MAIL SURVEYS AND NON-REPONSE BIAS[*]
REPORT ON A N.S.W. MAIL SURVEY

J. W. FREEBAIRN
N.S.W. Department of Agriculture, Sydney

All farmers in one stratum of a wheat storage survey who failed to answer a mail questionnaire were personally interviewed. Answers obtained from the mail respondents and from non-respondents are compared to assess the extent of non-response bias. Five suggestions for planning "high return" mail surveys are discussed.

The purpose of this article, which reports on response to a recent mail survey, is to throw more light on the use of this technique by indicating (a) how a high rate of return may be obtained and (b) the extent of non-response bias. It is hoped that this discussion, in conjunction with that by Dillon and Jarrett,[1] will provide a greater understanding of, and confidence in, the use of farm economic mail surveys as a research technique.

Details of Wheat Storage Survey

During August 1965, questionnaires were mailed to a sample of 1,580 wheat farmers in N.S.W. Information sought included the farmer's wheat acreage, yield per harvested acre, harvesting and grain handling method and equipment, the type and capacity of on-farm bulk wheat storage, the costs of delivery delays incurred during the 1964-65 season, and expected "normal" wheat acreage[2] in each of the next five years. No financial data was sought.

A pilot survey of 30 randomly selected farmers was carried out to pre-test the questionnaire and to assess the likely response rate. A stratified random sample design was used in the main survey, the strata giving proportionate representation from the main wheatgrowing regions of the State. Within each stratum survey farmers were selected at random from a list of eligible voters compiled by the Australian Wheat Board for the election of N.S.W. farmer representatives in 1965.

Each survey farmer was sent a six-page questionnaire, a general introductory letter explaining the questionnaire, its objectives and the importance of every farmer completing it, and a stamped addressed envelope in which to return the completed questionnaire. Three weeks later a reminder letter requesting co-operation, a new questionnaire, and a stamped addressed envelope were sent to all non-respondents. This was repeated three weeks later. The questionnaires were printed on coloured

1 Dillon, J. L. and Jarrett, F. J. Response Patterns in Some Australian Farm Economic Mail Surveys. *Aust. J. Agric. Econ.*, Vol. 8, No 1 (June 1964).

2 Farmers were asked: "Weather and other conditions permitting, about how many acres of wheat do you anticipate sowing each year over the next five years?"

stationery with pale green, pink and yellow being used for the first, second and third mailings respectively.

Rate of Response

Sample size and the time rate of response for the seven strata are shown in Table 1. All replies received up to and including three days after the second mailing are included in the first response and similarly for the second response. All subsequent replies are included in the third response.

TABLE 1

Pattern of Response by Strata and Time of Response to Wheat Storage Mail Survey

Aspect of response		I	II	III	IV	V	VI	VII	N.S.W.
Sample size (no.)		230	240	240	190	240	230	210	1,580
Response to:									
First mailing	(%)	27·8	33·3	37·9	33·2	28·8	27·8	36·7	32·2
Second mailing	(%)	35·2	28·8	30·8	32·1	35·4	31·7	22·4	31·0
Third mailing	(%)	10·4	10·4	6·7	8·9	10·4	10·0	10·5	9·6
Total response	(%)	73·4	72·5	75·4	74·2	74·6	69·5	69·6	72·8
Dead wood[a]	(%)	4·3	4·2	5·8	3·7	5·0	3·9	5·2	4·6

[a] Returns from dead letter office and from sample members who were deceased, retired or otherwise inactive or ineligible.

Perusal of Table 1 indicates a worthwhile response in all strata to the two reminder letters. From the number mailed, 45·7 and 16·2 per cent of the second and third mailings were returned respectively, compared with 32·2 per cent from the original mailing. The relative success of the second mailing follows the pattern found by Dillon and Jarrett[3] for N.S.W., although this was not the observed pattern for the whole of Australia.

The purposeful selection of three weeks between each mailing is interesting. The pattern of response encountered in this survey approximated the conclusion drawn by Robinson and Agisim that "by the end of one week the researcher can expect about half of all questionnaires that will come in; in two weeks approximately 90 per cent of the response will be returned; at the end of three weeks almost all (around 98 per cent) who will answer the questionnaire will have sent in their replies".[4] This appeared to be the pattern for the first and second mailings where the rate of returns reached a peak in the second week and fell to negligible levels by the third week. It was confirmed by the third mailing where 45 per cent of the third response was received within a week, 89 per cent within a fortnight, and 94 per cent by the end of three weeks.

Comparison of Responses

Because of the large number of questions asked, and the limitations of space, comparisons shown in Table 2 have been restricted to four

[3] Dillon and Jarrett, *op. cit.*

[4] Robinson, R. A. and Agisim, P. Making Mail Surveys More Reliable. *J. Marketing*, Vol. 15 (April 1951), p. 424.

representative answers. In general, other survey data showed a high degree of correlation with one or a combination of the variables shown.

TABLE 2

Time Pattern of Response by Average Wheat Acreage per Holding, Yield per Acre, Proposed "Normal" Future Wheat Acreage, and Proportion of Holdings which met Delivery Delays in 1964-65

	Strata	First response	Second response	Third response
Average wheat	I	284·5	273·2	292·4
area (acs.)	II	336·9	253·0	315·3
	III	524·7	443·3	505·6
	IV	524·8	358·6	558·9
	V	218·6	189·2	210·6
	VI	298·3	325·3	274·1
	VII	600·6	550·0	516·7
Average yield	I	26·6	28·0	26·7
per harvested	II	26·0	24·8	25·3
acre (bu.)	III	30·7	28·7	31·6
	IV	24·4	24·9	26·5
	V	28·7	25·3	26·3
	VI	28·0	26·4	25·2
	VII	26·3	23·0	23·5
Proposed "normal"	I	108·4	106·8	105·2
acreage relative	II	101·0	105·4	93·2
to 1964-65 (per	III	120·6	132·1	141·0
cent)	IV	156·2	154·8	125·9
	V	135·1	112·5	148·4
	VI	106·2	106·3	101·7
	VII	117·9	114·4	120·8
Proportion of farms	I	83·9	74·6	92·5
with costly	II	74·7	80·0	66·7
delivery delays	III	86·3	79·4	55·0
in 1964-65 (per	IV	77·8	58·1	71·4
cent)	V	95·1	74·3	71·4
	VI	79·2	81·7	73·3
	VII	84·1	78·4	86·2

Differences can be seen in the value of the variables shown in Table 2 for the three response stages. However, in practically every instance the difference is not statistically significant, i.e. it is due to chance or sampling error. For any one variable the difference between response stages is not significant for more than two of the seven strata. Again the values of the differences are positive for some, and negative for other strata.

An important corollary arising from the last paragraph is that no definite pattern exists between the time of response and the type of farm or farmer. This contrasts with the tentative hypothesis advanced by Dillon and Jarrett[5] that larger properties tend to be predominantly slower responders than small holdings.

[5] Dillon and Jarrett, *op. cit.*

Non-Response Bias

During February 1966 all non-repliers in stratum 1 (or their agents) were personally interviewed. Of these 51 farmers, five had never grown any wheat, four had either died or sold out and three had already made replies in conjunction with a sharefarmer or partner. With the exception of one farmer who could not be contacted, the remaining farmers completed the questionnaire with the author's assistance.

How did the answers of mail respondents compare with those given by non-respondents? Some such information is presented in Table 3. In general, data obtained from other questions can be related to one or a combination of the variables shown.

TABLE 3

Comparison Between Mail Respondents, Non-respondents and Total Sample

Characteristic	Respondents	Non-respondents	Total sample
Number analysed	169	38	207
1. Wheat acreage			
Mean (acs.)	282·0	286·0	282·6
Standard deviation			209·1
2. Handling method			
Per cent of holdings using:			
(a) complete bag handling	10·1	13·2	10·6
(b) bag to bulk	13·8	21·0	15·0
(c) bulk with only paddock storage	52·7	55·3	53·1
(d) bulk with some permanent (weather-proof) storage	23·4	10·5	21·3
	100·0	100·0	100·0
3. Average number of days required to harvest a 30 bushel crop in average condition			
Mean	11·84	12·05	11·87
Standard deviation			6·75
4. Average yield per acre (bu.)			
Mean	27·36	23·84	26·74
Standard deviation			8·66
5. Percentage of holdings affected by grain delivery delays 1964-65	82·4	82·1	82·4
6. Percentage of total wheat crop delayed	31·0	38·4	32·3
7. Proposed "normal" acreage relative to 1964-65 (per cent)			
Mean	106·97	99·96	105·7
Standard deviation			36·40

Average wheat acreage per holding (and though not shown, also the distribution of holdings for different levels of wheat acreage) are almost identical, thus refuting the presumption that small scale producers tend

to be in preponderance among non-respondents. Comparisons 2 and 3 of Table 3 relate to the level of capital intensity in the wheat enterprise. Handling methods are listed in ascending order of their capital requirements. On this evidence it would appear that non-respondent holdings are relatively less capital intensive and more labour intensive than respondent holdings. However, another yardstick of capital, the capacity of harvesting equipment measured as the average number of days required to harvest each holding's crop does not substantiate this hypothesis.

If yield per acre be regarded as an approximate measure of managerial ability, it is apparent that the average level of managerial success and efficiency is lower for non-respondents than for respondents. The value of this comparison is, however, limited to the extent that yield is only an approximate measure of managerial efficiency, and secondly it assumes that both sets of farmers are on the same production function, i.e. other factors are comparable. While this assumption may be true in some cases, it does not necessarily hold for all factors.

The remaining comparisons in Table 3 are designed to assess if there exists any set of circumstances which would encourage one set of farmers to respond and not another. Comparisons 5 and 6 do not support any hypothesis that the presence of a "felt" problem (in this case the inadequacy of the wheat storage system in 1964-65) enhances the response. The last comparison suggests that respondents may have a relatively greater interest in the future development of the industry than non-respondents.

Even though the answers of respondents and non-respondents differ, it is apparent from Table 3 that the answers of non-respondents have very little effect on the total sample statistics. But, are the differences that exist between these groups the result of variations in sampling (i.e. chance alone), or are they the result of inherent differences between the respondents and non-respondents themselves? If chance alone caused these differences then it can be assumed that both groups are samples of the same population and that the statistics derived from the respondent group are unbiased estimates of their respective population parameters. If the reverse is true, then the assumption that the answers of respondents are representative of the total mailing becomes a false one.

Chi-square and T tests[6] were made to compare the answers given by respondents, non-respondents and the total sample. Using these tests the probability with which the observed differences would occur due to chance alone are shown in Table 4.

The probability level at or below which differences are considered significant, i.e. due to reasons other than chance, is often set at 0·05. The exact significance level chosen is, however, arbitrary and will vary for the particular parameter and its distribution depending principally on the probability of making a type I or II error, together with the relative costs of making each type of error. For the most part the results in Table 4 indicate that the differences between respondents and non-respondents are attributable to chance alone and not to inherent variations in characteristics of the two sub-samples. Still more satisfactory

[6] For purposes of analysis the total sample is regarded as the population and therefore a T test is appropriate.

TABLE 4

Probability of Observed Difference Between Answers of Respondents, Non-respondents and Total Sample Being Due to Chance

Characteristic	Total sample versus respondents	Total sample versus non-respondents	Respondents versus non-respondents
1. Wheat acreage			
(a) Distribution	>0·95	>0·50	>0·50
(b) Mean	>0·95	>0·95	>0·90
2. Handling method	>0·95	>0·02	<0·01
3. Number of harvesting days	>0·95	>0·70	>0·85
4. Yield per acre	>0·35	>0·04	>0·02
5. Percentage of holdings affected by delivery delays	>0·95	>0·95	>0·95
6. Percentage of wheat crop delayed	>0·70	>0·40	>0·35
7. Anticipated "normal" future annual wheat acreage relative to 1964-65	>0·70	>0·30	>0·25

from the point of view of mail surveys is the high probability with which the observed differences between respondents and the total sample could have occurred by chance alone.

From the study of non-response bias for stratum 1 it seems safe to conclude that if returns to the mail questionnaire are relatively high, the answers from respondents can be used to obtain unbiased estimates of the population parameters from which the sample was drawn.

Why People Do Not Respond

Non-respondents were not asked directly why they failed to reply because it was anticipated that face-saving excuses rather than true reasons would be given. From the attitude of non-respondents and subsequent discussion, the main reasons for non-response were deduced as follows:

32 per cent overlooked answering or were too busy;

26 per cent thought the questionnaire was not applicable to them;

21 per cent had difficulty contacting their sharefarmer who had much of the information;

11 per cent found the questionnaire too difficult;

5 per cent had already filled out the questionnaire; and

5 per cent feared reprisals from the Taxation Office.

From the point of view of conduct of mail surveys the most important finding arising from these reasons is the necessity to initially interest all farmers in the survey and impress upon each farmer the importance and necessity of completing the questionnaire. This can be done through an introductory letter.

Techniques for High Response

To a large extent the success and usefulness of a mail survey depend on the attainment of a high response rate. From this and other studies which have produced high returns, it is possible to formulate certain

rules or suggestions that are helpful in planning "high return" mail surveys.

Pre-testing of the questionnaire is essential. This has a two-fold function which will more than repay the time and cost involved: it provides evidence of the extent of understanding of the questions by respondents; and it gives a rough guide to the likely final response rate. It is here that one finds out which questions are difficult to answer, which ones will produce biased, misdirected, or ambiguous replies, what terms can be misinterpreted, whether the questionnaire is too long, etc.

The principles of questionnaire design, both of individual questions and of format, are adequately discussed elsewhere[7] and need not be reiterated here. One particular point is that in the present survey, space was provided for respondents to add their own comments. Over 90 per cent of the survey farmers did take this opportunity. It is suggested that this practice probably enhances the response rate by adding an element of individuality and personal expression to the questionnaire.

A priori one would anticipate the response rate to decline with increasing size of the questionnaire. Compared with the response in other Australian mail surveys,[8] the return from the six-page wheat storage survey indicates that decreasing returns to questionnaire size may not be relevant up to six pages.

Although coloured paper was used in the two successful Australian economic mail surveys conducted by the author and by Dillon and Jarrett, it is not clear that it necessarily increased the response rate. For example, Bender[9] in experiments with military personnel using six different coloured stationery combinations found no significantly higher returns than by using plain white stationery. However, it is arguable that the sample he used was atypical. On the other hand, Robinson and Agisim in discussing this question summarize the results obtained from a series of experiments by Eastwood that "yellow coloured stock will bring a higher per cent return than any other colour, with pink running second. Generally speaking dark colours rank low in the scale of effectiveness, light colours rank high."[10] While the issue is still inconclusive it would seem worthwhile to use coloured stationery, if only because coloured stationery must stand out relative to ordinary correspondence on a farmer's desk.

The main functions of the accompanying letter are to interest the respondent in the survey and to persuade him to answer the questionnaire. Kimball[11] and Weilbacher and Walsh[12] found no significant response rate difference between a personally addressed letter and a general Dear Sir letter.

Fourth, inclusion of a stamped addressed return envelope greatly enhances the response. Ample evidence supporting the greater effective-

[7] For example: Karmel, P. H. *Applied Statistics for Economists.* Pitman, Melbourne, 1959, Chapter 3.

[8] Summarized in Dillon and Jarrett, *op. cit.*

[9] Bender, D. H. Coloured Stationery in Direct-Mail Advertising. *J. Applied Psych.*, Vol. 25 (October 1961).

[10] Robinson and Agisim, *op. cit.*, p. 419 discussing Eastwood, R. P. *Sales Control by Quantitative Methods.* Columbia Univ. Press, N.Y., 1940, pp. 288-93.

[11] Kimball, A. E. Increasing the Rate of Return in Mail Surveys. *J. Marketing*, Vol. 25 (October 1961).

[12] Weilbacher, W. M. and Walsh, H. R. Mail Questionnaires and the Personalized Letter of Transmittal. *J. Marketing*, Vol. 16 (January 1952).

ness of stamped addressed envelopes as compared to reply-paid envelopes exists.[13]

The most important technique by which returns may be enhanced is through the use of follow-up reminder mailings. Since the most important reasons for non-reply involve misplacing or overlooking the questionnaire,[14] it is necessary to include a questionnaire and stamped addressed envelope with each reminder letter or notice. The apparent high cost per respondent obtained from a third mailing, relative to earlier mailings, may preclude its use. However, by reducing the possibility of non-response bias this consideration may be relegated in importance. Approximately three weeks, and no less than two weeks should elapse between each mailing.

[13] For example, Robinson and Agisim, *op. cit.*, p. 420, report that a return of 73·8 per cent was obtained using stamped envelopes compared to 66·8 per cent from reply-paid envelopes.

[14] See, for example, Robinson and Agisim, *op. cit.*, p. 418.