The Statistician (1991) 40, pp. 325-331

Attrition and rotation in panel surveys

TREVOR SHAROT

AGB Television, The Research Centre, West Gate, London W3 3HH, UK

Abstract. This paper draws on recent research into these issues in the commercial research sector; specifically, in the measurement of television viewing. Different analyses for detecting measurement errors associated with time-in-sample are described.

1 Introduction

Panel samples offer unique attractions for the measurement of change. Estimates of net change have low sampling variation, while gross change and changes at the individual respondent level may also be measured. Their drawbacks, at least for panels of households or individuals, include low response rates (compared with single-interview surveys), sample attrition and the possibility of conditioning. In such panels, attrition rates vary between demographic subgroups and this can give rise to sample imbalance and bias in estimation. Conditioning also gives rise to bias.

Even in the absence of measurement error, differential rates of attrition demand an active replacement sampling policy if a representative panel is to be maintained. Sample rotation represents one method. An optimal method for introducing rotation is described. However, rotation incurs higher cost and sample turnover rates than a selective discard policy. As a result rotation has seen limited application to date.

2 Why panels?

From a statistician's point of view, the essence of a panel survey is that repeated measurement on a single sample provides the following:

(a) reduced sampling variation in the measurement of change;

(b) the ability to analyse behaviour in general, and change in particular, at the individual respondent level.

For panel operators and data users, there are three further attractions:

(c) the ability to 'train' respondents to perform relatively complex tasks within the data collection process (such as completing a special diary);

(d) the ability to collect a greater range of data than is possible at a single contact;

(e) the spreading of the not insubstantial costs over a long time period and a potentially large user base.

Against these advantages must be set some features of panel research which act against accuracy in estimation:

(a) the initial cooperation rate is lower than for a single contact;

(b) the sample undergoes attrition over time;

(c) there may be response conditioning (e.g. fatigue);

(d) there may be behavioural conditioning.

Panel operators are not complacent about these problems and seek to employ methods which minimise or compensate for these factors. This paper describes some current practices in media research in the commercial research sector.

3 Case-study-the measurement of television viewing

AGB operates in the UK a national television viewing panel of 3000 households, under contract from the Broadcasters' Audience Research Board (BARB), and similar panels are operated in other countries, including Eire, Italy, Holland, Belgium, Greece, Turkey, Portugal, Australia, New Zealand and Thailand.

A brief word about the method of collection of television viewing data is first required. Panel households agree to have an AGB Peoplemeter attached to each of their television sets; the main component of this is an electronic device which sits on top of each television set and detects when the set is switched on or off and the channel tuned. In addition, each resident indicates their viewing sessions by pressing a button allocated to him or her on a remote control handset, once when they start viewing and again when they stop. Their viewing status is displayed for confirmation on an LCD display on the set-top box. The set-top unit sends this information via the 240 V supply line to an electronic memory located inconspicuously somewhere in the home, which is connected to the telephone line. Stored information is retrieved from the memory each night during the small hours by means of a telephone call to or from the host computer (the bell is prevented from ringing). The resulting data are essential information to broadcasters as input to programming policy and programme scheduling, and to broadcasters, advertisers and their agencies for the efficient sale and purchase of commercial air-time.

4 Sampling procedure

Sampling is performed in two phases. In the first phase a face-to-face interview survey (the Establishment Survey) is carried out each year. In the UK, this is based on a sample of approximately 20000 households (successful interviews) using a two-stage stratified cluster sample of addresses from the Postcode Address File. This survey provides certain universe estimates, such as the distribution of number of TV sets per household. It also provides a presurveyed bank of potential recruits for the viewing panel. In the second phase, homes are selected from this address bank for recruitment, selection being made at random within control strata.

Households which move home, out of their PSU, are dropped from the panel and replaced by another.

5 Cooperation and attrition rates

The monitoring process described in Section 2 requires a degree of cooperation from respondents, both because of the need to install the meter and because of the requirement to register viewing on the handset. The average cooperation rate for the UK panel is about 50%, though this varies between different subgroups of the population. Information on non-response is necessarily limited, but most non-response occurs in the second phase, panel recruitment, so the characteristics of those who agree to the survey but not to panel membership are known.

Panel representativeness is also affected by attrition. This occurs for a variety of reasons, at rates in the UK typically as shown in Table 1.

Table 1. Annual rates of panel attrition

House moving	10%
Inability to continue	1%
Fatigue (drop-out)	6%
Fatigue (discard)	2%
Discard for stratification	8%
Total	27%

It should be noted that in some instances, fatigue causes the household to ask to be dropped from the panel, while in other cases monitoring procedures detect unsatisfactory response and the household is discarded. Rates of attrition also vary between different subgroups of the population, as is shown in Table 2.

Table 2. Annual attrition rates by subgroup (indices)

	Drop-outs	Discards
N umber in household		
1	91	70
2	73	99
3	126	103
4	128	108
5+	112	153
Age of individual		
16-24	125	89
25-34	136	124
35-44	121	106
45-54	100	106
55-64	55	98
65+	63	80

The immediate impact of these differences is that even if a new panel could be recruited with 100% cooperation, it would, over time, become increasingly unbalanced demographically. However, there is a need to maintain overall demographic balance to be able to gross up the data to produce audience estimates.

6 Control of panel balance by stratification

The approach used by AGB to balance the panels is to make use of stratification in replacement sampling. Strata are defined by relevant demographic criteria such as household size, as well as by other relevant indicators such as the number of television sets in the home. The actual stratum definitions naturally vary from country to country depending on the sample size and local demographic structure.

Further stratification is performed using a behavioural control. Many years ago we established that the cooperation rate is positively correlated with the amount of television viewed, i.e. light TV viewers exhibit lower response rates then do heavy viewers. This relationship is not accounted for by demographic differences between light and heavy viewers. There are at least two mechanisms at work here: one is that light television viewers include those who are out of the home a lot and therefore unavailable for interview; the second is the tendency for light viewers to believe that their cooperation is less important.

Of course, at the time of recruitment to the panel, the level of viewing of the respondent is not known. It is possible to elicit data on claimed viewing behaviour at the initial survey interview which correlates well with subsequent actual viewing. Further strata are therefore defined within the demographic strata, based on the claimed viewing profile and the viewing panel is controlled to these targets. Other relevant characteristics, such as the number of TV sets owned, also contribute to the control strata.

This stratification for replacement is complemented by a selective discard procedure. Homes are dropped from the panel either to reduce excessive stratum sizes, or because their cooperation in using the Peoplemeter fails to meet certain standards of acceptability, and after attempts to correct this by appropriate feedback have been unsuccessful.

7 Control using forced rotation

It is also possible to employ forced rotation of the panel to control both the demographic imbalances arising through attrition and any fatigue among panel-members. Under rotation, respondents are discarded at a given maximum time-in-sample. In principle, the limit can be of any value, though in the case-studies discussed here the range of interest is between one and five years. The extent to which such rotation reduces demographic imbalance depends on the differentials in attrition rates, as described in Section 4. Similarly, the extent to which rotation reduces fatigue effects depends on the rate that this occurs after joining the panel.

Panel rotation has, however, some disadvantages of its own, arising from the fact that it increases the overall rate of panel turnover. The resulting loss of sample continuity increases the sampling error in estimates of change, to an extent depending on the correlation in the variables of interest over time. While this effect is modest from week-to-week or month-to-month, it can be significant for year-to-year changes given the rates of rotation considered.

Reduced panel continuity also decreases the sample size available for analyses of behaviour at the respondent level over time. For example, much use is made of the viewing panel to assess the cumulative audience size and frequency of viewing to episodic programmes and advertising campaigns. Differential rates of attrition also mean that the continuously reporting subsample available for such analyses may be unrepresentative.

Finally, increased panel turnover implies higher costs, arising from increases in many activities- sampling, interviewing, recruitment, installation and removal of the metering equipment, and so on. Associated with this is an increase in burden on the research contractor, which may distract resources from ensuring the underlying quality of its operations.

8 Experience with rotation

Panel rotation was adopted in 1987 by the AC Nielsen company in the USA, where it operates a national television audience panel of 4000 households. This was in response to evidence such as that produced by NBC Research (Cook, 1988, 1989), which shows that audience estimates from this panel tend to decline with increasing time-in-sample. More recent evidence from CONTAM (the Committee for National Television Audience Measurement in the USA), provides supporting evidence for this phenomenon (CONTAM, 1989).

Current policy is that panel households are retired after two years time-in-sample, though a one-year maximum has been suggested by Cook (1988), reflecting the fact that time-in-sample effects are most vigorous soon after sign-up. In addition to the disadvantages of forced rotation described in Section 7, Cook (1988) points out that in the presence of time-in-sample effects, any change in the time-in-sample profile of the panel will affect estimates; with a new panel it is not possible to perform rotation so as to avoid these, a point expanded in Section 9.

An alternative method currently receiving consideration in the USA is to employ two samples. One would be a panel (which might or might not employ rotation) using a meter that simply measures tuning of the TV with no demand on respondents to indicate their own viewing. The intention is that a higher cooperation rate and lower attrition rate can be achieved, to the benefit of the measurement of the number of households viewing each channel and programme. The other sample would be a rapidly rotating sample capturing both set-tuning and respondents' viewing, possible employing Peoplemeters. Audience estimates from this sample would then be subject to ratio estimation, based on the common measurement of set-tuning, not to reduce sampling error, but to reduce non- response bias and measurement error.

This strategy is similar in sampling terms to that used in the USA prior to 1987 (and in current use there for some local television measurements), under which data derived from two-week diary samples are weighted to a tuning-only panel sample.

9 The UK experience

The analyses cited from the USA have not been carried out in the UK. However, various verification exercises have been carried out on the AGB/BARB viewing panel to ascertain the accuracy of the measurement. The most valuable of these is an annual survey carried out on the panel itself, referred to as a 'coincidental study'. Each panel home is telephoned at some time between 9 a.m. and 9 p.m. The person who answers the telephone is asked to report which

television sets were on when the telephone rang and who was watching each, leaving the 'phone to check if necessary. The time of the call is noted exactly and the responses received are subsequently compared with the data retrieved from the Peoplemeter. The telephone responses are considered to show the 'actual' viewing of the household, though of course these too carry a degree of measurement error. The degree of correspondence typically is as shown in Table 3.

Table 3. Coincidental survey results

	'Actual viewing'	
Button pressed	Yes	No
Yes	50%	5%
No	5%	40%

The audience level recorded in this way (a rating of 55% in this example) is of course much higher than the true level since households with no one at home, and therefore no one watching television, do not respond to the telephone call. Among responding households, however, there is overall about 90% correspondence between the two measures, while the index of 'actual' to recorded viewers is very close to 100%, typically within 1%-3%.

In order to throw more light on the effect of fatigue on audience estimates, this index is also calculated for subgroups defined by time-in-sample. The results obtained, averaged over the last two surveys, are shown in Table 4.

Table 4. Recorded viewing index by time-in-sample

Time-in-sample	
(years)	Index
0-1	103
1-2	100
2-3	102
3-4	94
4-5	100
5+	99

10 The mechanics of panel rotation

In the UK, the current panel of 3000 households is shortly to be replaced by a new panel of 4500 households, using a new version of the AGB Peoplemeter and a different sample design. Other countries are also moving to Peoplemeter panels. If rotation were to be adopted from the outset, how does one reach the steady state situation of a rotating panel with say a three-year limit to time-in-sample? One can clearly not wait until three years have elapsed and then retire all respondents. It is necessary to consider retiring some respondents before three years and others after more than three years, with a view to reaching the steady-state as fast as possible while

(a) not retiring any households earlier than necessary,

(b) avoiding large fluctuations in the mean time-in-sample,

(c) avoiding large fluctuations in the combined rate of panel loss due to both rotation and attrition, both for the sake of data stability and for operational reasons.

In addition, the required rate of rotation for a given maximum time-in-sample will depend on the rate of natural attrition, for if this increases the required rate of rotation decreases. Some thought resulted in the discovery (or possibly rediscovery) of an optimal solution, in the sense that it provides a perfectly constant rate of panel replacement. It can be expressed as follows:

(a) only the initial sample need be subject to 'early retirement';

(b) all replacement homes serve their full three years;

(c) the required rate of rotation is given by

$$d = \frac{r.(1-r)^k}{1-(1-r)^k}$$

where d is the required rate of rotation, k the required maximum time-in-sample and r the natural rate of attrition.

For example, if k = 3 years and r = 0.25 (25 %p.a.) then d = 0.18 (18% p.a.). Hence 43% of the panel need to be replaced each year. The resulting rotation policy is given in Table 5.

Table 5. Three-year rotation policy for a new panel

		Time-in-sample (years)				Sample
Year	Recruits	<1	1-2	2-3	3+	size
0	100	100				100
1	43	43	57			100
2	43	43	33	24		100
3	43	43	33	24	0	100
4	43	43	33	24	0	100

Thus at the end of year 1, 25 of the initial 100 sample have dropped out and another 18 are retired. In year 2, 25% of the remaining 57 (i.e. 15) drop out and 18 are retired, leaving 24. In year 3, 25% of these (i.e. 6) drop out and the remaining 18 are retired. It can also be seen that the panel reaches its steady state distribution of time-in-sample at the end of year 2.

11 Conclusions

Whether or not measurement errors associated with time-in-sample are present, differential rates of attrition are commonplace and alone demand an active replacement policy if a representative panel is to be maintained. Various operating methodologies are employed for television viewing panels in different countries. Sample rotation represents one method, and an optimal method of introducing rotation for new samples has been described. However, rotation incurs higher cost and sample turnover rates than a selective discard policy. It may also be impracticable to rotate the sample fast

enough to deal with fatigue effects. A selective policy requires an ongoing monitor of data quality at respondent level, which in itself has merit.

Thus as yet rotation has seen limited implementation, and success, for the applications described.

Acknowledgement

Permission from the Broadcasters' Audience Research Board (BARB) to reproduce certain material in this paper is gratefully acknowledged.

References

CONTAM (1989) *Nielsen People Meter Review 1988-1989*, New York.
Cook, B. (1988) *Proceedings of the Sixth Annual ARF Research Quality Workshop, New York.*
Cook, B. (1989) Debugging the People Meter, *Proceedings of the Television Research International Symposium, Tarry town, New York* (New York, Communication Options).