Journal of Official Statistics, Vol. 15, No. 2, 1999, pp. 199–216

Interviewers' Calling Strategies on Face-to-Face Interview Surveys

Susan Purdon, Pamela Campanelli, and Patrick Sturgis[1]

In July 1995, Social and Community Planning Research (SCPR) began a programme of research to take an in-depth look at the role of interviewers in the survey nonresponse process. The programme is divided into three sub-projects. This article describes the analysis undertaken for Sub-Project 2 which deals mainly with the noncontact component of nonresponse. The analysis is based on call record data from the random half of the 1995/96 Family Resources Survey (FRS) undertaken by SCPR. The particular aim of the analysis is to examine how day of week and time of day of calling affect (a) the probability of contact and (b) the probability of subsequently achieving a productive interview. Combining these estimates with information on costs and survey designs, definitions of efficient calling strategies can, in principle, be developed.

Key words: Interviewer effects; noncontact rates; survey nonresponse; discrete hazard rate logistic regression.

1. Background

In probability sample-based surveys of the general population a primary objective is to maximise response rates for minimum cost.[2] To achieve this objective on interview surveys requires good information on the means or strategies by which interviewers gain responses from individuals, and the costs associated with each of these strategies. In principle, all interviewers can then be briefed on the most efficient strategy.

The role of the interviewer can conveniently be divided into two parts: making initial contact with the potential respondents, and actually achieving an interview with that person or persons. Lack of success in these endeavours can either result in a final outcome which is a noncontact or one which is a refusal, respectively.

This article focuses mainly on the issue of making initial contact. Details about how interviewers handle the actual persuasion of the respondent during the doorstep interaction are focused on in a separate article (see Campanelli, Sturgis, and Purdon, 1997). Both this article and the doorstep article were inspired by a special 18-month programme of research conducted by SCPR into the role of the interviewer in the survey nonresponse process. The programme was funded by a grant from the UK Economic and Social Research Council and has been in co-operation with the ESRC Centre on Micro-Social Change at the University of Essex and NOP Research.

[1] Survey Methods Centre at SCPR, 35 Northampton Square, London, EC1V OAX, UK. Correspondence should be addressed to Susan Purdon. E-mail: s.purdon@mds.qmw.ac.uk

[2] The underlying assumption is that reducing nonresponse will reduce nonresponse bias.

Acknowledgments: The support of the Economic and Social Research Council (ESRC) is gratefully acknowledged. The work was funded by ESRC award number: R000235776.

Other articles have also looked at the issue of making initial contact, detailing how the timing of calls to an address influences the probability of contact. Weeks, Jones, Folsom, and Benrud (1980) found, for example, that in a 1975 U.S. survey conducted in the spring, the chances of finding at least one household member at home improved significantly if calls were made in the late afternoon or evening. In the UK, Swires-Hennessy and Drake (1992) found in 1986 that ''the highest probability of a successful outcome was between 17.00 and 22.00 hours.'' See also Weber and Burt (1972), Vigderhouse (1981), and Kulka and Weeks (1988). Data on hours of ''wakeful occupancy'' from time use studies can also be used to determine probability of contact (cf. Hill, 1978). Knowledge about the best times to call has also been fed into the development of optimal call scheduling algorithms for telephone surveys (see, for example, Groves and Robinson, 1983, Weeks, Kulka and Pierson, 1987), but face-to-face surveys still often rely on a basic suggestion to interviewers to call at different times of day and on different days of the week. Lievesley (1986) found, however, that interviewers with the lowest non-contact rates were not calling more frequently than other interviewers, but were calling at times when people were more likely to be at home.

In this article we describe the development of a model that predicts the probability of making initial contact at an address. This is discussed in terms of the factors that interviewers have control over. As opposed to the numerous options available to the interviewer once contact has been made (see Morton-Williams, 1993; Groves, Cialdini and Couper, 1992), there are basically only three options available to the interviewer aiming to make initial contact. The interviewer can vary the time of day or day of the week he/she calls, how many times to call before giving up, and how much time to leave between calls. The model can be used to predict, for any particular interviewer strategy, the expected number of calls needed to make initial contact. These issues are discussed in Section 2. In principle, armed with an estimate of the number of calls which will need to be made, together with cost information, the total costs of different calling strategies can be evaluated. This is discussed further in Section 3. Although this article focuses on making initial contact, rather than doorstep persuasion, we can still check whether the time of day or day of the week of initial contact affects the final outcome at an address. This is considered in Section 4.

2. Making Initial Contact with Householders

2.1. The data

The analysis is based upon data for the random half of the 1995/96 financial year Family Resources Survey (FRS) conducted by SCPR on behalf of the UK Department for Social Security. The FRS is a continuing survey of households implemented via computer-assisted personal interviewing. The detailed financial interview covers topics such as income, housing costs, child care responsibilities, pensions and state benefits in considerable detail. The survey takes place in a selection of 1,752 sampling points each year covering all of Great Britain, spread systematically throughout the year on a monthly basis with a random half of the sampling points being allocated to SCPR. The fieldwork period for interviewers for each monthly assignment is essentially a month although some

time allowance is made for any remaining hard-to-reach respondents. For the purpose of the analysis reported here, a ''positive response'' is defined as a full or partial interview with any adult member of the household.

Call record data from the FRS is available on over 17,000 addresses (after excluding vacant and nonresidential addresses, re-issues, and unusable records) and over 61,000 individual calls. There are several advantages to using the FRS to study interviewers' calling patterns. It is an easily available source of a large number of call records as, unlike many surveys, the call records are keyed by interviewers as part of their regular administrative tasks. There is also a relatively large number of calls made per address before an address is coded as a ''noncontact''[3], 6.6 on average, compared to the minimum of 4 required of SCPR's interviewers. A slight disadvantage from the analysis standpoint is that the FRS therefore has a low noncontact rate (just 4%[4]).

2.2. *Assumptions behind the analysis*

The outputs of most interest are the estimates of probabilities of contact for individual calls. However, if these estimates are to be unbiased then we must make an assumption that the interviewer is acting as if he/she knows nothing about the household at the first call, and that at subsequent calls the only extra information available is that the household could not be contacted at one or several previous calls. If this assumption is incorrect, which is in reality very probable, then the estimates of probabilities of contact made are likely to be somewhat too large. For example, when an interviewer calls on a weekday morning it may be because he/she has reason to believe the householder does not work full-time. Assuming that the interviewer is more often right than wrong in any such belief, the probability of contact achieved by the interviewer will be higher than it would be had the interviewer simply called at a random selection of addresses on weekday mornings.

2.3. *Controlling contact rates – the role of the interviewer*

In this section we consider the role that the interviewer plays in controlling contact rates.

The factors affecting the contact rate for any particular interviewer's assignment can be divided into two categories: (a) those within his/her control and (b) those outside of his/her control. The latter include factors such as the nature of the area (for example, householders are less likely to be found at home if they live in urban areas) and the time of year. The factors over which the interviewer *does* have control are:

(i) the day and the time of day at which calls are made
(ii) the number of calls made
(iii) the length of time between calls.

Interviewers have to make decisions about the timing of calls in the light of several

[3] An address is coded as a ''noncontact'' if no-one was at home whenever the interviewer called.
[4] This noncontact rate is slightly higher than that published for the FRS. An address has been defined as a non-contact in this article if the *first* interviewer to whom the address is issued fails to make contact. In practice any such addresses will usually be re-issued to a second interviewer who may subsequently make contact. All data on re-issues has been excluded from the analysis reported here.

considerations. Firstly they must decide at what time of day or week the first call to an address should be made. Previous research (Weeks et al. 1990; Swires-Hennessy and Drake 1992) suggests that the evening is optimal. However, since there are fewer available hours in the evening than in the daytime, making all first calls in the evenings may actually increase the number of times an interviewer must visit a sampling point, if not the total number of calls, to reduce non-contact levels to acceptable levels. Furthermore, interviewers often prefer not to make first calls in the dark, and some interviewers do not like to work in the evenings or during weekends.

Having made a decision about when the first calls should be made, the interviewer must then make decisions about the timing of any necessary second and subsequent calls. SCPR's own interviewer manual is reasonably clear on this issue:

''An address or a selected person cannot be coded as a ''noncontact'' until at least four calls have been made at that address. These must be at different times of the day, including at least one evening call, and on different days of the week including at least one call at the weekend.''

Nevertheless some interviewers may conclude that only in the case of the third or fourth call should evening or weekend calls be resorted to. Others may choose to ignore the implied instruction that at least one call should be made during a weekday morning or afternoon and choose to make all their calls during the evening or at the weekend.

A possibly more important decision is whether or not to call again at an address at all. On the surface the most rational approach would appear to be to call at each address a minimum number of times (the minimum expected being four). However, there may be reasons why interviewers choose to concentrate their calls on just a subset of previous noncontacts. For example, some addresses may be more accessible than others.

Finally, interviewers must decide upon the length of time to leave between calls. If the interviewer has reason to believe that the household is away from home for a few days then it makes sense for him/her to wait for a week or so before recalling. However, if there is any doubt about whether or not the household is away it may be sensible to call earlier because otherwise the total number of calls that can be made within the available fieldwork period may be substantially reduced.

2.4. Estimated probabilities of contact by time of call, number of call and time between calls

Table 1 shows the estimated probabilities of contact at the first call by time of day and day of week. The figures largely confirm the findings of Weeks (1980) and Swires-Hennessy and Drake (1992), with the largest probabilities of contact being associated with weekday and weekend evenings. The probability of contact is particularly high on Sunday, Monday and Tuesday evenings. Weekday mornings and afternoons appear to be particularly poor times to call.

Tables 2 and 3 show, respectively, the estimated probabilities of contact at the second and third calls, in both cases excluding all cases where contact was made at a previous call. The probabilities are given as conditional on the previous call time. Because of small

Table 1. Estimated probabilities of contact on first calls, by time of call

		Estimated probability of contact	Sample size
Saturday	Morning	0.50	381
	Afternoon	0.53	1,089
	Evening	0.61	98
Sunday	Morning	0.57	121
	Afternoon	0.60	459
	Evening	0.65	65
Monday	Morning	0.45	654
	Afternoon	0.50	2,216
	Evening	0.66	549
Tuesday	Morning	0.45	702
	Afternoon	0.52	2,124
	Evening	0.67	497
Wednesday	Morning	0.47	659
	Afternoon	0.52	2,125
	Evening	0.63	401
Thursday	Morning	0.48	639
	Afternoon	0.50	1,912
	Evening	0.64	376
Friday	Morning	0.47	582
	Afternoon	0.48	1,807
	Evening	0.59	336

Table 2. Estimated probability of contact at second call

Second call	First call					
	Saturday	Sunday	Weekday morning	Weekday afternoon	Weekday evening	Overall
Saturday	0.38	–	0.50	0.44	0.41	0.45
Sunday	0.49	–	0.51	0.51	0.38	0.48
Weekday morning	0.31	0.44	0.37	0.30	0.26	0.32
Weekday afternoon	0.40	0.39	0.38	0.37	0.38	0.38
Weekday evening	0.54	0.52	0.54	0.56	0.48	0.54

'–' indicates cells with sample sizes of less than 30.

Table 3. Estimated probability of contact at third call

Third call	Second call					
	Saturday	Sunday	Weekday morning	Weekday afternoon	Weekday evening	Overall
Saturday	0.35	–	0.40	0.39	0.26	0.36
Sunday	0.44	–	0.34	0.43	0.30	0.40
Weekday morning	0.32	0.26	0.20	0.33	0.29	0.29
Weekday afternoon	0.27	0.29	0.31	0.30	0.26	0.29
Weekday evening	0.52	0.37	0.50	0.52	0.42	0.47

'–' indicates cells with sample sizes of less than 30.

Table 4. Contact rate at each call

Call No.	1	2	3	4	5
Proportion contacted	0.52	0.42	0.38	0.33	0.31
Sample size	17,792	8,529	4,881	2,988	1,922
Call No.	6	7	8	9	
Proportion contacted	0.27	0.25	0.22	0.18	
Sample size	1,211	799	507	318	

sample sizes the time of day categories have been collapsed to just three. A number of observations can be made:

- irrespective of the time of the previous call the best time to call is always a weekday evening. Almost without exception calling on a weekday morning or afternoon is a sub-optimal strategy;
- the probability of contact reduces as the number of calls increases – in other words, if a householder has failed to be in on two previous occasions then he/she is also likely to be out at the third call.

The latter point is confirmed in Table 4 which gives the overall (or average) probability of contact at each call. The probabilities show a very marked decline as the number of calls increases even though interviewers are more likely to call in the evening on later calls.

Tables 2 and 3 demonstrate the dependence of contact probabilities both on the time of the current call and on the time of the previous call. In theory it is possible that the dependency may be more complex than this, with probabilities being conditional upon the times of *all* previous calls. Small sample sizes mean that this cannot be tested using cross-tabular outputs. However, an informal examination of the data suggests that the rules noted above still hold. For example, if we examine those addresses where the first two calls were noncontacts, then the optimal time for the third call still appears to be a weekday evening, irrespective of when the first two calls were made, even though the sample numbers for many combinations are far too small to test anything but very general hypotheses. These findings are consistent with the estimates of conditional probabilities of contact calculated on a large U.S. telephone survey (Kulka and Weeks 1988).

Table 5 gives estimates of the probability of contact at the second call by the length of time since the previous call (in days). These estimates suggest that the highest probabilities of contact are associated with either calling on the same day as the first call, or after an interval of two weeks or more. The interpretation of these estimates is somewhat ambiguous however, since, for example, second calls on the same day as the first call must by definition be made later in the day than the first call and hence are more likely to be evening calls. And calls within the same week will include a proportion of addresses where the residents are away from home for a period, whereas we might assume that if a second call is not made for at least two weeks then this is because the interviewer has obtained information that the household is away in that period.

2.5. Modelling the probability of contact

The probability of contact at a particular call, assuming noncontact prior to that call, can be modelled as a discrete hazard rate logistic regression model (see Blossfield, Hemerle, Mayer 1989). The hazard rate is defined as

$$\lambda(t|x) = P(T = t|T \geq t, x) = \frac{\exp(\beta_{0t} + x'\beta)}{1 + \exp(\beta_{0t} + x'\beta)}; \qquad t = 1, \ldots, q$$

where t is the call number. In words, the hazard rate represents the probability that contact will be made at time (or call) t given that no contact has occurred before time (or call) t.

Possible predictors of contact at call t include those factors that interviewers have control over, namely:

- time of day/week of the current call
- time of day/week of the previous call(s)
- number of previous calls $(= t - 1)$
- length of time since the previous call.

In recognition of the fact that time of year and the nature of the area may affect probabilities of contact, the following independent variables were also tested in the model:

- month
- area classifiers (e.g., population density, percentage of flats in area, percentage in non-manual occupations, percentage single person households, percentage owning a car – in total around 15 different area classifiers, based on the 1991 UK Census of Population, were tested in the model).

Finally, although with no prior hypothesis about their operation, two interviewer classifiers were added: grade and sex, these being the two classifiers readily available for all interviewers working on the survey. We would not in principle expect the characteristics of the interviewer to be a significant factor in the model since, although interviewer characteristics may influence the times when an interviewer chooses to attempt to make contact, once that decision has been made whether or not contact is made is out of the hands of the interviewer.

Table 5. Estimated probability of contact at second call by number of days since first call

Number of days	P (contact)	Sample size
0	0.50	863
1	0.42	1,952
2	0.42	1,189
3	0.43	994
4	0.40	759
5	0.40	613
6	0.41	529
7	0.39	425
8–13	0.41	941
14+	0.51	258

Discrete hazard rate logistic regression models can be fitted easily within any package offering logistic regression if every call is set up as a separate record. This approach, akin to ''episode splitting'' in event history analysis, allows for the use of time-dependent covariates, namely in this case present and previous call times[5]. The model was fitted in SPSS using forward selection procedures.

Table 6 gives the full details of the final model, the main points of which can be summarised as follows:

1) The odds of a contact at a particular call are dependent upon
   - the time of the current call
   - the time of the immediately preceding call
   - the number of previous calls
   - the month
   - the nature of the area (in terms of population density, % in non-manual occupations, % aged 65 and over, % living alone, % living in detached houses)
   - interviewer grade.
2) In terms of the current call, the odds of making contact with a householder are highest if the call is made on a Monday or Tuesday evening, although the odds are higher than average for any evening. The odds of contact are particularly low for weekday mornings and afternoons.
3) If the previous call was made on a Saturday evening, or at any time on a Sunday, then the odds of making a contact at the current call are significantly lower than average. (In other words, being out at any of these times marks a household as being ''difficult to contact.'') It is worth noting that no significant previous call/present call interaction was found. This, if correct, has implications for how much interviewers can control their non-contact rates by manipulation of time of day call patterns (see Section 3). The timing of the current calls and the timing of the most recent calls are the two most important predictors in the model of whether or not contact is made.
4) Although tested in the model, the timing of calls prior to the most recent call was not found to significantly aid the prediction of contact rates.
5) Contact becomes less likely as the number of calls increases (every extra call after the first is associated with a 17% reduction in the odds of contact).
6) Contacts are more difficult to make at some times of the year. April, May and September are particularly difficult months, whereas January is significantly easier than average. Perhaps surprisingly, no significant holiday effect was found in July/August.
7) It is more difficult to make contact with householders in areas of high population density, in areas where a high proportion of residents live alone, and in areas where there is a high proportion of persons in nonmanual occupations. In contrast, in areas

[5] In terms of event history analysis the single ''event'' is a successful contact, and the single episode is the length of time (measured as the number of calls) until the contact occurs. Some of the covariates, such as time of present call, are time-dependent in the sense that they differ each time a call is made. Maximum likelihood estimation of the hazard rate is achieved by splitting episodes at the points in time (i.e., calls) when change in the covariates occurs and then fitting a time-constant covariate model to the sub-episodic data. (See, for example, Blossfeld, Hamerle, Mayer (1989).)

Table 6. Discrete hazard rate logistic regression model for contacts

		β	S.E.
Time of current call[a]	Saturday morning	−0.16	0.06
	Saturday afternoon	−0.10	0.04
	Saturday evening	0.10	0.11
	Sunday morning	0.10	0.10
	Sunday afternoon	0.06	0.06
	Sunday evening	0.18	0.14
	Weekday morning	−0.50	0.04
	Weekday afternoon	−0.35	0.03
	Monday or Tuesday evening	0.41	0.04
	Other weekday evening	0.25	0.04
Time of previous call[a]	Saturday morning	0.18	0.08
	Saturday afternoon	0.12	0.06
	Saturday evening	−0.27	0.17
	Sunday morning	−0.07	0.16
	Sunday afternoon	−0.07	0.09
	Sunday evening	−0.31	0.22
	Weekday morning	0.21	0.05
	Weekday afternoon	0.20	0.04
	Monday or Tuesday evening	−0.02	0.06
	Other weekday evening	0.02	0.05
First call	Yes	0.51	0.05
	No	0	–
Number of previous calls		−0.18	0.009
Month	January	0.11	0.04
	February	0.03	0.04
	March	0.03	0.04
	April	−0.07	0.03
	May	−0.07	0.03
	June	−0.00	0.04
	July	0.01	0.03
	August	0.03	0.04
	September	−0.10	0.03
	October	0.07	0.04
	November	−0.02	0.03
	December	−0.03	0.06
% in area[b] in nonmanual occupations		−0.004	0.0009
% persons in area aged 65 and over		−0.010	0.0026
% one person households in area		−0.009	0.0024
% detached houses in area		0.003	0.0009
Population density of area (1000 persons/sq.km)		−0.016	0.005
Interviewer grade[c]	A	−0.05	0.03
	B	−0.03	0.02
	C	−0.05	0.02
	D	0.13	0.03
	Supervisor	0.00	0.02
CONSTANT		0.03	0.08

[a]Mornings are defined as up to 12.00, afternoons from 12.00 to 5.00.

[b]An area is defined as the postcode sector (the primary sampling unit for the FRS).

[c]SCPR interviewers are graded A, B, C, D and supervisor, the least experienced interviewers being graded as A.

with a higher than average number of elderly people making contact is easier than average.

8) Somewhat inexplicably, the model suggests that the grade of the SCPR interviewer making the call may influence the probability of contact on the FRS even after the timing of calls and area characteristics have been taken into account – interviewers just below supervisor grade being more likely to make contact. However, this result should probably not be taken at face value since relatively few interviewers are of this grade and the significant result may simply reflect some inadequacy in the model in controlling for, for instance, area effects.
9) The length of time between calls was not found to be a significant predictor of contact rates (even after transformation to a quadratic).

3. Implications of the Model

The model described above can be used to explore the implications (in term of contact rates) of a number of different scenarios. As a simple illustrative example we might envisage a theoretical situation where an interviewer has split his/her first calls equally between weekend, weekday daytime and weekday evening calls and for the second calls intends to make a similar split. The interviewer might legitimately ask whether, for example, the evening calls at the second call should be made predominantly at those addresses that were previously called on at the weekend, or at those previously called on on a weekday daytime, or at those previously called on in the evening. Or would it in fact be better to distribute the calls evenly across all previous call times? The aim, of course, being to maximise the number of contacts at the second call.

Because the two predictors of contact, ''current call time'' and ''previous call time,'' act independently in the model (and because a constant increase in the log odds leads to a roughly constant increase in the probabilities for the range of probabilities we are considering), the model suggests that it actually makes no real difference how interviewers distribute their calls. It appears, then, that the range of control interviewers have over their contact rates is somewhat limited, decisions only really being necessary on:

(a) what proportion of calls should be made on weekday evenings, at weekends, and on weekday mornings or afternoons, with weekday evening calls being the most profitable; and
(b) for individual dwellings, whether an additional call should be made at all.

From the point of view of the analyst, perhaps the most useful consequence of the fact that the list of factors that can be controlled by interviewers is very limited is that the model can be used to predict how many calls on average an interviewer will need to make to get the noncontact rate down to some specified level. Just one piece of information is needed: namely, at each call number, what proportion of calls will be made at each time of day and each day of the week.

To illustrate this we might imagine that interviewers are divided into two extreme groups, the first group being interviewers who consistently make 90% of their calls on weekday evenings and the rest of their calls on weekday mornings and afternoons, and

the second group being interviewers who consistently make just 10% of their calls on weekday evenings and the rest of their calls on weekday mornings and afternoons. Under this very simplistic scenario, to achieve a noncontact rate of 10% the model suggests that the first group would have to call at addresses a minimum of three times before declaring the address a ''noncontact,'' whereas the second group would have to call five times. This is a large difference, but perhaps not as large as might have been anticipated.

3.1. Interviewer calling strategies on the FRS

In practice interviewers would rarely use such extreme calling strategies. Nevertheless, between these extremes there is, on the FRS at least, some variation in the calling strategies that interviewers use. The logistic regression model described in Section 2.5 provides a simple and convenient way to discriminate between calling strategies. The categorisation can be made as detailed as is thought desirable, but for our purposes we chose to divide the 224 interviewers who worked on the FRS in 1995/96 into just four groups based upon their calling times over the first four calls. These four groups appear to discriminate reasonably well between interviewers, with relatively little variation in behaviour *within* the groups.

The four groups defined have been labelled AA, AB, BA and BB. Interviews fall into the AA category if they use a calling strategy that gives an ''above average'' overall probability of contact on their first and second calls and on their third and fourth calls[6]. Relative to this, groups AB, BA and BB are defined as follows:

Interviewer group	First and second calls (probability of contact)	Third and fourth calls (probability of contact)
AA	High	High
AB	High	Low
BA	Low	High
BB	Low	Low

The different styles of the four groups can be very clearly seen by comparing the times of their calls over the first four calls. Table 7 gives the details. Interviewers in Group AA begin with a higher than average proportion of evening calls yet gradually increase this proportion as the number of calls increases. Those in Group BA, in contrast, start with relatively few evening calls but rapidly build this proportion up in later calls. Those in Group BB start with very few evening calls, and although they increase this proportion over time, they do not do so to the same extent as the BAs.

The overall noncontact rate on the 1995/96 FRS was 4% (see Footnote[4]). For each of the four groups to achieve this would require the number and type of calls to be made shown in Table 8 (under the simplifying assumption that interviewers make an equal number of calls to each address before declaring it a ''noncontact'')[7].

[6] An ''above average'' overall probability of contact on the first and second calls is defined by standardising the average probability of contact for each call in turn and then taking the average of these two standardised values. Those ''above average'' are those above the median value for all interviewers. Similarly for the third and fourth calls.

[7] It is worth noting that four calls per noncontact, which is the minimum specified in SCPR's interviewer manual, would give an average noncontact rate of around 11%.

Table 7. Distribution of calls by interviewer calling strategy

Interviewer group	Percentage of calls made:	Call number			
		1	2	3	4
AA	Weekday evenings	20	40	51	55
	Weekday mornings and afternoons	66	40	29	24
	Weekends	14	20	20	21
Sample size		5,641	2,549	1,352	512
AB	Weekday evenings	17	34	31	41
	Weekday mornings and afternoons	65	41	46	36
	Weekends	19	25	23	23
Sample size		3,728	1,737	956	395
BA	Weekday evenings	4	16	43	58
	Weekday mornings and afternoons	89	72	39	26
	Weekends	7	12	18	16
Sample size		4,008	2,058	1,246	435
BB	Weekday evenings	5	10	21	34
	Weekday mornings and afternoons	85	73	55	41
	Weekends	10	17	24	25
Sample size		4,339	2,167	1,327	580

The calculations suggest that to achieve a noncontact rate of 4%, a BB interviewer must make around 11% more calls than an AA interviewer, and 1.3 additional calls to each non-contact, but make less than half the number of evening calls. The other strategies lie between these extremes.

It is not clear why interviewers choose one strategy over another. In terms of interviewer characteristics, male interviewers on the FRS were more likely to adopt the AA and AB strategies than were female interviewers, which suggests that men may be more comfortable than are women in making a large number of evening calls. However, it was also the case that SCPR's less experienced interviewers were more likely to adopt strategy AB than were their more experienced interviewers, which may suggest that as they gain experience, interviewers begin to change their calling strategy.

3.2. Assessing the cost of different call strategies

What is apparent from Table 8 is that interviewers who prefer not to make a large proportion of their calls in the evening can still achieve the same contact rates as other interviewers if they are prepared to call more often. There would seem to be no reason

Table 8. Expected number of calls required to achieve a noncontact rate of 4%, by interviewer calling strategy

Interviewer strategy	Expected number of calls made per address	Percentage of calls made on weekday evenings	Expected number of calls made to each of the 4% of non-contacts
AA	2.13	35	6.8
AB	2.21	27	7.5
BA	2.30	22	7.6
BB	2.37	15	8.1

to divert interviewers from this strategy unless such a strategy significantly increases survey costs.

The amount of cost incurred in achieving a pre-specified noncontact rate under different interviewer calling strategies will depend upon the payment system in operation at a given organisation. For organisations that pay by the interview the difference in costs by strategy will be a function of travel costs only, of which the largest component will be travel to the primary sampling unit rather than travel within it. For organisations that pay by the hour the difference in costs will be a function of travel costs *and* the time taken to make contact (which will depend on the number of calls made). In neither case is it clear, a priori, which interviewer strategy will lead to the lowest cost: interviewers making a large number of evening calls have to make fewer calls to achieve a pre-specified non-contact rate, but constraining interviewers to working within just a few hours a day may mean they have to make more journeys to the sample point, which in turn will push up travel costs. The ideal would be a reasonably large number of evening calls but not to the extent that travel costs are adversely affected. The most efficient interviewer calling strategy is likely to differ from organisation to organisation.

A very simplistic model of survey costs will serve to illustrate the approach in principle. Suppose, for example, that the main determinant of total survey cost (excluding interview fees) is travel costs to and from sampling points. Then the most efficient interviewer strategy would be the strategy that minimised the number of visits made to a sampling point. If it is assumed that eight calls can be made during a single daytime visit (one per hour between 9 a.m. and 5 p.m.) and four during an evening visit, and it is assumed that an interviewer will not make both daytime and evening calls on a single visit, then, based on the figures of Table 8, an interviewer following strategy AA would have to make

$$2.13n\left(\frac{0.35}{4}+\frac{0.65}{8}\right)$$

visits to the sampling point (where n is the number of addresses in the sample) to achieve a noncontact rate of 4%.

An interviewer following strategy BB would, in contrast, have to make

$$2.37n\left(\frac{0.15}{4}+\frac{0.85}{8}\right)$$

visits. Taking the ratio of the two, it follows that an interviewer following strategy AA would need to make 5% more visits to the sampling point than an interviewer following strategy BB. So, under this model, BB is the more efficient strategy.

If, more realistically, it is assumed that travel costs within an area contribute to total survey costs, then the BB strategy will only be more efficient than the AA strategy whenever the following inequality is satisfied:

$$2.37\left(\frac{0.15(1+3x)}{4}+\frac{0.85(1+7x)}{8}\right) < 2.13\left(\frac{0.35(1+3x)}{4}+\frac{0.65(1+7x)}{8}\right)$$

where x is the average cost of travel between addresses within sampling points written as a proportion of the cost of travel to and from the sampling point. The inequality is satisfied

whenever $x < 0.07$. So, under this model, strategy BB is more efficient than strategy AA whenever the average cost of travelling between addresses within sampling points is less than 7% of the cost of travel to and from the sampling point.

To make this model more realistic, account would have to be taken of factors such as the effect on costs of making appointments with respondents rather than carrying out immediate interviews. Some modification would also be needed for organisations who pay interviewers by the hour. Complicating factors, such as the fact that interviewers will, on occasion, visit a sampling point in order to carry out an interview by appointment and will use the opportunity to try and make contact at other addresses, may prove impossible to model.

4. Effect of Calling Times on Survey Response

An interviewer's task is not restricted, of course, to reducing noncontact rates – he/she must also obtain productive interviews. Reducing noncontact rates by calling on householders when they are most likely to be at home may, in principle, prove counterproductive if at certain times of day interviewers are more likely to get a refusal. It is common practice for interviewers to be advised not to call at unsociable hours such as very early morning or late evening. In this section we look at whether or not the FRS data gives evidence to either support or refute this view.

Although interest naturally lies with refusal rates, there are in fact three possible outcomes to a contact on the FRS: an interview, a refusal, and the making of an appointment. We consider this three-way split rather than the refusal/nonrefusal dichotomy. The ''refusal'' category on the FRS at individual calls is somewhat ambiguous because it is in fact an ''other'' category rather than a genuine ''refusal'' category, and includes cases where the interviewer did not get an outright refusal but nevertheless thought it sensible to withdraw from the address, perhaps to call again later. The FRS data-set does not allow us to distinguish between these.

Tables 9 and 10 give the outcome at the first contact by time of day and day of the week. Table 9 gives the percentages of people falling into each of the three outcome categories

Table 9. Outcome at time of first contact by call time (1)

Time of call	Outcome			
	Interview	Appointment	Other (including refusals)	Sample size
Saturday morning	27	41	32	536
Saturday afternoon	26	40	34	1,281
Saturday evening	20	49	31	144
Sunday morning	29	40	31	164
Sunday afternoon	24	43	33	594
Sunday evening	27	43	30	83
Weekday morning	35	33	31	2,217
Weekday afternoon	34	33	33	7,322
Weekday evening	25	42	33	4,684
Overall	31	37	33	17,025

Table 10. Outcome at time of first contact by call time (2)

Time of call	Outcome			
	Interview	Appointment	Other (including refusals)	Sample size
9.00–10.00	26	46	28	174
10.00–11.00	35	34	30	1,009
11.00–12.00	33	34	33	1,734
12.00–13.00	30	36	34	2,014
13.00–14.00	32	35	33	1,823
14.00–15.00	36	33	31	1,865
15.00–16.00	34	33	33	1,810
16.00–17.00	31	36	33	1,685
17.00–18.00	26	41	34	1,551
18.00–19.00	25	43	33	1,716
19.00–20.00	25	42	33	1,176
20.00–21.00	27	40	33	422
21.00–22.00	14	61	25	44
Overall	31	37	33	17,025

by time of call when first contact was made. Table 10 gives a similar breakdown, this time by hour of call. The main points to note are:

- As expected, interviewers tend not to call very early in the morning or very late in the evening. This makes it difficult to test for peaks in refusal rates at these times. The figures give no evidence for any such peaks, however.
- There is very little variation in the refusal rate by time of contact.
- Householders are significantly more likely to give an immediate interview if contact is made in the daytime rather than in the evening. Conversely, householders are significantly more likely to make an appointment with the interviewer if contact is made in the evening rather than in the daytime.

There are problems with looking at the data in this way. If the probability of refusal differs by household type or economic status then any differences in refusal rates we observe by time of day may be attributable to this difference. Unfortunately data on the characteristics of households and individuals is limited to responding units and is not available for non-responding ones.

An alternative way to look at the impact of contact times on survey outcome is to consider final outcome rather than immediate outcome. If it is the case that time of contact affects final outcome then we would expect those interviewers who most often make contact on weekday mornings and afternoons to have the highest response rates (although the differences would be expected to be small). In this context the four interviewer calling strategy groups identified in Section 3.1 prove useful.

The times of day at which first contacts are made vary considerably by interviewer calling strategy group, with BA and BB interviewers making around two-thirds of their first contacts on a weekday morning or afternoon as compared to 46% for AA interviewers and 48% for AB interviewers. Given that not all appointments lead on to interviews we

might expect to find that BA and BB interviewers achieve a higher response rate by the end of fieldwork.

In actual fact the FRS data shows almost no difference between response rates by interviewer strategy, the response rates by group differing by no more than one percentage point. This result remains even when interviewer experience is controlled for. Additional analysis suggests that whether or not an appointment is made is related more closely to the type of respondent than it is to the time of day of contact. In other words, some types of respondents simply prefer to make appointments irrespective of when they are contacted.

In conclusion, therefore, our analysis suggests that changing the time at which households are contacted has no detectable effect on survey response rates.

5. Discussion and Conclusions

The main finding of the analysis described in this article confirms what other survey researchers have found (see, for example, Weeks et al. 1990; Swires-Hennessy and Drake 1992), namely that, if the probability of contact at any one call is to be maximised then calls should be made on weekday evenings. If interviewers make use of this finding then it will be possible for them to reduce the total number of calls they need to make to meet a target noncontact rate. Furthermore, the logistic regression model described in Section 2.5 allows us to make estimates of the expected number of calls they will have to make to achieve this target.

Nevertheless it is not obvious, a priori, that instructing interviewers to increase the proportion of calls they make in the evening will lead to gains in efficiency. Restricting calls to evenings reduces the length of the working day and may result in interviewers having to make extra visits to sampling points to complete their assignments. This will increase travel costs and may, in principle at least, increase the number of hours spent working on the survey. Searching for the ''most efficient'' calling strategy would have to be an organisation specific activity, incorporating information on the organisations' payment system, interviewer workloads, and travel time and distance to sampling points. Organisations attempting to enforce efficiency in calling strategies would also have to be sensitive to the wishes and concerns of interviewers.

Were organisations to adopt extreme versions of calling strategies, then some consideration would need to be paid to the question of survey bias. Our analysis suggests that the time of contact does not impact on final response rates, but restricting calls to certain times of day may effectively exclude certain sub-groups from the survey (e.g., shift workers and those who refuse to answer their doors after dark). The profile of respondents to the FRS does not differ significantly by our four interviewer strategy groups but it is not possible to determine whether this would be the case under more extreme calling strategies.

This discussion of bias above refers to differences in the distribution of the type of people included in the survey based on different interviewer calling strategies. It does not address the overall issue of noncontact bias. We know from past research that noncontact bias depends on the allowable number of call-backs as respondents that are reached on the first few calls are often different from those found at home later in the

interviewing process (see, for example, Dunkelberg and Day 1973). Our study of the FRS contact patterns suggests that to achieve a four per cent noncontact rate (with a three weeks plus fieldwork period), then four calls should be considered the minimum rather than the maximum. A minimum of four calls would yield a response noncontact rate close to 11 per cent and potentially a higher noncontact bias. Arguably, ''efficient interviewer calling strategies'' should be defined in terms of bias reduction as well as in terms of cost (cf. Durbin 1954).

Discussions of overall bias should also take into consideration the work of Wilcox (1977) and Lievesley (1986), who found evidence to suggest that the bias incurred from noncontacts and the bias incurred from refusals may to some extent be compensating rather than additive. This suggests that reducing the noncontact rate and thus the noncontact bias, could actually increase the overall nonresponse bias.

6. References

Blossfield, H.P., Hamerle, A., and Mayer, K.U. (1989). Event History Analysis. Lawrence Erlbaum Associates, Publishers.

Campanelli, P., Sturgis, P., and Purdon, S. (1997). Can You Hear Me Knocking: An Investigation into the Impact of Interviewers on Survey Response Rates. London: SCPR.

Dunkelberg, W. and Day, G. (1973). Nonresponse Bias and Callbacks in Sample Surveys. Journal of Marketing Research, 10, 160–168.

Durbin, J. (1954). Nonresponse and Call-backs in Surveys. Bulletin of the International Statistical Institute, 34(2), 72–86.

Groves, R.M., Cialdini, R.B., and Couper, M.P. (1992). Understanding the Decision to Participate in a Survey. Public Opinion Quarterly, 56(4), 475–495.

Groves, R. and Robinson, D. (1983). Final Report on Callback Algorithms on CATI Systems. Report of the U.S. Bureau of the Census, Washington, DC.

Hill, D.H. (1978). Home Production and the Residential Electric Load Curve. Resources and Energy, 1, 339–358.

Kulka, R.A. and Weeks, M.F. (1988). Towards the Development of Optimal Calling Protocols for Telephone Surveys: A Conditional Probabilities Approach. Journal of Official Statistics, 319–332.

Lievesley, D. (1986). Unit Nonresponse in Interview Surveys. JCSM. Unpublished paper. London: SCPR.

Morton-Williams, J. (1993). Interviewer Approaches. Aldershot: Dartmouth Publishing Company Limited.

Swires-Hennessy, E. and Drake, M. (1992). The Optimum Time at Which to Conduct Interviews. Journal of the Market Research Society, 34(1), 61–72.

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

Weeks, M.F., Jones, B.L., Folsom, R.E., and Benrud, C.H. (1980). Optimal Times to Contact Sample Households. Public Opinion Quarterly, 44(1), 101–114.

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

Wilcox, J. (1977). The Interaction of Refusal and Not-at-home Sources of Nonresponse Bias. Journal of Marketing Research, 14, 592–597.
Vigderhous, G. (1981). Scheduling Telephone Interviews: A Study of Seasonal Patterns. Public Opinion Quarterly, 45, 250–259.

Received March 1997
Revised March 1998