Optimizing Call Time Lags by Modeling the Probability of Call Outcomes

August 2005

Brian J. Meekins, Bureau of Labor Statistics
Roberta L. Sangster, Bureau of Labor Statistics
John F. Meekins, Aerospace Corporation

In recent years the productivity of RDD telephone survey samples has declined. Two factors have largely contributed to this decline: a modest but consistent increase in refusals, and a more dramatic increase in phone numbers not able to be contacted. The dramatic increase in phone numbers that are not able to be contacted is largely due to the recent proliferation of sparsely populated 100 banks, leading to a decline in the proportion of actual household numbers in a given RDD sample. The decrease in nonresponse rates, increasing noncontact rates and poor calling efficiency is well documented in RDD telephone surveys and will only be briefly mentioned here in the context of the Telephone Point-of-Purchase Survey (TPOPS).

In light of this decrease in calling efficiency, the authors examine wait times or time lags between call attempts in order to find the optimal lag time. In the end the authors hope to find locate optimal wait times for different calling scenarios and thus inform operational guidelines used by calling centers.

DATA

The authors use data from the Telephone Point-of-Purchase Survey (TPOPS) The TPOPS is a nationally representative list-assisted RDD telephone survey conducted by the Census for the Bureau of Labor Statistics. The purpose of the survey is to collect outlet information (names and addresses) of the places where respondents shop. In addition, the amount of the purchase is also recorded. The TPOPS is a rotating panel survey conducted each quarter. Panel recruitment is conducted from a list-assisted RDD sample. TPOPS has an eight week calling period and does not use any within household selection. Approximately 40% of each quarter's sample is "new" RDD sample. The total sample size per quarter varies, but is close to 38,000. This analysis includes data from RDD (first interview) portion of the sample from 2001, 2002, and the first quarter of 2003.

Figure 1 shows the response rate, the percent of the total sample determined to be eligible, and the percentage of the current sample (2004 quarter 4) used to obtain the necessary completions. It is clear from this chart that the response rate for the TPOPS has been declining since 2001, but this is roughly on the same scale as the decline in the percentage of eligible households in the sample. In response the amount of RDD sample used in the TPOPS has greatly increased. Indeed the amount of sample of used in 2001 was only about 60% of the amount used in 2004.

Table 1 shows the outcome of attempts, grouped by their attempt number in the call history, indicating that 6.7% of the first three attempts to all households result in a completion. One can easily see that completions compose a small number of the total attempts to households and as calling increases those resulting in completions become even less frequent. Answering machines and ring-no-answers (RNAs) compose the largest percentage at all call attempts. Note also that nonworking and ineligible numbers are removed quickly from the sample. Our ultimate goal in this research is to decrease the number of total call attempts to the sample without decreasing the number of completions.

The TPOPS uses some “calling rules” in order to limit inefficiency in calling. After 12 consecutive RNAs calling is stopped on a sample unit. Calling is also stopped on a sample unit if 30 attempts are reached. However, in some circumstances this rule is violated (e.g. pending callbacks). Adamant or “hard” refusals are removed immediately, but the TPOPS receives very few of these. Sample units are removed after two tentative or “soft” refusals.

Most importantly for this research, the TPOPS has a number of time lags that are pre-set in its computerized calling system. Call attempts to a sampling unit following an attempt resulting in answering machine should occur at least three days after. Refusals are not to be called before a wait of one day. Immediate hang-ups also wait at least one day and numbers that were temporarily out of service wait 2 days. No sampling unit is called more than five times within the same day. The delay on busy signals does not appear to be set, and may, in fact, be at the interviewers’ discretion. The same is true for RNAs, although these seemed to be returned to the calling queue immediately (and called based on priority). In general, calling rules appear to be frequently violated. While this may be problematic from an operational standpoint, it enables the authors to examine time lags of varying length to determine optimality given the immediate call history. Over 140,000 sample units and total of about 764,000 attempts on these units are examined.

METHODOLOGY

Figure 2 shows a model of call attempts outcomes and subsequent action ($delay_x$ or drop). The length of the time lag is dependent only upon this previous call attempts. As indicated by the differing subscripts the hypothesized time lags are different for RNA, busy, phone problem, bad number, answering machine, language problems, deaf respondent, and temporarily sick. Note that soft refusals and callbacks without an appointment are thought to have similar lag times.

Although such a model can be simultaneously estimated, this papers examines each call outcome separately. We focus on those that are of primary interest: RNA, busy, answering machine, and tentative refusal or callback. For each of these call dispositions the primary mode of inquiry are scatterplots of the probability of either contact or completion by lag time. These probabilities are calculated for changes in time lag times creating “bins” with a sufficient number of cases to calculate probabilities of contact or completion based on the observed number of contacts or completions obtained in that lag time period divided by the total number of calls made during the lag time period. These lag time “bins” are small relative to the total observed lag time. In some cases, weighted least squares and smoothing is used to assess fit. For some of the same time lag bins

RESULTS

RNA dispositions

Figure 3 plots the probability of contact given a previous attempt of RNA by lag time up to 20,000 (14 days). In this case the, the lag time bins used were 5 to 10 minutes in size depending on the density of calling. As shown in figure 3 there is a consistent linear relationship between lag time and the probability of contact. The longer the delay, the more likely contact with a sample unit is achieved. The weighted lease squares (WLS) estimate for beta is 0.569, indicating a fairly strong positive relationship. The value of R^2 is 0.341. Upon examining this plot, however, it becomes clear that there are periodic areas of high error associated with the calculated probabilities. In between these areas there is some evidence of a curvilinear relationship (e.g. between 0-700 minutes). Results from WLS regression indicates and R^2 of around 0.15 for these areas. This relationship appears consistently even after smoothing using running medians and Hanning and collapsing lag times into larger bins. Figures 4 & 5 show this relationship for time lags of 750-2,000 minutes and 3,500-5,000 minutes, respectively. One possible explanation is that time lag is related to the probability of contact in the manner shown in Figure 6. That is, the relationship is a combination of both a periodic fluctuation and a linear relationship in the probability of contact. This would be consistent with the hypothesis that some sampling units are easier to reach during some times of the day, and in other sampling units members are absent for long periods of time.

Figure 7. examines one of the these abbreviated time lag periods (between 750 and 2,000 minutes) and superimposes a plot of calling density, calculated as the proportion of the total calls made during each lag time bin (re-scaled). From Figure 7, we see that TPOPS actually calls most during lag times of the lowest probability.

Busy Dispositions

Except for the first two hundred minutes, the relationship of contact to lag time following a busy disposition is quite flat (Figure 8). In the first 200 minutes there is a fairly strong negative relationship between contact and lag time, where WLS R^2 is .431 and the standardized slope is -.508. This relationship is illustrated in Figure 9, together with the calling density. For sampling units with a busy disposition, it appears that most of the subsequent call attempts are occurring at the time of greatest probability.

Answering Machine Dispositions

The probability of contact following the disposition of an answering machine is modestly and positively related to lag time. This relationship is shown clearly in Figure 10. The WLS standardized regression coefficient is .208, with an R^2 of .252. The majority of calls to answering machines, however, occur within the first 5,000 minutes (3-4 days) at the time when the probability of contact is lowest.

Tentative Refusal or Indefinite Callback

In analyzing the lag time for tentative refusals or indefinite callbacks we are concerned more with completion rather than contact. Contact has already been established with the sampling unit and there is not value in additional contacts without a completion (unlike busy, for example, where TPOPS may be able to classify that household as ineligible).

Figure 11 shows the probability of contact and completion given a tentative refusal or indefinite callback (where no appointment is made) by lag time for 10,000 minutes (1 week). We can see that contact decreases in the first 500 minutes and then remains quite flat, whereas the probability of completion is essentially flat across all lag times. Also plotted is the calling density which indicates a large amount of early calling and a very clear periodicity. The large amount of very early calling is a result of the practice of confirming the number via an immediate callback to a refusal. Usually cooperation for an interview is not solicited at these times. The periodicity in the call density is the result of a fairly strict one day calling delay rule. However, this periodicity seems unnecessary in light of the probability of completion.

Conclusions and Recommendations

Some fairly clear conclusions arise from these results. Firstly, with calls that result in busy dispositions, the TPOPS call centers appear to be calling at the appropriate times to maximize the likelihood of a contact. However, some sampling units with busy disposition wait a long time. This should be rectified.

The probability of completing and interview with sampling units that initially refuse or asked to be called back without setting up an appointment is quite flat across lag times. Therefore, the calling rule can be relaxed in this case.

The lag time for sampling units with answering machine dispositions should be at least 3 to 4 days based on the probability of contact. Coincidentally, this is supposed to be the calling rule in place. However, as shown in Figure 10 a large amount of calling occurs earlier then 3 or 4 days.

For sampling units with a disposition of RNA the results are more complex. The relationship of the probability of contact with lag time is strongly positive. However, given the limited calling period, a long delay is not necessarily recommended. Table 2 shows the mean lag time by attempt number. Following the first attempt sampling units wait, on average 3,000 minutes (~2 days) until the next attempt. Although the standard deviation is quite large this table is useful in estimating the number of attempts that can be made in a give calling period. The TPOPS calling period is 8 weeks or 80,640 minutes. Clearly the most efficient method of calling would be to have long lag times (with high probabilities of contact) and few attempts. However, as Figure 12 shows the relationship between the probability of contact and attempt number is fairly flat, arguing for a large number of attempts. Given the relatively low cost of a RNA disposition, the data in its entirety argue for a large number of attempts to be made.

Figure 1.

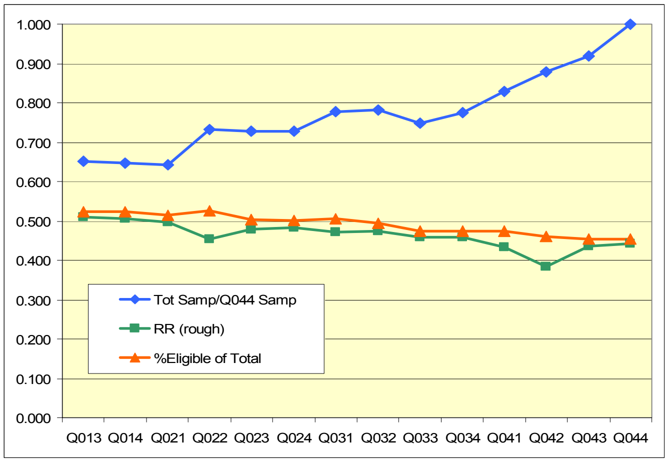

Table 1

Percent Outcomes by Call Attempts						
Outcomes	1 - 3	4 - 6	7 - 12	13 - 20	21 - 32	33 to 45
Complete	6.7	5.0	2.6	2.5	1.6	0.5
Refusal	7.3	5.8	3.3	3.7	2.7	1.2
Callback	11.1	12.0	8.8	13.5	10.9	9.1
Answ. Mach.	22.3	26.9	24.4	48.3	50.5	18.0
Ring-No-Ans	21.1	34.7	44.3	22.7	26.2	54.9
Busy	6.5	5.8	6.2	3.9	3.7	6.0
Other NCs	1.1	1.2	0.9	1.5	1.2	1.1
Phone Prob.	3.8	5.8	8.0	2.7	2.6	8.5
NonWork\Cell	16.3	1.2	0.7	0.6	0.4	0.6
Inelig. Other	3.8	1.5	0.8	0.6	0.3	0.1
Total Attempts	302785	149462	196485	71301	42545	2082
n=	143616	56646	39327	12246	5575	783

Figure 2

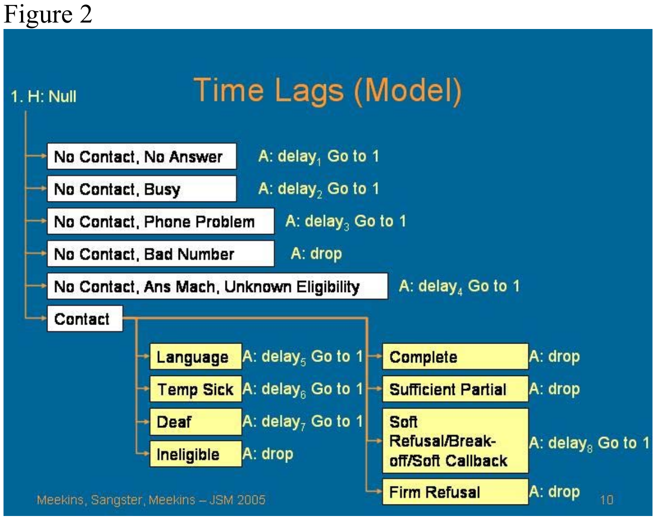

Figure 3.

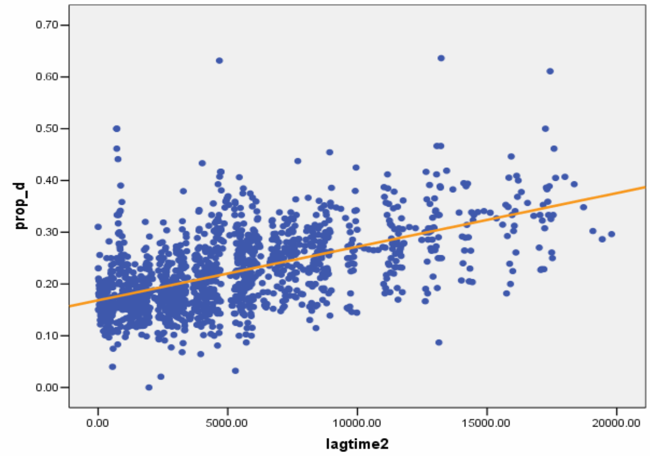

Figure 4.

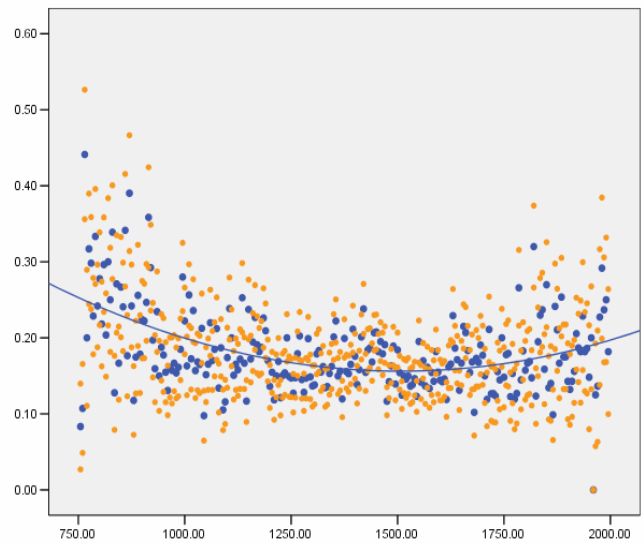

Figure 5

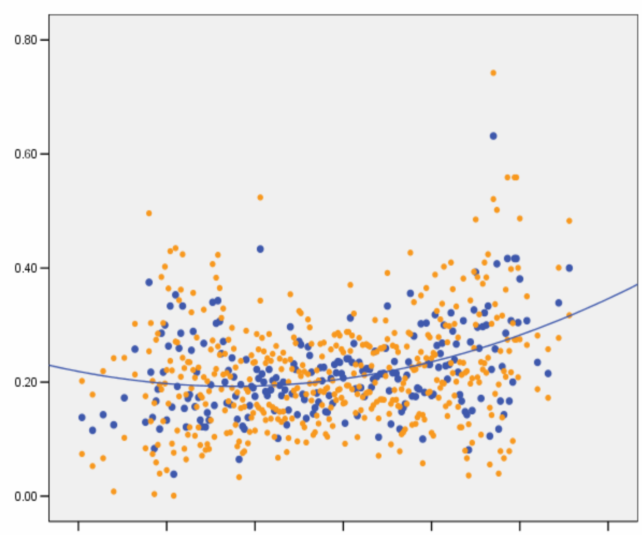

Figure 6.

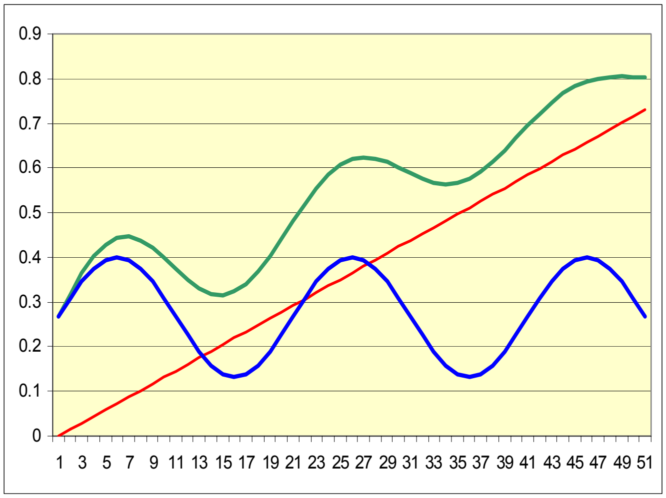

Figure7.

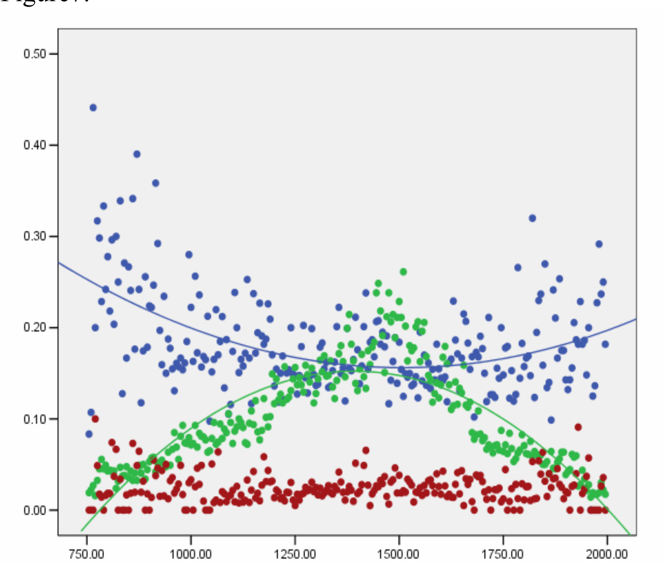

Figure 8.

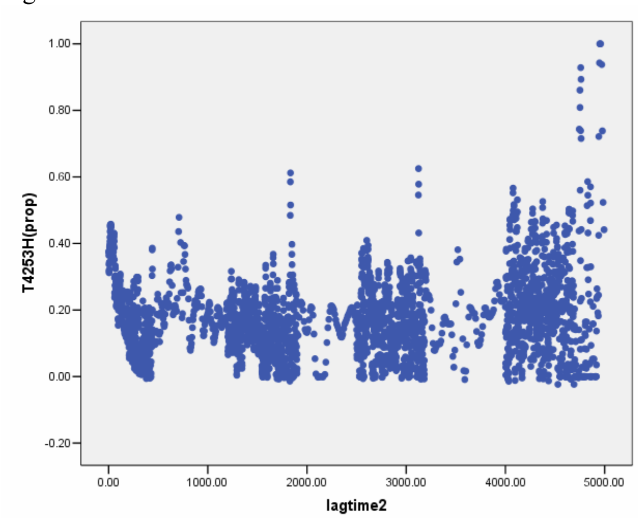

Figure 9.

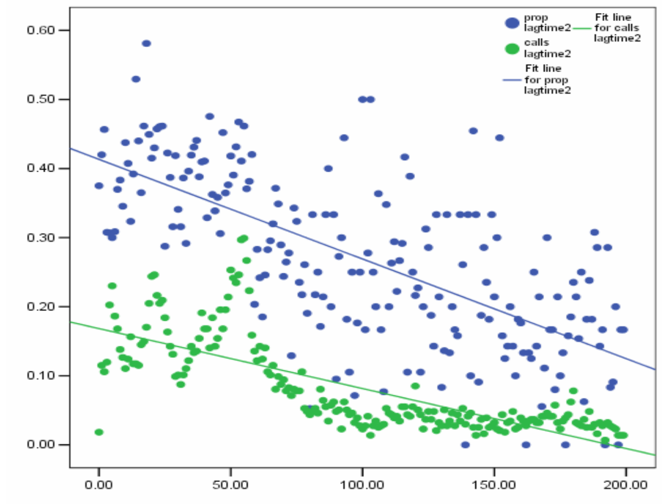

Figure 10

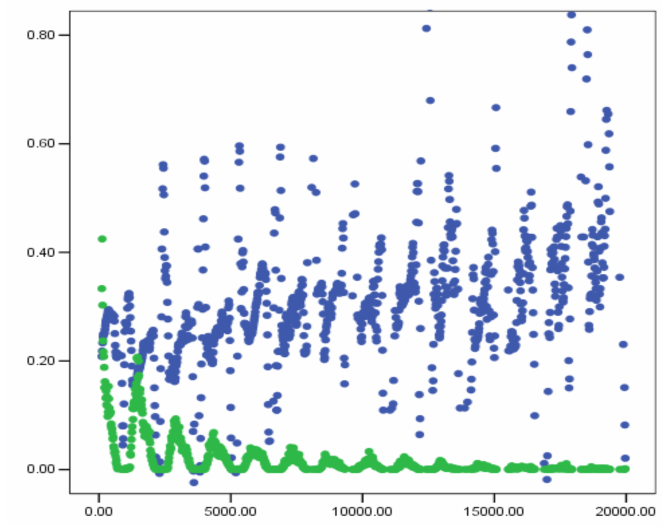

Figure 11.

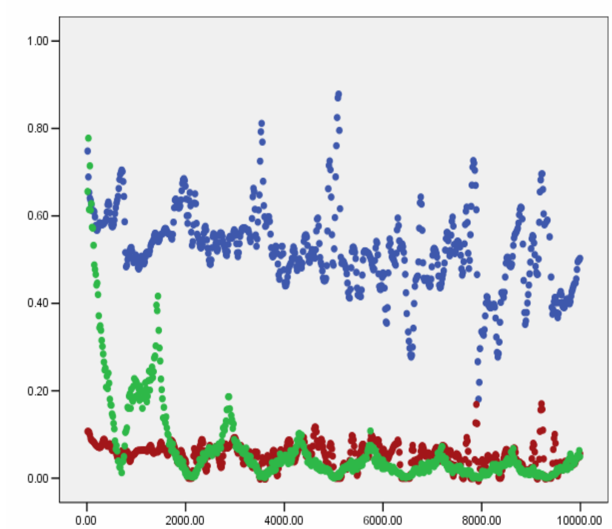

Table 2.

atttime

lagatt	Mean	N	Std. Deviation
1.00	3038.7582	111200	3929.34515
2.00	5848.1310	98532	6211.47267
3.00	8364.5067	87070	7757.51716
4.00	10657.79	76853	8954.29184
5.00	12790.26	67785	10052.56031
6.00	14829.72	59715	11086.86123
7.00	16689.55	52378	12026.10671
8.00	18324.16	45549	12846.82701
9.00	19687.35	39156	13552.36641
10.00	20785.04	33328	14189.76534

Figure 12.