

Output Quality
in a
Survey Organisation
such as
NSSO

SDRD, NSSO

Quality in general and in surveys

There is no doubt that all human progress is the outcome of some human beings' striving for quality.

But products of different human endeavours vary a lot.

A survey organisation's output has certain special features.

Output quality in a survey organisation has certain special problems.

Content

- Definition
- Assessment: special problems
- Assessment: possibilities
- Improvement
- Quality vs. quantity
- Role of competition
- Presentation of output
- Instructions to data users

Definition

- Quality of survey output is easier to define than to actually achieve.
- A survey aims at estimating some broad or summary features of one or more populations.
- Over a short period such as a year, these features can be regarded as unchanging parameters.
- **The quality of any estimate, obviously, is its closeness to the parameter it seeks to estimate.**
- **The problem of quality measurement stems from the fact that the parameters are unknown.**

Non-observability of output quality

- Most things are judged by their performance (leaving out purely decorative products).
- In case of many products, the quality is manifested as soon as they are used.
- But, for other products, it may take time for deficiencies to come to light.
- There is rarely a simple test of goodness of recorded data, or goodness of estimated aggregates or averages.
- So, one may use or apply wrong data to arrive at misleading results and wrong policy implications for a long time, without knowing it.

Assessment of quality: the basic problem

Thus the basic problem of assessment of quality of survey output is its non-observable nature.

Good quality and bad quality, in case of estimates, have no visible features by which they can be recognized.

Quality assessment possibilities - 1

Estimating RSEs of estimates

- ❖ To judge the extent of error due to sampling, Mahalanobis introduced the method of interpenetrating sub-samples (IPNS) to estimate standard errors of the estimates of the target parameters.
- ❖ Relative standard errors (RSEs) are still widely used in NSS to judge the quality of the estimates.

**Relative Standard Errors (RSE) of estimates of Proportion of Ailing Persons (PAP),
Rural, selected States**

State	No. of sample villages	RSE (%) of rural PAP	State	No. of sample villages	RSE (%) of rural PAP
UP	616	4.75	JHK	104	14.47
MAH	340	6.08	TRP	104	13.13
WB	324	5.32	MAN	96	13.37
BHR	264	8.93	PUN	96	5.56
MP	248	8.35	TEL	94	11.08
TN	246	7.08	J&K	92	16.55
ODI	212	7.34	HAR	90	15.64
ASM	212	13.94	HP	88	13.29
RAJ	210	10.80	CTG	85	15.33
KTK	186	9.95	MEG	68	19.22
GUJ	182	7.76	MIZ	48	29.80

Relative Standard Errors (RSE) of estimates of Proportion of Ailing Persons (PAP), Rural, selected States and UTs

State	No. of sample villages	RSE (%) of rural PAP	State	No. of sample villages	RSE (%) of rural PAP
ARP	48	15.44	CHN	8	2.30
NAG	44	56.70	D&NH	8	36.71
UTK	44	17.64	D&D	8	7.27
SIK	40	33.26	LAK	8	27.10
A&N	20	15.50	PUD	8	54.25
GOA	12	19.21	INDIA	4577	1.84

A low RSE reassures us that the estimate is probably not affected appreciably by sampling fluctuations.

But estimates may be affected by other errors, e.g. reporting errors, that have nothing to do with sampling.

Pervasive reporting biases can affect estimates very seriously.

Limitations of RSEs

RSEs cannot be used to detect or measure

- ✓ systematic respondent biases, e.g., general tendencies to under-report expenditures, savings and asset holdings

(these shift the location of the distribution of the estimator without affecting its variability)

- ✓ biases (if they exist) such as deliberate under-enumeration

(which may cause aggregates to be underestimated without affecting estimation of averages)

**Estimates of PAP based on (1) self-reporting (2) proxy reporting
Rural, selected States**

State	PAP estimate based on		ratio	State	PAP estimate based on		ratio
	self	proxy			self	proxy	
UP	101.7	55.0	1.8	JHK	90.7	31.9	2.8
MAH	107.4	66.3	1.6	TRP	40.0	33.5	1.2
WB	217.4	114.8	1.9	MAN	19.9	28.5	0.7
BHR	92.7	42.0	2.2	PUN	245.8	116.6	2.1
MP	89.8	38.4	2.3	TEL	143.9	70.7	2.0
TN	197.4	111.4	1.8	J&K	101.6	49.5	2.1
ODI	128.0	87.1	1.5	HAR	120.4	32.6	3.7
ASM	30.8	32.0	1.0	HP	109.9	70.7	1.6
RAJ	86.2	39.8	2.2	CTG	39.0	40.2	1.0
KTK	126.1	76.7	1.6	MEG	50.5	23.1	2.2
GUJ	136.6	58.4	2.3	MIZ	17.8	28.5	0.6

**Estimates of PAP based on (1) self-reporting (2) proxy reporting
Rural, selected States**

State	PAP estimate based on		ratio	State	PAP estimate based on		ratio
	self	proxy			self	proxy	
ARP	87.3	103.9	0.8	CHN	149.2	92.1	1.6
NAG	14.7	37.2	0.4	D&NH	58.2	53.3	1.1
UTK	119.4	63.1	1.9	D&D	46.0	34.5	1.3
SIK	73.1	18.2	4.0	LAK	150.6	161.9	0.9
A&N	342.7	99.2	3.5	PUD	175.4	175.0	1.0
GOA	184.5	145.1	1.3				

For India as a whole, self-reporting-based estimate of PAP is 147.1 and proxy-reporting-based estimate of PAP is 72.2 (a ratio of 2.0). This exercise shows that true PAP is either underestimated by proxy reporting, or overestimated by self-reporting, or both.

Limitations of RSE estimates

RSE estimates assess sampling errors but cannot detect reporting biases.

In the above example, the magnitude of a special kind of reporting bias was assessed through special tabulation.

This possibility may not exist for other kinds of reporting bias, e.g. deliberate and widespread under-reporting (of savings, asset holdings, jewellery purchases, etc.).

Quality assessment possibilities - 2

Comparison with Census data

- ❖ This has been tried out for such parameters as population and number of households, literacy, employment, etc., where the decennial census, too, gives comparable estimates.
- ❖ However, the Census estimates, too, are subject to various systematic non-sampling errors.
- ❖ Comparison with Census estimates has created a general impression that NSS surveys underestimate population, which may or may not be true.

Quality assessment possibilities - 3

Comparison with administrative data

- ❖ Requires corresponding administrative data to be made available to NSSO so that it may be compiled and compared with NSS estimates
 - NSS estimates of cons. exp. on railway fare/ bus fare can be compared with data available with M/o Railways/ Road Transport authorities
 - NSS estimates of no. of households with bank accounts can be compared with data available with RBI
- ❖ Cannot be used as a general method of assessing NSS data quality

Quality assessment possibilities - 4

Re-surveys on a limited scale

- ❖ Sample households of some sample villages (say) can be re-surveyed at a different time to see if
 - change in informant results in different data
 - survey by a different (supervisory) officer elicits the same data
- ❖ Such re-surveys have great potential, but require the consent of the sample households

Scope for quality improvement

Fortunately, it is possible to bring about improvement of quality even if measurement of quality is difficult.

Improvement of quality of inputs should always bring about some improvement in output quality.

Improving input quality

- ✓ improving a survey frame
- ✓ changing the sample design
- ✓ modifying the data collection process
- ✓ improving follow-up routines
- ✓ changing the processing procedures
- ✓ revising the design of the questionnaire

Improving input quality – 1

Sample design improvements

- Enterprise surveys: Greatest need is for up-to-date first-stage sampling frames
- Household surveys: Recent innovation in tackling FSU size variation (sub-FSU selection by SRS instead of FSU selection by PPS) is expected to improve estimates

Improving input quality – 2

Schedule design improvements

- Adoption of questionnaire format should do away with the need to convey difficult concepts to the informant
- Questions whose answers require tedious recall and research by the informant should not be asked – if necessary, shorter reference periods or as-on-date-of-survey questions may be used

Improving input quality – 3

Modifying the data collection process

- Adoption of e-schedule for data is expected to improve data quality, reduce processing time, which in turn would lead to timely publication of survey results and faster dissemination of unit level data

Improving input quality – 4

Measures for better quality field work

- Continuous monitoring of primary field work, following concurrent inspection system
- Longer and more comprehensive training of Field Investigators in general and survey-specific concepts, our survey instruments should be 100% free from ambiguity
- Training of field workers in how to interact with informants and reassure them that divulging information will not harm their interests
- Creating greater public awareness of NSSO through publicity campaigns this includes impressing upon the respondents about the importance of their participation in the survey

Quantity vs. quality

More is not always better. There may be a trade-off between the two. For example,

- When estimates are required for even the smallest State/UTs, giving adequate sample size to ensure reliable estimates for them will result in smaller sample sizes for the other States.
- Covering too many subjects in the same round will mean too few sample households per FSU for each schedule.
- The effort to generate more frequent estimates (e.g. quarterly estimates) may produce unreliable estimates and thus be self-defeating.

Lack of competition

- Competition is healthy because it keeps up quality.
- But a national survey organisation like NSSO is in a monopolistic position – no one can compete with it in terms of resources.
- This means that NSSO does not get the stimulus for improvement of quality that it would have got if it had to compete with other data producers.

Presentation of output

- The printed report remains a very important output of NSSO.
- Over the years, what may be called packaging has improved a great deal.
- While some emphasis on presentation is good, packaging can be overemphasised.
- No artwork can compensate for badly composed and incoherent text.

Instructions for data users

- The more complex and sophisticated a product, the greater is the danger that it may be improperly used.
- Instructions to the user in the use of the product are correctly regarded nowadays as a part of product quality.
- In NSSO, special care needs to be taken in preparing the instructions for use of unit level data.
- Many users find the estimation procedure difficult to comprehend, and instead need step-by-step worked-out examples.
- Specialised cells can be set up to respond to users' queries, though their duties should not include the checking of calculations made by users.

Summary

- Data quality is not easy to assess.
- All possible ways of checking quality, even if in limited spheres, need to be explored, including comparison with administrative data, and special post-survey investigations.
- **Much needs to be done to reduce respondent bias and unwillingness. This includes creating greater public awareness of NSSO, and better training of field workers.**
- It should be seen that attempts to get more data in a shorter time do not sacrifice quality.
- In preparing survey reports, good finishing should not take precedence over the content. Content should be as lucid as possible.
- In case of unit level data, providing instructions for proper use of the product is part of product quality.

Thank you