates into virtually a geographical blanket coverage, and creates issues on the reliability of the findings in terms of sampling methodology and sample sizes.

Some of the factors which impact media reach estimates even within a State are :

- Variances in literacy levels
- Differential economic growth patterns
- Village infrastructure
- Village class.

Apart from these issues, the most pertinent to the readership measurement methodology, is the variation in the retail distribution systems of publications.

In Urban India, the retail distribution systems are well organised and efficient. Consumers either get the publication delivered at their doorstep, or it is picked up by the reader from the kiosk.

As yet, there is no validated data on rural distribution patterns. In most cases, the publications move from the nearest Urban centre either by a bus, or an auto rickshaw, or even a bullock cart! Often there are days when the bus doesn't go to the village at all or days when the last three days' newspapers arrive bundled together. In addition, the fieldwork in a sampled village is likely to be completed in a day and not spread over the entire week, because the sample per village is small.

Therefore, using the Recent Reading Method (RRM) in a scenario with an erratic distribution system could theoretically impact readership estimates - especially newspapers. For example, in a sample village, if the day previous to the interview is one of the days when the newspaper was not delivered, there would be no 'yesterday' readers. On the other hand, if the newspapers had reached the village the day previous to fieldwork, all readers would be Average Issue Readers (AIR).

To sum up, erratic distribution in rural markets could be a major factor impacting readership estimates and also creating significant variances between checks, in a continuous study.

In NRS 1999, we collected information on :

- How newspapers reach the village
- Whether they arrive on the day of the issue or later.

In the readership interview we also collected information on :

- Primary and pass-on readership i.e. Source of copy
- Place of reading i.e. at home or outside.

This paper reports on the magnitude of the erratic distribution patterns, if any, and whether there are variances in readership between villages with 'good' and 'bad' distribution.

If distribution emerges as a factor impacting readership estimates in rural markets, one of the options would be to use Frequency of Reading distribution to model Average Issue Readership.

Methodology

A brief description of the NRS 1999 Round 1 Fieldwork

The NRS 1999 - 2000 is a continuous study comprising of four half-yearly rounds. Round I is now complete with fieldwork conducted between April - June, 1999. The Round I of the NRS 1999 covered 818 towns and 2,058 villages across the Indian continent. The total sample of 1,72,905 individuals comprised of 1,10,351 urban interviews and 62,554 rural interviews.

A stratified multi-stage sampling, procedure was used for the survey to select the towns and villages. At the first level, the administrative districts which constituted a State were grouped together to create homogeneous regions based on :

- Language
- Geographical contiguity
- Financial, economic and administrative homogeneity
- Regionalisation of culture and lifestyle, which make them unique from other districts
- Caste and class considerations.

The group of homogeneous districts are referred to as Socio-Cultural Regions (SCRs).

Within each SCR in a State, a fixed number of villages by population class were selected randomly following the PPS (Population Proportionate to Size) method.

The All India distribution of the villages by population classes for this round of fieldwork is given below :

Village classes	No of villages	Sample size
Population 5000+	309	11,021
Population 2001- 5000	516	18,249
Population 1001- 2000	625	16,907
Population upto 1000	608	16,377
TOTAL	2,058	62,554

Given the high illiteracy levels in Rural India, a totally random sampling procedure for individual interviews would have yielded less than 50% of literate respondents in some villages. This means that we would be measuring a large and varied number of publications within a small literate sample.

In order to obtain reliable readership estimates based on a larger reader sample, a three stage sampling method was developed for the selection of homes and individuals for readership interviews within a village. This is described briefly as follows :

1. Random selection of homes.
2. Collection of household demographics including literacy level of every individual adult (15+ years) person in the household.
3. Random selection (following systematic random sampling procedure), of a fixed number of adults for the Readership Interview within each stratum of individuals - literate and illiterate, based on the household listings at stage one above. The skew in over-sampling and under-sampling by literacy is corrected at the analysis stage, based on the survey data for homes at the aggregate SCR x Village population class level.

The assessment of distribution of publications and the impact on readership

We have selected for this study, four out of the total 25 States reported in Round I of NRS 1999. Their salient population demographics are given as Appendix with this paper.

The States / State groups and the rationale for selection is given below :

1. **Andhra Pradesh (AP)** : High TV and High Cable and Satellite penetration.
: One of the leaders in information technology.
: Opening up for more industrialisation and inviting foreign investments.
2. **Madhya Pradesh (MP)** : Heterogeneity of language and lifestyles across SCRs.
: Primarily agricultural based but also rich in natural resources - iron, coal, other minerals. Yet high variance between SCRs in economic development.
: One pre-dominant language (Hindi) across State.
: Fairly low penetration of TV - especially Cable and Satellite.
3. **North States and Assam (NESA)** : Hilly forests and Tea plantations.
: Heterogeneous population consisting partly of Bengalis and immigrants from surrounding Tibeto - Burmese regions.
: Politically unstable and geographically inaccessible in some regions.
: A greater proportion of population than elsewhere is familiar with English.
4. **Uttar Pradesh (UP)** : This State accounts for 18% of the country's total rural population.
: Geographically too, a large State which could be divided in 3 distinct regions - each different from the other.
: Historically very important, and even in modern times drives the political scene in the country.
: Despite lower than average literacy, it is a fountain head for as many as 7,895 newspapers, with a total circulation of over 30 million copies and 38,607 periodicals.

During the NRS fieldwork, village information was collected on availability of newspapers and on several parameters which affect media exposure and lifestyles. The newspaper availability was checked out for : Frequency and Source / Mode of receipt.

- Frequency**
- a) "Which newspaper, if any, does the village receive ?"
 - b) For each Daily : "Does the day's issue, say today's issue of...(DAILY) generally reach the village on the **same day** or later ?"
- IF LATER** : "When / after what interval does the village receive ...(DAILY) ?"
- (a) Everyday / but the previous day's issue
 - (b) Every alternate day - 2 issues together
 - (c) 2 or 3 times a week - on fixed days
 - (d) 2 or 3 times a week - no fixed days
 - (e) Paper comes only if somebody brings it from next village / town - no fixed pattern
 - (f) Other
- Source / Mode of receipt**
- For each Daily : "Mainly / most times how does the village get ... (DAILY) ?"
- SINGLE ANSWER :**
- (i) From Agent / News Vendor or / Distributor
 - (ii) Visiting hawkers bring paper with their products for selling
 - (iii) Bus drivers passing on route
 - (iv) Village shopkeepers - with other products, bring daily from next town / village
 - (v) Villagers visiting next village / town get / bring the paper
 - (vi) (Other) Visitors to this village get the paper from next village / town
 - (vii) When there is a market / haat (local fair)
 - (viii) Other

Based on this data, the villages within each SCR in a State are classified as those with good, average and poor availability of newspaper. We used the availability of dailies for this categorisation rather than magazines as Recent Reading Method (RRM) has a greater impact on AIR of dailies, by definition.

The conditions for classifying the villages by availability of newspapers are given below :

Villages with :

1. **Good availability / Distribution of dailies** (EITHER : Daily received on the same day of issue, OR Later, but everyday, the previous day's issue)

AND ANY / EITHER OF FOLLOWING MODE OF RECEIPT
 - (i) From Agent / News Vendor or / Distributor
 - (ii) Visiting hawkers bring paper with their products for selling
 - (iii) Bus drivers passing on route
 - (iv) Village shopkeepers - with other products, bring daily from next town / village
2. **Average availability / Distribution of dailies** : All other villages who receive any daily irregularly i.e. not on the day of issue or everyday.
3. **Poor** : All villages where no daily is received. Respondents interviewed in these villages could get a copy of the daily during their visit to a nearby town.

We first examined the effect of distribution on readership estimates based on the Recent Reading Method. That is, we examined whether the Average Issue Readership for the leading dailies and magazines varied between villages with good, average and bad distribution.

In order to determine that the variation in AIR, if any, between the three village categories is not due to differences in the literacy levels, the study looked at the differences in readership, based on :

- All respondents within the three village categories and also
- based on literate respondents within each distribution strata.

The next module of analysis examined the gap between AIR estimates based on Recent Reading Method and the readership derived using the 'Frequency of Reading', that is, modeled on theoretical probability.

In other words, we examined whether in markets with erratic distribution, the Recent Reading Model is equally valid, or whether there is a larger gap in these markets between Recent Reading Model and Frequency of Reading, when compared with the regions where retail distribution is more organised.

The findings are projected to adult population (Sex x Age x Literacy) by SCR for the group of villages clubbed as those with good, average or bad distribution.

The Findings

Classification of villages by distribution levels

Across States, there were significant differences in the distribution structure. A State such as AP had a majority of villages (72%) with regular delivery of publications.

On the other hand, more than half (54%) of the villages in MP did not receive a newspaper daily on the same day. Although UP boasts the largest number of publications, and huge circulations, it still has nearly 50% of its villages, uncovered.

Table 1 : Classification of villages by distribution levels

	AP	MP	NESA	UP
Base : Number of villages covered	159	195	73	352
	100%	100%	100%	100%
% with distribution of Publication:	%	%	%	%
Good	72	28	45	35
Average	16	19	51	18
Poor	12	54	4	47

The larger villages tended to have a better distribution compared to the smaller villages. The range varied by State.

Table 2 : Distribution Levels and Village population size

	Large Village (Population 5001+)	Small Village (Population < 1000)
% of villages with good distribution :	%	%
AP	88	53
NESA	66	38
UP	50	23
% of villages with poor distribution :	%	%
AP	4	29
NESA	-	4
UP	25	60

The impact of erratic distribution is reflected in the lower incidence of primary readership and the increasing readership outside home especially in AP and MP. The relatively small samples of villages with Poor distribution in NESA (2 villages) slightly distorts this picture (in fact, estimates on this segment should be ignored for the rest of the paper).

Table 3 : Distribution Levels and Access to Publications

STATE	DISTRIBUTION LEVELS											
	AP			MP			NESA			UP		
	Good	Avg.	Poor	Good	Avg.	Poor	Good	Avg.	Poor	Good	Avg.	Poor
% Estimated Adults in...	%	%	%	%	%	%	%	%	%	%	%	%
Any Daily : A.I.R.	14	8	8	10	7	4	15	8	10	12	8	6
Place of Reading :												
Home	21	15	10	33	27	17	52	59	58	26	16	10
Outside	79	85	90	67	73	83	48	41	42	74	84	90
Source of Copy :												
Bought	14	10	9	29	29	15	46	52	58	25	18	10
Borrowed	86	90	91	71	71	85	54	48	42	75	82	90
Number of Issues Read	1.59	1.36	1.63	1.71	1.75	1.59	1.78	1.46	1.20	1.58	1.66	1.45

Is Readership lower when distribution is erratic ?

Readership levels overall for both dailies and magazines decline with an average and poor distribution. In States such as AP where a regional daily has a virtual monopoly, the difference is not as marked as in MP where the market is divided between a clutch of Hindi newspapers.

Also, one could argue that villages with 'poor' distribution could be remote and having lower literacy rates. We therefore, analysed the data for the literate universe to demonstrate that the differences are dominantly due to distribution.

Table 4 : AIR of Publications / Any Daily / Any Magazine

	AP		MP		NESA		UP	
Base : All Individuals	All	Lit	All	Lit	All	Lit	All	Lit
Any Daily in villages with distribution levels :	%	%	%	%	%	%	%	%
Good	14.1	31.1	9.6	19.4	15.2	23.3	12.0	22.8
Average	7.9	19.4	7.0	14.6	8.2	12.9	8.0	16.2
Poor	7.5	17.4	3.6	8.3	10.0	15.7	5.5	11.0
Any Magazine in villages with distribution levels :	%	%	%	%	%	%	%	%
Good	10.5	23.3	8.8	17.8	19.0	29.3	11.2	21.2
Average	7.6	18.7	5.9	12.3	15.4	24.3	7.7	15.5
Poor	7.8	18.1	3.4	7.7	10.8	16.8	6.6	13.1
All = Literate + Illiterate								
Lit = Literate								

In Table 5 overleaf, the AIR values for each publication in the 'good' column is taken as 100. The corresponding AIR estimate in 'average' and 'poor' market is shown as Index.

Within a State, readership levels for individual publications decline as we go from good to poor distribution areas.

But, dailies which have a nearby town as the centre of publication or has accessed a poor distribution region, would create a larger AIR for itself compared to the norm. For instance, in NESA, Magazine I.N. has a higher index of 75 than Magazine B.S. in the same region (47).

Similarly, we find that the erosion in readership as we go from good to the poor distribution regions is much slower, among literates.

Table 5 : AIR of Selected Titles By Distribution Levels

	ALL RESPONDENTS			LITERATE RESPONDENTS		
Distribution	Good	Avg.	Poor	Good	Avg.	Poor
AIR %	%	Index	Index	%	Index	Index
AP						
Daily N.	12.2	56	56	27.1	62	59
Magazines :						
Data	3.6	114	75	8.0	125	79
N.D.	1.7	35	29	3.8	37	29
A.S.V.	4.3	19	28	2.7	74	67
MP						
Daily K. K.	4.9	65	27	9.9	67	31
Daily N. T.	4.9	47	31	9.9	48	34
Magazines :						
M.K.	1.1	145	82	2.2	150	91
NESA						
Daily M.O.	6.7	54	31	10.3	54	31
Magazines :						
B.S.	12.1	92	46	18.7	94	47
I.N.	8.6	77	74	13.3	78	75
UP						
Daily D.J.	8.4	71	48	15.9	77	50
Daily P.Q.	5.2	58	29	9.9	62	30
Magazines :						
S.S.	4.5	64	49	8.6	70	52

Is the Recent Reading Model valid in Rural India ?

For this analysis, we compared the readership estimates of selected publications (dailies and magazines) based on two main methods :

- Recent Reading Model** : The audience of an average issue of a publication is measured by establishing how many people have read it in a time interval equal to the publishing interval.
- Frequency Method** : Readership is estimated by allocating an average reading probability, (depending on the frequency scale used), to all the respondents who claim that particular frequency of reading.

We assume that where the distribution is irregular, the difference between the readership estimates based on theoretical probability i.e. Frequency Method and the Recent Reading (publication read with period of publication), would be small. This is because the actual Reading Frequency is likely to be closer to the expected Reading Behaviour i.e. theoretical probability.

On this assumption, we would expect that the gap between estimates based on observed probability i.e. RRM and theoretical probability would be :

- Greater where the distribution is erratic
- Less where the retail setup is organised.

We calculated for selected publications within each good, average, poor category of villages at the State level :

- The difference in A.I.R. based on the two methods
- Indexed the Frequency Method estimate on the RRM estimate

a) RRM estimates for publication 'X' in :

- Good market : 2,205
('000 readers)

b) Frequency Method / Theoretical Probability : 2,862
('000 readers)

c) Index (b/a) : 129.8

d) Similar Index for the publication in :

- Average market: 209.2

The above Index indicates that the difference between the readership estimates in an average market is nearly 65% (209 / 129) higher in the less organised regions compared to the good distribution centres.

For the purpose of this analysis, we have clubbed the poor and the average regions in AP, MP and NESa. For UP, since 47% of the villages have poor press distribution, the balance villages are clubbed together as Good and Average.

The salient findings based on this analysis are :

1. Newspapers tend to have a higher variance in readership in the less developed distribution regions. This suggests a need to consider an alternative to RRM.
2. The differences for magazines are less pronounced , since even in the poor distribution areas, readers would visit a town every 4 to 5 days.
3. Within a State, the indice for individual publications vary by very narrow margins for some (e.g. B.S. in NESa) and three times for others. For example, Daily 'J' readership index in AP is 420 in the Average / Poor sector. On the other hand, Daily 'N' reveals only a 60% variance between RRM and Frequency based estimates, in the same sector.

This also suggests that Daily 'N' has a stronger distribution network even in the less developed regions.

Table 6 : Frequency of Reading Indexed to RRM

	Distribution levels	
	Good	Average / Poor
AP		
Any Daily	106	126
Daily J.	163	420
Daily N.	130	159
Magazines :		
Data	90	84
N.D.	142	363
A.S.V.	166	200
MP		
Any Daily	117	122
Daily K. K.	174	189
Daily N. T.	144	150
Magazines :		
M.K.	74	69
S.T.	47	66
NESA		
Any Daily	108	130
Daily M.O.	165	235
Daily Ad	169	228
Magazines :		
B.S.	84	70
I.N.	97	98
	Good / Average	Poor
UP		
Any Daily	109	120
Daily D.J.	143	156
Daily P.Q.	139	180
Magazines :		
M.K.	100	101
S.S.	105	93

Conclusions

1. Undoubtedly, distribution and readership are inversely correlated, better and regular distribution resulting in higher readership levels. This calls for extreme care in ensuring that the villages selected in a State and Socio-Cultural Region (SCR) at the sampling stage 'truly' represent the distribution patterns of the State. This variable alone could cause readership estimates to swing considerably from one round of NRS to another.

We propose therefore that a large scale exercise be conducted in all States covering atleast a 10% sample of all villages to establish the distribution pattern. An alternative could be to use Census based infrastructure development indicators at the village level and correlate these with the distribution patterns observed in NRS 1999. Perhaps, it would be possible to develop a model which can be used to predict print media distribution patterns at each village level. This would then enable either stratification by distribution patterns or atleast we can ensure that the sample / villages selected represent the universe of villages on this parameter.

2. Undoubtedly, the gap between estimating readership (AIR) using the 'Recent Reading Model' and theoretical probabilities widens as distribution becomes erratic and poor. We therefore recommend that for the next wave of the NRS this be revalidated using an alternative method (e.g. diaries and a short term panel) in some of the sample villages.

A case could also be built for adjusting the readership estimates for villages with poor distribution by using the claimed reading frequency and not the theoretic probabilities but the 'observed' probabilities at each point on the scale in sample villages with good distribution.

However, such an adjustment could be controversial unless we have more empirical evidence and experimental data measuring readership in the same village using alternative methods (e.g. a Diary Panel).

References

- Reliability of Response in Readership Research, Wally Langschmidt
- Predicting Magazine Audience, Helen Johnston, Grey Advertising
1993 Readership Research Symposium 6
- Magazine Sourcing, Peter Walsh
Readership Research : Theory and Practice, Third International Symposium 1985
- Circulation and Readership : Trend Analysis, Ingemar Lindberg
Readership Research : Theory and Practice, First International Symposium 1981
- A Study of overclaiming readership using a Recent Readership Technique, Clark Schiller
Readership Research : Theory and Practice, First International Symposium 1981

APPENDICES**Appendix I- Demographic profile of states reviewed for distribution network of publications in rural India**

All Figures in '000s		AP	MP	NESA	UP
Est. Homes	(Total)	17,090	14,147	5,153	25,370
	(Urban)	4,932	3,476	1,208	5,209
	(Rural)	12,158	10,671	3,945	20,161
Est. Adults	(Total)	49,785	48,620	17,721	1,01,061
	(Urban)	16,014	12,489	4,296	21,726
	(Rural)	33,771	36,131	13,425	79,335

URBAN				
All Figures in '000s	AP	MP	NESA	UP
Urban - All	16,014	12,489	4,296	21,726
% Literate	75.7	76.5	91.3	69.9
Urban - Men	8,086	6,594	2,267	11,743
% Literate	84.5	87.9	95.5	78.4
Urban - Women	7,928	5,895	2,028	9,983
% Literate	66.7	63.7	86.6	59.8

RURAL				
All Figures in '000s	AP	MP	NESA	UP
Rural - All	33,771	36,131	13,425	79,335
% Literate	45.9	47.6	64.1	51.1
Rural - Men	17,092	18,652	7,034	42,224
% Literate	59.3	64.2	74.8	67.6
Rural - Women	16,679	17,479	6,391	37,111
% Literate	32.2	29.8	52.3	32.3

Source : NRS 1999

Appendix II- Demographic profile of states reviewed for distribution network of publications in rural India

All Figures in '000s (Main / Regional Language)	AP (Telugu)	MP (Hindi)	NESA (Assamese)	UP (Hindi)
Exposure to Press (URBAN)				
Any Daily	5,894	4,724	1,536	7,577
Any English Daily	1,295	280	672	638
Any Main Language Daily	5,134	4,667	759	7,395
Any Magazine	3,874	2,981	1,883	5,683
Any English Magazine	1,654	673	1,212	1,222
Any Main Language Magazine	3,041	2,836	817	5,416
Any Publication	6,844	5,446	2,304	9,173
Any English Publication	2,135	765	1,377	1,469
Any Main Language Publication	5,943	5,379	1,063	8,994
Exposure to TV (URBAN)	13,252	15,136	10,086	15,136
Exposure to Radio (URBAN)	3,776	1,731	1,683	4,115
Exposure to Press (RURAL)				
Any Daily	4,287	2,301	1,533	6,767
Any Main Language Daily	4,261	2,286	1,373	6,750
Any Magazine	3,615	1,796	2,261	6,600
Any Main Language Magazine	3,508	1,755	2,024	6,533
Any Publication	5,752	3,261	2,762	10,743
Any Main Language Publication	5,680	3,245	2,507	10,688
Exposure to TV (RURAL)	18,007	14,984	6,421	25,266
Exposure to Radio (RURAL)	9,165	5,182	6,861	18,126
Literacy (URBAN + RURAL)	55.5	55.0	70.7	55.1

Source NRS 1999

