Journal of Official Statistics
Vol 4. No. 4, 1988, pp. 319–332

Toward the Development of Optimal Calling Protocols for Telephone Surveys: A Conditional Probabilities Approach

Richard A. Kulka and Michael F. Weeks[1]

Abstract: This study confirms the results of previous research indicating that the chances of finding a respondent at home and conducting a telephone interview on the first call are much better on weekday evenings and on weekends than during weekday daytime hours. By examining call outcomes across a series of three unanswered calls, it was also demonstrated that this general principle applies to second and third "cold" calls as well, while also providing some evidence that repeated calls during the same "optimal" time period may yield diminishing returns. We also see how one might best "recover" from a nonoptimal initial call through more efficient scheduling of second and third calls. However, when optimal and nonoptimal time slots were mixed, better results were obtained when the former were called earlier in the sequence, thereby emphasizing the critical importance of optimal timing for the first call in particular.

Key words: Telephone surveys; call scheduling; contact rates.

1. Introduction

The past decade has seen a steady growth of telephone interviewing and extensive methodological research on this data collection mode (cf. Groves et al. (1988)). Yet, one issue that has received relatively scant attention is the scheduling of calls. The practical significance of this issue is reflected in the survey research literature (cf. Weeks (1988)). For example, Groves and Kahn (1979) reported that an average of 3.4 calls was required per sample telephone number to successfully complete two national household surveys by random-digit-dialing (RDD). Weeks, Jones, Folsom, and Benrud (1980) found in a national face to face household survey that the average probability of finding anyone at home aged 14 and over was only a bit over 50-50 between the hours of 8:00 a.m. and 9:00 p.m. Marcus and Crane (1986, p. 108) observed that "the efficiency of telephone surveys can be significantly improved if interviewers make their calls during times when respondents are most likely to be at home."

[1] Richard A. Kulka is a Chief Scientist and Michael F. Weeks a Senior Survey Specialist in the Center for Survey Research at Research Triangle Institute, P.O. Box 12194, Research Triangle Park, NC 27709, U.S.A.

Acknowledgements: The authors wish to express their appreciation to James L. Knight for his assistance in data processing and to Graham Burkheimer, Jr., Jay R. Levinsohn, and Frederick W. Immerman for their assistance in conceptualizing portions of the analyses presented here.

Research to date on call scheduling has focused almost exclusively on the timing of calls to sample households and the effects of a variety of independent variables on call outcomes, including time of day, day of the week, seasonality, and the type of area surveyed (e.g., urban or rural). Weber and Burt (1972) and Weeks et al. (1980) investigated optimal timing of interviewer visits in personal interview surveys. For telephone surveys, a number of investigators have analyzed and presented call outcome results for a variety of different survey designs and populations (e.g., Falthzik (1972), Rogers (1976), Groves and Kahn (1979), Fitti (1979), Wiseman and McDonald (1979), Vigderhous (1981), Kerin and Peterson (1983), and Warde (1986)). With the exception of Kerin and Peterson (1983), the literature confirms the intuitive hypothesis that the best time to contact an adult respondent by telephone in a cross-sectional household survey is on weekday evenings and weekends, as opposed to weekday daytime slots (i.e., "normal business hours"). However, optimal timing may vary for special populations (e.g., Falthzik (1972)).

An interesting point is that most researchers have focused only on the first call to a household or have lumped all calls together and presented "static" or unconditional probabilities for obtaining a specific call outcome (answer, interview, refusal, etc.). This approach ignores the distinct possibility of the influence of conditional probabilities; that is, changes in probabilities based on the timing of previous (no answer) calls to a given number. To our knowledge nothing is published that evaluates algorithms for a series of calls to a sample case[2].

[2] However, Groves and Robinson (1982) have conducted some unpublished research on this issue.

Weeks, Kulka, and Pierson (1987) were the first to examine both first and second call outcomes, while also distinguishing between "answered" and "interviewed" outcomes in a large national telephone survey of male U.S. Army veterans aged 32–45, conducted in 1985–86. Consistent with the research cited above, that study found that the chances of finding a respondent at home and conducting an interview on the first call were much better on weekday evenings and on weekends than during weekday daytime hours. Some variation, however, was observed in the best times to call for contacting versus interviewing purposes. Weekday evenings were significantly better than all other times (except for Saturday mornings and evenings) for contacting households by telephone. On the other hand, to obtain an actual interview, Sunday afternoons and evenings along with weekday and Saturday evenings were significantly more productive than other calling periods. Moreover, the study found some evidence that the relative superiority of weekday evenings and weekends for successfully making contact on the initial call also applied to second calls made to first-call no answers, although the pattern was somewhat weaker.

While these results suggested improvements in the efficiency of telephone data collection by more optimal scheduling of calls, that study was subject to a number of limitations. First, as is the case with virtually all studies of call scheduling, initial calls were not assigned to different time periods in a strictly random manner. Therefore, the presumption that a more systematic assignment of first calls to those time periods having the highest contact rates would increase the efficiency of calling was speculative in the absence of a field test or experiment.

Fortunately, shortly after completion of the study, an opportunity arose to conduct a field test in connection with a second phase of the same study. A small (415 cases) sample of

U.S. Army male Vietnam era veterans was contacted using a specific three-call scheduling algorithm based on the probabilities of success for various time periods reported in Weeks et al. (1987). In this study initial calls and callbacks to unanswered numbers were scheduled with little regard to potential variation among time slots in their probabilities of achieving a successful contact. By comparison, the results of our field test were fairly dramatic, indicating a significant increase in the proportion of calls answered, not only for the first call (from 62 to 80 %), but for the second and third calls as well. Overall, we saw an almost 25 % reduction in the average number of calls required to contact a household. Thus, although the two samples of veterans are not strictly comparable, the results of this field test provide preliminary support for the assumption that even a crude call scheduling algorithm, based on the general pattern of call outcome probabilities suggested by prior research, may substantially improve the efficiency of calling in a telephone survey relative to a quasi-random assignment of initial calls followed by a scattering of subsequent calls across different time periods.

Second, due to the special nature of the sample of veterans described in Weeks, et al. (1987), it is difficult to assess the extent to which the results observed in these analyses are generalizable to other special populations or to the general household population. Third, the database used in the Weeks et al. (1987) study (approximately 8 450 cases) was too small to permit more than a cursory investigation of the optimal timing of a series of calls. This was unfortunate since a series of calls was precisely the type of data required to evaluate the utility of a "static" versus a "conditional" probability approach to call scheduling. The next step was to conduct an analysis using the same approach followed in Weeks et al. (1987) on a significantly larger database from a national household telephone survey. The goal was to identify several optimal calling algorithms for multiple call attempts, and further assess the potential efficacy or efficiency of a "conditional probabilities" approach to call scheduling.

2. Methods

The basis for this new study was the 1986 Youth Attitude Tracking Study II (YATS II), a nationwide household telephone survey conducted by Research Triangle Institute and Amrigon Enterprises, Inc. between July and November 1986. The target populations for this study were men and women aged 16–24. The first two phases in this three phase data collection involved the screening of almost 190 000 telephone numbers via random-digit-dialing (RDD) to identify working residential numbers. The database for our analyses consisted only of the 78 682 sampled numbers ultimately identified as households. It was necessary to exclude nonresidential and nonworking as well as no-contact numbers because the call record data for these cases were not available or readily accessible. This analysis also focuses only on "answered" rather than "interviewed" outcomes because of the extensive screening for eligibles in this survey as well as the specialized population actually surveyed in its interview phase. Thus, in contrast with the specialized sample used in Weeks et al. (1987), the results in this article are based on a national probability sample of households with telephones, and are thereby generalizable to the general population.

The survey was conducted from a central telephone interviewing facility. Initial calls to respondents were made between 8:00 a.m. and 10:00 p.m. (local time for the respondent) Monday through Friday, between 8:00 a.m. and 5:00 p.m. on Saturday, and between 12:00 p.m. and 10:00 p.m. on Sunday. Initial calls to newly activated cases were assigned

rather arbitrarily to the earliest available time slot during these time periods. Subsequently, "no contact" cases were rotated systematically through successive daytime and evening shifts (even on the same day) until a contact with someone in the household was achieved.[3] Multiple calls were permitted on the same shift if work on other cases was completed prior to the end of the shift.

This approach resulted in the following distribution of initial calls: 14.7 % were made on weekday mornings, 26.0 % on weekday afternoons, 38.1 % on weekday evenings, 10.3 % on Saturdays, and 10.9 % on Sundays. This distribution corresponds roughly to the proportion of total interviewing hours for a given week (approximately 89 hours) represented by each time slot. Weekday evenings received somewhat more than their proportional share of calls, and weekday mornings somewhat less, than would be expected under a strict allocation of cases by time and number of interviewers for each time slot.

In addition, a stratified and nonequal probability sample was drawn so as to represent the geographic areas of 66 Military Entrance Processing Stations (MEPS). The differential sampling weights generated by this design were not used in the analyses presented in this article. These weights were omitted because a preliminary comparison of the weighted and unweighted data indicated no substantial differences for these results.

3. Results

Reflecting a more "optimal" scheduling of calls in YATS II as opposed to phase one of the survey of veterans, the results of the first three calls for YATS II were somewhat better than those presented in Weeks et al. (1987). The percentages of residential numbers answered for the first three calls in YATS II were 65.5, 56.5, and 54.9 resulting in a cumulative contact rate after three calls of 93.2 %. The average number of calls required to find a respondent at home was 1.59.

The first-call answer rates by time of day (one-hour periods) and day of the week for YATS II are presented in Table 1. All time slots presented in this and subsequent tables are local time for the respondent. These data indicate that the probability of finding a respondent on the first call is much higher on weekday evenings (5:00 to 10:00 p.m.) than during the weekday daytime hours (8:00 a.m. to 5:00 p.m.). For example, 17 of the 25 weekday evening one-hour periods (68 %) had an answer rate of 70 % or higher, in comparison to none of the 45 weekday daytime time slots. Contact probabilities for the weekend time slots reported in the table (Saturday daytime and Sunday afternoon and evening) were also better than the weekday daytime hours, though not as good as weekday evenings. Both of these patterns have intuitive appeal and are consistent with the results presented in Weeks et al. (1987) and other reports in the research literature.

Table 2 presents the same data collapsed into seven day-and-time calling periods. Since these periods span four to five hours, they are easier to use in scheduling telephone interviewer workshifts than either the hourly or day of the week data presented in Table 1. As indicated, weekday evenings had the highest answer rate (73 %) followed by the four weekend time slots [62–71 %, with a weighted (by number of calls) mean of 65 %]. The two worst periods are clearly the weekday daytime intervals (59–60 %, with a weighted mean of 59 %). These differences are all statistically significant ($p<.0001$) and are also consistent with the results presented in Weeks et al. (1987).

[3] This is one variation on the "scatter" approach to scheduling callbacks on cases for which earlier calls have produced no contact, as recently described by Weeks (1988).

Table 1. Percentage of first calls answered by time of day and day of week

	First calls answered													
	Monday		Tuesday		Wednesday		Thursday		Friday		Saturday		Sunday	
Time of day	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*
Morning														
8:00– 8:59 a.m.	59	61	62	108	65	95	66	100	68	134	86	42	B	B
9:00– 9:59 a.m.	58	214	64	434	63	609	60	836	62	690	76	470	B	B
10:00–10:59 a.m.	59	346	61	688	61	866	60	1 051	60	945	68	800	B	B
11:00–11:59 a.m.	56	493	58	805	60	1 030	58	1 083	54	895	63	998	B	B
Afternoon														
12:00–12:59 p.m.	61	589	56	793	58	906	57	973	55	718	64	1 080	62	632
1:00– 1:59 p.m.	60	663	57	798	55	992	57	1 115	57	759	64	1 036	61	718
2:00– 2:59 p.m.	57	824	55	950	55	975	56	1 124	56	786	60	1 241	64	830
3:00– 3:59 p.m.	65	806	60	990	58	890	56	945	59	845	63	1 312	60	1 068
4:00– 4:59 p.m.	61	592	65	524	61	609	62	530	57	608	63	1 081	63	1 029
Evening														
5:00– 5:59 p.m.	69	293	70	475	71	340	68	485	63	372	B	B	69	1 101
6:00– 6:59 p.m.	75	568	76	432	73	782	69	838	69	892	B	B	69	1 124
7:00– 7:59 p.m.	75	1 147	72	900	73	1 389	69	1 317	67	1 611	B	B	73	1 110
8:00– 8:59 p.m.	75	2 161	76	1 699	75	2 560	73	2 453	67	2 625	B	B	75	761
9:00– 9:59 p.m.	80	1 106	79	1 047	78	1 518	78	1 359	74	1 417	B	B	82	111
All times	68	9 863	66	10 643	66	13 561	64	14 209	63	13 297	64	8 060	67	8 484

Notes: (1): B = base < 20; percentages are not provided where $N < 20$.

(2): Nonsubstantive outcomes such as "busy," "intercept," etc., have been ignored.

(3): Approximate standard errors may be computed for each cell as SE = $[(P)(1\text{-}P)/N]^{1/2}$.

Table 2. Percentage of first calls answered by time category

Day-and-time period	First calls answered	
	%	N
Weekday morning (8:00–11:59)	60 ± 1	11 483
Weekday afternoon (12:00–4:59)	59 ± 1	20 304
Weekday evening (5:00–9:59)	73 ± 1	29 786
Saturday morning (8:00–11:59)	68 ± 1	2 310
Saturday afternoon (12:00–4:59)	63 ± 1	5 750
Sunday afternoon (12:00–4:59)	62 ± 1	4 277
Sunday evening (5:00–9:59)	71 ± 1	4 207

Notes: (1): Nonsubstantive outcomes such as "busy," "intercept," etc., have been ignored.
(2): Approximate standard errors were computed for each cell as SE = $[(P)(1\text{-}P)/N]^{1/2}$. Approximate 95 % confidence intervals were computed as ±2SE.

While Tables 1 and 2 deal only with first calls, the answer rate for second calls conditional on the time of the first no-answer call was also examined. Because of the small cell sizes, it was necessary to collapse the two Saturday time slots into one and to do the same for Sunday. These analyses indicated that weekday evenings are still the best time to make a second call, regardless of when the first call was made. However, this time period performed least well when the first call was also made on a weekday evening. Likewise, the weekend time slots generally out-performed the weekday daytime slots for second calls, except when the first call was made during the same time period. In fact, for every second call time period, the lowest answer rate for that period was produced when the first call was also made during that period. This suggests that answer rates for second calls are affected by the timing of the first call.

It was also possible to carry these analyses one step further by looking at the answer rate for third calls conditional on the time of the first and second not-answered calls. Although this tabulation had several cells with small sample sizes, the available data were consistent with the observations made with regard to conditional second calls. Weekday evenings consistently out-performed the weekday daytime time slots on the third call regardless of the timing of the first and second calls. However, weekday evenings performed best when both of the earlier call attempts were made during the weekday daytime period, less well when one or the other of the previous calls was also made during a weekday evening, and least well when both of the previous calls were made during a weekday evening. In fact, in the latter circumstance, the weekday daytime periods actually performed almost as well as the weekday evening and weekend periods on the third call.

In order to summarize these data in a more manageable form, we examined all combinations of the five day-and-time periods over a series of three calls. Of the 125 possible combinations, 65 were eliminated due to small sample sizes (at least one of the component time slots had fewer than 20 calls). The remaining 60 three-call algorithms are presented in Table 3, which ranks these various combinations according to three basic criteria: (1)

Table 3. Ranking of three-call algorithms

Calling period			Calls attempted			Percent of calls answered (cumulative %)			Rank by cumulative % answer	Mean number of calls/ answer	Rank by mean calls/ answer	Approx. cost/ answer in USD	Rank by cost	Sum of ranks	Overall rank
First call	Second call	Third call	First call	Second call	Third call	First call	Second call	Third call							
WDE	SUN	SUN	29 786	429	38	73 (73)	54 (88)	63 (95)	4	1.375	1	1.39	10	15	1
SUN	WDE	WDE	8 484	324	20	67 (67)	72 (91)	70 (97)	2	1.494	20	1.24	2	24	2.5
SUN	WDE	WDM	8 484	324	49	67 (67)	72 (91)	73 (98)	1	1.493	19	1.27	4	24	2.5
WDE	WDE	SUN	29 786	1 837	39	73 (73)	59 (89)	56 (95)	9	1.386	3	1.44	14	26	4
WDE	SAT	SUN	29 786	1 385	308	73 (73)	59 (89)	53 (95)	16	1.386	4	1.39	7	27	5
SUN	WDM	WDE	8 484	402	129	67 (67)	58 (86)	69 (96)	3	1.507	22	1.35	6	31	6
WDE	WDE	WDE	29 786	1 837	216	73 (73)	59 (89)	55 (95)	11	1.387	7	1.47	16	34	7
WDE	WDE	SAT	29 786	1 837	137	73 (73)	59 (89)	53 (95)	14	1.387	6	1.45	15	35	8
WDE	WDE	WDA	29 786	1 837	124	73 (73)	59 (89)	52 (95)	15	1.387	5	1.50	17	37	9
WDE	WDE	WDM	29 786	1 837	210	73 (73)	59 (89)	53 (95)	13	1.388	8	1.50	18	39	10.5
WDE	WDA	SAT	29 786	1 615	77	73 (73)	50 (86)	64 (95)	10	1.393	10	1.51	19	39	10.5
WDE	WDM	SUN	29 786	2 520	29	73 (73)	57 (88)	59 (95)	8	1.394	11	1.52	22	41	12.5
WDE	WDA	SUN	29 786	1 615	27	73 (73)	50 (86)	63 (95)	12	1.393	9	1.51	20	41	12.5
SUN	WDE	WDA	8 484	324	24	67 (67)	72 (91)	42 (95)	18	1.495	21	1.27	3	42	14
WDE	SAT	SAT	29 786	1 385	38	73 (73)	59 (89)	42 (94)	29	1.383	2	1.40	12	43	15
WDE	WDM	WDE	29 786	2 520	525	73 (73)	57 (88)	59 (95)	6	1.397	17	1.54	23	46	16
SUN	WDA	WDE	8 484	225	118	67 (67)	49 (83)	65 (94)	22	1.509	25	1.35	5	52	17
WDE	WDM	SAT	29 786	2 520	90	73 (73)	57 (88)	52 (94)	19	1.394	13	1.52	21	53	18
WDE	WDA	WDE	29 786	1 615	431	73 (73)	50 (86)	58 (94)	20	1.396	16	1.54	24	60	19.5
SUN	WDA	WDM	8 484	225	34	67 (67)	49 (83)	65 (94)	23	1.509	24	1.40	13	60	19.5
SUN	WDM	WDA	8 484	402	30	67 (67)	58 (86)	53 (93)	32	1.508	23	1.39	9	64	21
WDE	WDM	WDA	29 786	2 520	169	73 (73)	57 (88)	47 (94)	25	1.396	15	1.58	25	65	22.5
SAT	SUN	SUN	8 625	1 125	45	64 (64)	56 (84)	56 (93)	36	1.584	28	1.18	1	65	22.5
WDE	WDA	WDM	29 786	1 615	130	73 (73)	50 (86)	54 (94)	26	1.394	12	1.58	28	66	24
WDM	WDE	WDE	11 483	1 903	117	60 (60)	68 (87)	64 (95)	5	1.641	29	1.79	35	69	25
WDE	WDM	WDM	29 786	2 520	240	73 (73)	57 (88)	44 (93)	30	1.397	18	1.58	27	75	26

(cont.)

Table 3. Ranking of three-call algorithms (continued)

Calling period			Calls attempted			Percent of calls answered (cumulative %)			Rank by cumulative % answer	Mean number of calls/answer	Rank by mean calls/answer	Approx. cost/answer in USD	Rank by cost	Sum of ranks	Overall rank
First call	Second call	Third call	First call	Second call	Third call	First call	Second call	Third call							
WDE	WDA	WDA	29 786	1 615	127	73 (73)	50 (86)	48 (93)	37	1.395	14	1.58	26	77	27
SUN	WDM	WDM	8 484	402	42	67 (67)	58 (86)	43 (92)	44	1.510	26	1.39	8	78	28.5
WDM	WDE	SUN	11 483	1 903	36	60 (60)	68 (87)	58 (95)	17	1.642	30	1.77	31	78	28.5
WDM	WDA	SUN	11 483	1 002	27	60 (60)	47 (78)	78 (95)	7	1.704	39	1.85	40	86	30
WDM	WDE	SAT	11 483	1 903	100	60 (60)	68 (87)	52 (94)	24	1.644	31	1.78	34	89	31
SUN	WDA	WDA	8 484	225	25	67 (67)	49 (83)	28 (88)	53	1.511	27	1.39	11	91	32
WDA	WDE	SUN	20 304	3 462	55	58 (58)	65 (85)	56 (93)	31	1.700	35	1.78	33	99	33
WDA	WDE	WDE	20 304	3 462	328	58 (58)	65 (85)	57 (94)	28	1.701	38	1.80	36	102	34.5
WDA	WDE	SAT	20 304	3 462	200	58 (58)	65 (85)	56 (93)	33	1.701	37	1.78	32	102	34.5
WDM	WDE	WDA	11 483	1 903	121	60 (60)	68 (87)	47 (93)	35	1.645	32	1.84	37	104	36
WDM	SAT	SUN	11 483	393	41	60 (60)	60 (84)	51 (92)	42	1.675	34	1.71	30	106	37
WDM	WDA	WDE	11 483	1 002	153	60 (60)	47 (78)	71 (94)	27	1.701	36	1.92	49	112	38.5
WDM	WDE	WDM	11 483	1 903	216	60 (60)	68 (87)	41 (92)	41	1.650	33	1.84	38	112	38.5
WDA	WDM	SUN	20 304	2 499	41	58 (58)	49 (78)	73 (94)	21	1.758	54	1.86	42	117	40
WDA	SAT	SUN	20 304	714	60	58 (58)	59 (83)	53 (92)	45	1.728	50	1.70	29	124	41
WDA	WDE	WDA	20 304	3 462	210	58 (58)	65 (85)	42 (91)	47	1.704	40	1.85	39	126	42.5
WDA	WDE	WDM	20 304	3 462	392	58 (58)	65 (85)	47 (92)	43	1.706	42	1.85	41	126	42.5
WDM	WDM	WDE	11 483	1 023	218	60 (60)	44 (77)	70 (93)	34	1.706	43	1.92	50	127	44
WDM	WDA	SAT	11 483	1 002	75	60 (60)	47 (78)	59 (91)	48	1.705	41	1.89	45	134	45
WDA	WDM	SAT	20 304	2 499	143	58 (58)	49 (78)	66 (93)	40	1.757	53	1.88	43	136	46
WDA	WDM	WDE	20 304	2 499	462	58 (58)	49 (78)	66 (93)	38	1.753	51	1.93	51	140	47
WDM	WDM	SAT	11 483	1 023	68	60 (60)	44 (77)	57 (90)	49	1.712	46	1.90	46	141	48
WDA	WDA	WDE	20 304	1 285	279	58 (58)	42 (75)	70 (93)	39	1.753	52	1.93	52	143	49
WDA	WDA	SUN	20 304	1 285	37	58 (58)	42 (75)	65 (91)	46	1.758	55	1.88	44	145	50
WDM	WDM	SUN	11 483	1 023	26	60 (60)	44 (77)	50 (89)	51	1.713	47	1.92	48	146	51

(cont.)

Table 3. Ranking of three-call algorithms (continued)

Calling period			Calls attempted			Percent of calls answered (cumulative %)			Rank by cumulative % answer	Mean number of calls/ answer	Rank by mean calls/ answer	Approx. cost/ answer in USD	Rank by cost	Sum of ranks	Overall rank
First call	Second call	Third call	First call	Second call	Third call	First call	Second call	Third call							
WDM	WDA	WDA	11 483	1 002	117	60 (60)	47 (78)	41 (87)	54	1.710	44	2.03	54	152	52
WDA	WDA	SAT	20 304	1 285	92	58 (58)	42 (75)	59 (90)	50	1.758	56	1.90	47	153	53
WDM	WDA	WDM	11 483	1 002	140	60 (60)	47 (78)	40 (87)	55	1.711	45	2.03	55	155	54
WDM	WDM	WDA	11 483	1 023	106	60 (60)	44 (77)	40 (86)	58	1.717	48	2.04	57	163	55
WDA	WDA	WDM	20 304	1 285	192	58 (58)	42 (75)	51 (88)	52	1.760	57	2.04	58	167	56.5
WDM	WDM	WDM	11 483	1 023	127	60 (60)	44 (77)	34 (85)	59	1.720	49	2.04	59	167	56.5
WDA	WDM	WDA	20 304	2 499	194	58 (58)	49 (78)	39 (87)	56	1.763	59	2.03	53	168	58
WDA	WDM	WDM	20 304	2 499	379	58 (58)	49 (78)	38 (87)	57	1.768	60	2.04	56	173	59
WDA	WDA	WDA	20 304	1 285	115	58 (58)	42 (75)	36 (84)	60	1.762	58	2.06	60	178	60

Notes: (1) Calling period codes are defined as follows:

WDM=Weekday morning (8:00–11:59 a.m.)
WDA=Weekday afternoon (12:00–4:59 p.m.)
WDE=Weekday evening (5:00–9:59 p.m.)
SAT=Saturday (8:00 a.m.–4:59 p.m.)
SUN=Sunday (12:00 noon–9:59 p.m.)

(2) Nonsubstantive outcomes such as "busy," "intercept," etc., have been ignored.

(3) Cost data include the estimated costs of both contacts and noncontacts. Contact costs include three minutes each of interviewer labor, telephone toll charges, and computer charges. Noncontact costs include one minute each of interviewer labor and computer charges, plus a component of telephone toll charges based on the assumption that 10 % of all noncontact calls required an average of 15 seconds of toll charges. It should be noted that interviewer labor charges remain constant across calling periods while telephone toll charges and computer charges vary somewhat and are cheaper in the evenings and weekends than during weekday daytime.

the cumulative proportion of calls answered after three call attempts; (2) the mean number of calls required to obtain an answer; and (3) the approximate cost per answer. Though not yet considered in our analyses, the last criterion is an important factor in evaluating these various combinations. Some time slots clearly cost less than others, due to variable rates for telephone tolls and computer charges related to the use of computer-assisted telephone interviewing, or CATI (see Note 3, Table 3).

In the table, the algorithms are ranked from 1 to 60 based on the sum of their rank scores on the three separate criteria. For example, ranked first is the "weekday evening/Sunday/Sunday" algorithm, which ranked fourth on the basis of cumulative proportion answered, first on average number of calls to obtain an answer, and tenth on cost per answer. This gives us an overall sum of ranks of 15, the lowest (best) rank of any three-call combination. While we recognize that this approach gives equal weight to the three criteria, the full data provided in Table 3 would allow a reordering of this list in several different ways should a different weighting or priority scheme be judged more appropriate for a particular survey.

A cursory review of Table 3 indicates that the algorithms which emphasize weekday evenings and weekends cluster near the top of the rankings, while those which emphasize weekday daytime slots tend to cluster near the bottom. The ranges on the three criteria run from 84 % to 98 % on the cumulative proportion answered, 1.38 to 1.77 on the average number of calls to obtain an answer, and from $1.24 to $2.06 on cost per answer obtained. In Table 4 these same algorithms are sorted by time of the first call. Since it may often be necessary to schedule initial calls at nonoptimal times due to study time constraints and limitations in facility capacity, this table offers some guidance on how best to "recover" in scheduling the second and third calls. However, the predominant influence of the timing of the first call is readily apparent through an examination of the means provided for each criterion on each set of algorithms.

Table 4. Three-call algorithms ordered by time of first call

Calling period			Cumulative % answered	Mean calls per answer	Approximate cost per answer in USD	Overall rank
First call	Second call	Third call				
WDM	WDE	WDE	95	1.641	1.79	25
	WDE	SUN	95	1.642	1.77	28.5
	WDA	SUN	95	1.704	1.85	30
	WDE	SAT	94	1.644	1.78	31
	WDE	WDA	93	1.645	1.84	36
	SAT	SUN	92	1.675	1.71	37
	WDA	WDE	94	1.701	1.92	38.5
	WDE	WDM	92	1.650	1.84	38.5
	WDM	WDE	93	1.706	1.92	44
	WDA	SAT	91	1.705	1.89	45
	WDM	SAT	90	1.712	1.90	48
	WDM	SUN	89	1.713	1.92	51
	WDA	WDA	87	1.710	2.03	52
	WDA	WDM	87	1.711	2.03	54
	WDM	WDA	86	1.717	2.04	55
	WDM	WDM	85	1.720	2.04	56.5
	All with first call on WDM		91	1.687	1.89	42

Table 4. Three-call algorithms ordered by time of first call (continued)

Calling period			Cumulative % answered	Mean calls per answer	Approximate cost per answer in USD	Overall rank
First call	Second call	Third call				
WDA	WDE	SUN	93	1.700	1.78	33
	WDE	WDE	94	1.701	1.80	34.5
	WDE	SAT	93	1.701	1.78	34.5
	WDM	SUN	94	1.758	1.86	40
	SAT	SUN	92	1.728	1.70	41
	WDE	WDA	91	1.704	1.85	42.5
	WDE	WDM	92	1.706	1.85	42.5
	WDM	SAT	93	1.757	1.88	46
	WDM	WDE	93	1.753	1.93	47
	WDA	WDE	93	1.753	1.93	49
	WDA	SUN	91	1.758	1.88	50
	WDA	SAT	90	1.758	1.90	53
	WDA	WDM	88	1.760	2.04	56.5
	WDM	WDA	87	1.763	2.03	58
	WDM	WDM	87	1.768	2.04	59
	WDA	WDA	84	1.762	2.06	60
	All with first call on WDA		91	1.739	1.89	47
WDE	SUN	SUN	95	1.375	1.39	1
	WDE	SUN	95	1.386	1.44	4
	SAT	SUN	95	1.386	1.39	5
	WDE	WDE	95	1.387	1.47	7
	WDE	SAT	95	1.387	1.45	8
	WDE	WDA	95	1.387	1.50	9
	WDE	WDM	95	1.388	1.50	10.5
	WDA	SAT	95	1.393	1.51	10.5
	WDM	SUN	95	1.394	1.52	12.5
	WDA	SUN	95	1.393	1.51	12.5
	SAT	SAT	94	1.383	1.40	15
	WDM	WDE	95	1.397	1.54	16
	WDM	SAT	94	1.394	1.52	18
	WDA	WDE	94	1.396	1.54	19.5
	WDM	WDA	94	1.396	1.58	22.5
	WDA	WDM	94	1.394	1.58	24
	WDM	WDM	93	1.397	1.58	26
	WDA	WDA	93	1.395	1.58	27
	All with first call on WDE		95	1.390	1.50	14
SAT	SUN	SUN	93	1.584	1.18	22.5
	All with first call on SAT		93	1.584	1.18	23
SUN	WDE	WDE	97	1.494	$1.24	2.5
	WDE	WDM	98	1.493	1.27	2.5
	WDM	WDE	96	1.507	1.35	6
	WDE	WDA	95	1.495	1.27	14
	WDA	WDE	94	1.509	1.35	17
	WDA	WDM	94	1.509	1.40	19.5
	WDM	WDA	93	1.508	1.39	21
	WDM	WDM	92	1.510	1.39	28.5
	WDA	WDA	88	1.511	1.39	32
	All with first call on SUN		94	1.504	$1.34	16

Notes: See footnotes to Table 3.

Table 5. Comparison of groups of three-call algorithms by optimal – and – nonoptimal call mix

Composition of group			Number of algorithms	Mean results			
First call	Second call	Third call		Number of calls	Cumulative per cent answered	Number of calls per answer	Approximate cost per call in USD
OPT	OPT	OPT	8	25 853	95	1.423	1.37
OPT	OPT	NON	4	20 317	95	1.441	1.39
OPT	NON	OPT	8	26 268	95	1.423	1.48
OPT	NON	NON	8	20 425	93	1.453	1.49
NON	OPT	OPT	8	18 161	94	1.679	1.76
NON	OPT	NON	4	18 811	92	1.676	1.85
NON	NON	OPT	12	17 481	92	1.731	1.90
NON	NON	NON	8	17 517	86	1.739	2.04

Notes: (1) OPT = Optimal calling periods (weekday evenings, Saturday and Sunday).
NON = Nonoptimal calling periods (weekday mornings and weekday afternoons).
(2) See footnotes to Table 3.

Finally, Table 5 compares groups of algorithms based on their use of time slots shown to be "optimal" (i.e., weekday evenings and weekends) versus those shown to be "nonoptimal" (i.e., weekday daytime). In general, this way of organizing the data confirms that calling households at optimal times improves calling efficiency, and that when optimal and nonoptimal time slots are mixed, better results are obtained when the optimal time slots are used earlier in the sequence. The particular importance of the timing of the first call is emphasized by the set of algorithms which involved an "optimal/nonoptimal/nonoptimal" sequence which was superior to the set which involved a nonoptimal/optimal/optimal sequence on two of the three criteria.

4. Summary and Conclusion

Overall, the results of this study are consistent with the preponderance of research literature, both for face to face and telephone interview surveys. Our study confirms the results of other studies that indicate that "the chances of obtaining an answer and conducting an interview on the first call are much better on weekday evenings and on weekends than they are during weekday daytime hours" (Weeks et al. (1987, p. 547)). Our study intended to move beyond this general conclusion, and did yield some new information of value when deriving optimal algorithms or protocols for the scheduling of no contact calls in telephone surveys.

Our most significant result was the thorough examination of call outcomes of up to three unanswered calls. We were able to do this because our database consisted of 78 000 cases. It was possible to examine the extent to which the probabilities of obtaining an answer for a second or third call are conditional on when the preceding call was made and, potentially, to identify optimal call scheduling algorithms for a series of "cold" calls. Analyses of the results of initial calls from this database were consistent with the results cited previously regarding the general superiority of weekday evenings and, to a lesser extent, weekends relative to weekday daytime for the scheduling of first calls. Moreover, analyses of second and third calls indicated that this finding also applies to second calls

made to first-call no-answers and to third calls made to first and second call no-answers. However, a notable exception to this general rule was also found. Even the most optimal time periods (i.e., evenings and weekends) tend to perform least well when earlier call attempts were made during the same time period, thereby providing some evidence for the importance of conditional probabilities.

To make better use of this multiple call data, all combinations of time slots over a series of three calls were examined, and 60 of the 125 possible combinations or algorithms were ranked according to: (a) the cumulative percent answered after three calls; (b) the mean number of calls to obtain an answer; and (c) the approximate cost per answered call. Based on an equally-weighted sum of ranks derived from these three criteria, the 60 algorithms were ordered from a high to low on call efficiency, indicating once again that algorithms that emphasize weekday evenings and weekends tend to cluster towards the top, while those that emphasize weekday daytime calling periods cluster near the bottom of these rankings. An alternative sorting of these rankings by time of the first call provided some indication of how one might best "recover" from a nonoptimal initial call through a more efficient scheduling of second and third calls. Similarly, a clustering of these algorithms according to their relative use of optimal time periods provided additional confirmation that calling households at optimal times clearly improves performance. When optimal and nonoptimal time slots are mixed, better results are obtained when the former are called earlier in the sequence. These and other data served to emphasize the critical importance of optimal timing for the first call in particular.

The data presented in Tables 3–5 should be of practical relevance and utility to survey practitioners who develop optimal calling protocols for both initial calls and "cold callbacks" (i.e., where no answer is received on prior calls). As noted by Weeks (1988), "calling protocols are the rules and algorithms that specify the calling efforts to be made on each type of case," and he distinguishes five basic approaches employed to schedule cold callbacks, ranging from "call every successive shift" or a "scatter" approach (spreading calls across some set of defined time periods) to more sophisticated protocols based on "conditional probabilities" or dynamic "priority scores." The data presented here lend support to those who emphasize the potential benefits of employing versions of the more complex calling protocols. Some of the evidence provided is clearly consistent with the logic of a "conditional probabilities" approach, which "requires that answer rate probabilities be reassessed after each successive call, since they may be conditional on the timing of previous no-answer calls (Weeks (1988)). At the same time, Tables 3–5 highlight the potential utility of even a relatively simple "priority score" approach, which may also include a variety of other factors relevant to the survey at hand. With the almost universal adoption of CATI systems to conduct telephone surveys, the implementation of these complex protocols has become increasingly straightforward (cf. Weeks (1988)).

Overall, the study presented here supports the approaches already being employed by several organizations, while suggesting directions for future research. More research is needed to refine and develop the algorithms and approaches suggested here. Special populations and survey conditions will undoubtedly require tailored rules and algorithms. Nevertheless, the increasing amount of research in this area speaks for the potential benefits yielded by the increasing rationalization of the call scheduling process.

5. References

Falthzik, A.M. (1972): When to Make Telephone Interviews. Journal of Marketing Research, 9, pp. 451–452.

Fitti, J.E. (1979): Some Results From the Telephone Health Interview System. American Statistical Association, Proceedings of the Section on Survey Research Methods, pp. 63–72.

Groves, R.M., Biemer, P.P., Lyberg, L.E., Massey, J.T., Nicholls, W.L., II, and Waksberg, J. (Eds.) (1988): Telephone Survey Methodology. John Wiley and Sons, New York.

Groves, R.M. and Kahn, R.L. (1979): Surveys by Telephone: A National Comparison With Personal Interviews. Academic Press, New York.

Groves, R.M. and Robinson, D. (1982): Final Report on Callback Algorithms on CATI Systems. Survey Research Center, University of Michigan, Ann Arbor.

Kerin, R.A. and Peterson, R.A. (1983): Scheduling Telephone Interviews: Lessons From 250 000 Dialings. Journal of Advertising Research, 23, pp. 41–47.

Marcus, A.C. and Crane, L.A. (1986): Telephone Surveys in Public Health Research. Medical Care, 24, pp. 97–112.

Rogers, T.F. (1976): Interviews by Telephone and in Person. Public Opinion Quarterly, 40, pp. 51–65.

Vigderhous, G. (1981): Scheduling Telephone Interviews: A Study of Seasonal Patterns. Public Opinion Quarterly, 45, pp. 250–259.

Warde, W. (1986): Problems With Telephone Surveys. U.S. Department of Agriculture, NASS Staff Report No. SRB-NERS-86-01, Washington, D.C.

Weber, D. and Burt, R.C. (1972): Who's Home When. U.S. Bureau of the Census, Washington, D.C.

Weeks, M.F. (1988): Call Scheduling With CATI: Current Capabilities and Methods. In Telephone Survey Methodology, edited by R.M. Groves, P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicholls, II, and J. Waksberg. John Wiley and Sons, New York.

Weeks, M.F., Jones, B.L., Folsom, R.E., Jr., and Benrud, C.H. (1980): Optimal Times to Contact Sample Households. Public Opinion Quarterly, 44, pp. 101–114.

Weeks, M.F., Kulka, R.A., and Pierson, S.A. (1987): Optimal Call Scheduling for a Telephone Survey. Public Opinion Quarterly, 51, pp. 540–549.

Wiseman, F. and McDonald, P. (1979): Noncontact and Refusal Rates in Consumer Telephone Surveys. Journal of Marketing Research, 16, pp. 478–484.

Received April 1988
Revised December 1988