# AN EXPERIMENTAL ATTEMPT TO REDUCE FIELD COSTS BY LIMITING CALLBACKS AND INCREASING CLUSTER SIZE

Morris Axelrod, Combined Jewish Philanthropies

#### PURPOSE

This study attempts to test some procedures which might effectively reduce the cost of interviewing without increasing the sampling error or non-response error beyond acceptable limits. At the outset may I note, however, that any such savings should not be completely offset by an increase in variance, or by the introduction of non-response error.

Interviewing expenditures loom large on most survey research budgets, and so are a prime target for cost-cutting. Field costs often account for a third or more of the total research budget. This is particularly true when the research entails the use of a strict probability model, in which only the designated person within a selected dwelling unit can be interviewed. Much effort and expense goes into getting such an interview because people vary considerably in the amount of time, and the particular hours they spend at home. It is not unusual for an interviewer to make as many as ten calls without being able to contact the designated respondent; thus, any attempt to approach the ideal of achieving a 100% response rate means correspondingly greater costs as the number of callbacks increase. Any lesser percentage, of course, increases the sampling error and may introduce a non-response bias. Actually, a 100% rate is rarely approached and need not be. Cost and other considerations dictate that we accept less than the perfect rate, and we must take this into account in determining a tolerable level of sampling error. When we go beyond a certain point, however, the marginal cost of additional interviews becomes prohibitively high and the percentage gain in response rate is out of proportion to the reduction in variance. For a sample of 2,000, an increase in response rate from 85% to 86% will reduce the variance by only a small fraction of a percentage point yet these twenty or so interviews will be tremendously costly ones.

Even if we assume optimum performance by the interviewer, there are still two practical means by which we may reduce costs and increase the efficiency of contacting households. One is to increase the number of addresses which are clustered together so that an interviewer's travel time is reduced. Another is to limit the number of callbacks an interviewer is permitted to make at a single address. (Presumably there is always a limitation imposed by calendar or money which translates into fewer calls - either selectively or systematically limited.)

#### CLUSTER SIZE AND EFFICIENCY

In a typical cross-section sample survey, selected addresses are clustered so that the interviewer can contact several on one trip. This increases the probability of his getting an interview on any single trip. But the more easily obtainable interviews tend to be picked off early. This means that clustering diminishes rapidly as more and more callbacks are made. Subsequent trips must then be made to individual addresses rather than to clusters of addresses. Although increasing the cluster size tends to reduce travel and interviewing costs, it also increases the sampling error and may thus negate or vitiate the saving in cost.

#### NUMBER OF CALLBACKS AND EFFICIENCY

Obviously a reduction in the number of callbacks will reduce per/interview field costs, since the marginal cost of each callback beyond three tends to be high. On the other hand, however, fewer calls will result in a lower response rate. This lower rate can lead to t wo undesirable consequences. First, insofar as the reduction in calls excludes persons who are least likely to be at home -- the single, the married without children, the more socially active -- it becomes selective and thus introduces nonresponse error. Second, it has a damaging effect on the variance. The smaller the number of interviews obtained, the larger the variance.

#### THE DESIGN OF THE EXPERIMENT

It was felt that the research design should try to capitalize on the effect of increased cluster size as well as the effect of reduced callbacks in combination, rather than to study the two independently. We made this decision for two reasons: one the relatively small number of sample areas; and two - the normal variability in average cost per interview among interviewers and among Primary Sampling Units. 1! Therefore, two situations were set up as follows:

- A control group of PSUs with the "usual" cluster size and the "usual" number of callbacks.
- (2) An experimental group of PSUs with twice the usual cluster size and a maximum of three calls to be made at an address.

The Survey Research Center's national sample, which at the time consisted of 66 PSUs, was used. This sample was divided into two halves, and was stratified by size of the PSU and the average cost per interview on previous studies. Each procedure was randomly assigned to the two halves. The four largest metropolitan areas were exceptions. There each PSU was split into matched halves, and both procedures were carried out in each half.

The decision to use no more than twice the usual cluster size was based primarily on the desire to minimize the effect on variance and at the same time to reduce the number of clusters to be visited. A second reason stemmed from the suspected "contamination" effect, which may result from preliminary and uncontrolled discussion between respondents who are neighbors. This effect shows up in two ways. It may sensitize the respondent to the subject matter and thus alter his responses. It may also crystallize his incipient resistance to the interview and increase the risk of getting a refusal.

1/Hereafter referred to as PSUs

A maximum of three calls was decided upon because, in a standard study, roughly 85% of all completed interviews are usually obtained by the third call, and the marginal costs are presumed to rise rapidly with the fourth and subsequent calls.

By way of review, then, the first set of PSUs was the control group in which the usual procedure was followed. The University of Michigan's Survey Research Center generally selects clusters of approximately four addresses and seeks a prescribed minimum response rate of 85% with no limit on callbacks, usually achieved by having the interviewer return to all unresolved addresses.

The second set of PSUs was the experimental group. The average cluster size was about eight addresses which is twice the usual size. The interviewer was instructed to make a maximum of three calls at each address which had not been resolved as an interview or a terminal non-interview on the first or second call. Any visit to the dwelling unit -whether or not the visit required a special trip -- was considered to be a call.

The experiment was carried out in the spring of 1961. The vehicle was a national study concerned mainly with economic matters.

#### THE RESULTS

#### EFFECT ON INTERVIEWING COSTS

As indicated earlier, cost reduction was the main concern. If nothing was saved there would be no need to study further the effects of the experimental variables on variance and non-response error.

The experimental group showed an average savings of one hour per interview, as compared with the control group (Table 1). This represents a savings of about 12% in total interviewing costs.

The size of the savings was found to vary with degree of urbanization. The PSUs were divided into three population categories: (1) large metropolitan areas, (2) other PSUs containing an urban place of at least 50,000, and (3) all other PSUs - small towns and rural areas. These will be referred to as "high urbanization", "moderate urbanization", and "low urbanization" areas respectively.

High urbanization areas showed the largest gain -- about 1 3/4 hours per interview. Moderate urbanization areas showed a negligible gain -- just a few minutes per interview -- and low urbanization areas showed a gain of about one hour per interview.

The sizeable gain seems reasonable for high urbanization areas, where, on the average, more calls per interview are necessary to produce an 85% response rate.

As you can see from Table 2, only 65% of the interviews ultimately obtained in these areas are obtained on the first three calls, as compared with 79% for moderate and 88% for low urbanization areas. Thus, any reduction in the number of calls required should have its greatest impact on the high urbanization areas, where such a reduction tends to eliminate the marginal and most costly calls.

The failure to show a large gain for the moderate urbanization areas suggests that the marginal cost of calls may not rise significantly for those areas after the first three. The small effect of clustering in these areas compared with the other two can probably be explained by their differing degrees of "friction of space".

Movement in high urbanization areas is hampered by dense traffic, narrow streets, pedestrian traffic, and such regulatory devices as signal lights, stop signs, and caution signs. Parking space can be obtained only at considerable cost in time or fees. Distances are so great that the interviewer often resorts to public transportation. Under these conditions, of which New York and Los Angeles afford prime examples, clustering has obvious advantages.

In low urbanization areas, which are relatively sparsely populated and rural, the interviewer is not faced with congestion, but is confronted with larger areas and greater distances. (Partly because of the larger areas and partly because of natural barriers such as mountains, rivers, hills, poor roads, and lack of roads, distances are greater to the cluster and within the cluster itself. Sample addresses will fall off the beaten path, away from the main routes. The unreliability of secondary roads and heightened susceptibility to seasonal hazards - mud, fog, snow and ice, floods, washouts -- compound the difficulties. Here also the advantages of clustering are obvious.)

The moderate urbanization areas, on the other hand, consist more typically of middlesize cities in which there is concentration of population without congestion. Sample addresses tend to fall in or near the central city. The traffic flow is smooth and rapid compared with both the metropolitan and rural areas. Such middle-size cities are easy to get around in; one can usually traverse them in ten or fifteen minutes. Since little time is spent in travel, the marginal cost of additional trips is minimal. Thus clustering is not likely to play as significant a part in reducing trip costs.

This view is supported by Table 3, which shows that travel costs are reduced substantially in high urbanization areas, reduced modestly in the low urbanization areas, and not at all reduced in moderate urbanization areas.

Paradoxically, moderate urbanization areas do not have the cheapest interviews, apparently because the first calls do not produce as many interviews as they do in the rural areas. Each call is less costly, however. In other words, it costs less to reach the cluster in a moderate urbanization PSU, but once the cluster has been reached a call is less likely to yield an interview.

#### RESPONSE RATE AND NON-RESPONSE ERROR

The reduction in cost is large enough to warrant further examination of the procedure, especially in regard to the consequence for non-response bias. The response rate for the control group was 10% higher than for the experimental group. This poses the question whether the 10% who have been sacrificed for the saving in field costs differs significantly from the total population.

Table 4 compares several attitudes and background characteristics for the control and experimental groups. The experimental group is composed of a higher per cent of persons who are married and have very young children (Table 4h). This is to be expected since their movement is relatively hampered and they are tied more to the home. Not unrelated is the very slight tendency for families with no children under 18 years of age to be excluded from the sample when calls are limited. (Table 4j) The experimental group is composed of a few more married persons than single persons. (Table 4b) This group may be slightly more selective of rural households (Table 4g) but there is no significant difference in sex or race between the two groups (Tables 4e and 4f). When calls are limited, two-adult households yield interviews relatively more frequently at the expense of one-adult households (Table 4a).

The educational distribution is somewhat weighted in favor of more education for the control group (Table 4d). College graduates apparently are more elusive than those with less education.

Despite the fact that a reduced response rate seems to be slightly more selective of certain types -- married persons, persons with children, persons with less education -the attitudes which were studied, on the whole, do not vary greatly between the two groups. The experimental group may be slightly less well off financially than the control group a year ago (Table 40) and perhaps slightly less affluent (Table 4s). More persons in the experimental group are pessimistic in their attitude toward the car market while more in the control group are unsure (Table 4k). The experimental group also tends more often to see business conditions as worse now compared with a year ago (Table 4n). However, in general, differences are very small, and are significant only by the application of most rigorous standards.

To the extent that there are differences between the two groups which are deemed to be significant, a correction factor is implicit in the control group. The two groups can be combined and all interviews obtained during the fourth or later call can be given a weight of two.

#### CONCLUSION

In general, a limitation on the number of calls, in combination with larger clusters, has effectively reduced the average cost of each interview while introducing what appears to be relatively small non-response error. The results are promising; however, further analysis of the differential effects for subgroups within our population and of the separate effects of varying the cluster size and limiting the number of callbacks is indicated. Equally important is analysis of the impact on variance.

Limiting the number of calls should be avoided whenever the subject matter under study is correlated with socio-economic characteristics which are underrepresented in the interviews completed on earlier calls.

We chose to experiment with a limitation on the number of calls on this study because it is one in a series of similar continuing studies for which we have information about the relationships which do exist among the variables being studied.

Let me summarize briefly. We clearly cannot proceed to limit the number of callbacks on a general basis. There are risks entailed where there is little previous experience with the relationship.

No gross limitation on the number of calls should be attempted unless it is preceded by clear experimental evidence that the biases introduced are minimal and the savings in cost are significant.

# TABLE 1

Difference in Interviewing Cost Per Interview Between Sample Areas Which Differ in Number of Calls Made and Clustering Size

Type of PSU	Difference in Cost Expressed in Minutes (Control minus experimental)	
High Urbanization	103	
Low Urbanization	65	
All Places	60	

# TABLE 2 $\frac{a}{}$

Of all Interviews Taken, Per Cent Completed on First Three Calls, by Degree of Urbanization

#### Type of PSU

High Urbanization	65%
Moderate Urbanization	79%
Low Urbanization	88%

a/ Source: Two large-scale typical national surveys which preceded this experiment.

### TABLE 3

Difference in Travel Costs Per Interview Between **S**ample Areas Which Differ in Number of Calls Made and Cluster Size

Difference in Cost Expressed in Minutes (Control minus experimental)		
49		
-4		
20		
20		

# TABLE 4

# Experimental and Control Group \_\_\_\_\_by Selected Characteristics

		PERCENTAC	GE DISTRIBUTION
a.	NUMBER OF ADULTS IN FAMILY	Control	Experimental
	One	21	17
	Two	63	68
	Three	12	12
	Four or more	4	3
	Not ascertained		*
		100%	100%
b.	MARITAL STATUS OF RESPONDENT		
	Single	9	6
	Married	74	77
	Divorced	2	3
	Widowed	13	11
	Separated	2	3
	-	100%	100%
c.	HOME OWNERSHIP		
	Owns home	65	65
	Pays rent	30	32
	Other	5	3
	Not ascertained	*	
		100%	100%
d.	EDUCATION OF HEAD OF FAMILY		
	Grade School	30	35
	Some High School	17	18
	Completed High School	23	24
	Some College	13	10
	Completed College	15	11
	Not ascertained	2	2
		100%	100%
e.	SEX OF RESPONDENT		
	Male	45	44
	Female	55	56
	Not ascertained		*
		100%	100%

\* Less than one-half of 1%

(Table 4 - continued)

f.	RACE	Control	Experimental
	White	86	87
	Negro	10	9
	Other	1	1
	Not ascertained	3	3
		100%	100%
g.	SIZE OF PLACE		
	Metropolitan areas	12	11
	Cities with 50,000 and over	17	16
	Cities with population 2, 500 - 49, 999	30	29
	Under 2,500	41	44
		100%	100%
h.	LIFE CYCLE		
	Under 45 and Single	5	5
	Under 45, married with no children	5	5
	Married, with children under 5	19	28
	Married, with children 5 to 15	21	16
	Married, with children 15 to 18	5	4
	45 and over, married with no children	23	23
	45 and over, and never married	16	13
	Other (divorced, widowed, separated)	5	6
	Not ascertained	1	1
		100%	100%
i.	AGE OF HEAD OF FAMILY		
	18 - 34	30	23
	35 - 49	30	33
	50 years or older	40	43
	Not ascertained		1
		100%	100%
j.	NUMBER OF CHILDREN UNDER 18 YEARS		
	None	49	46
	One	17	17
	Two or three	26	28
	Four or more	7	8
	Not ascertained	1	1
		100%	100%

\* Less than one-half of 1%

(Table 4 - continued)

k.	ATTITUDE TOWARD CAR MARKET	Control	Experimental
	Good	43	45
	Pro-con	10	12
	Bad	12	16
	Depends	1	*
	Don't know	32	26
	Not ascertained	<u>2</u> 100%	1 100%
1.	EXPECTATIONS OF BUSINESS CONDITIONS A YEAR FROM NOW		
		10	
	Better	49	47
	About the same	38	42
	Worse	6	4
	Don't know or depends	6	6
	Not ascertained		
		100%	100%
m,	EXPECTATIONS OF BUSINESS CONDITIONS DURING NEXT TWELVE MONTHS		
	Cood times	41	40
	Good times	41	40
	Good, with qualifications	21	19
	Pro-con Rad with evalifications	0 E	8
	Bad, with qualifications	9	10
	Dau Danit know, uncontain	15	10
	Not accertained	2	10
	not ascertamen	100%	100%
n.	BUSINESS CONDITIONS NOW COMPARED WITH A YEAR AGO		
	Better	30	25
	About the same	35	35
	Worse	32	36
	Don't know, or depends	3	3
	Not ascertained	*	1
		100%	100%
٥.	FAMILY'S FINANCIAL SITUATION NOW COMPARED WITH A YEAR AGO	-	
	Better	30	29
	Same. pro-con	48	43
	Worse	21	27
	Uncertain	1	 1
	Not ascertained	- *	- *
		100%	100%

<sup>\*</sup> Less than one-half of 1%

(Table 4 -	continue	d)
------------	----------	----

p.	WHETHER GOOD TIME TO BUY LARGE HOUSEHOLD ITEMS	Control	Experimental
	Good	46	45
	Pro-con	13	16
	Bad	16	19
	Uncertain	22	17
	Not ascertained	3	3
		100%	100%
q.	EXPECTATIONS REGARDING FAMILY'S FINANCIAL SITUATION		
	Better off	38	38
	Same	45	45
	Worse off	6	6
	Uncertain	11	11
	Not ascertained	*	*
		100%	100%
r.	WHETHER PLAN TO BUY NEW CAR OR USED CAR		
	New	11	8
	Used	8	9
	Uncertain	1	1
	Not ascertained	*	1
	Inapplicable	80	81
s.	ANNUAL FAMILY INCOME	20070	20075
	Less than \$3000	24	27
	\$3000	20	21
	\$5000	24	26
	\$7500	12	11
	\$10,000 or more	15	12
	Not ascertained	5	3
		100%	100%
	Number of cases	703	577

\* Less than one-half of 1%