Are Two Feet in the Door Better than One? Using Process Data to Examine Interviewer Effort and Nonresponse Bias

Kevin Wang, Joe Murphy, Rodney Baxter, Jeremy Aldworth

RTI International
3040 Cornwallis Road, RTP, NC 27709-2194
kwang@rti.org
jmurphy@rti.org
baxter@rti.org
jaldworth@rti.org

Introduction

Survey organizations employ numerous tactics to reduce the potential for bias due to unit nonresponse. After data collection, nonresponse and population weighting adjustments are often utilized to reduce potential bias from nonresponse and undercoverage. Prior to and during data collection, interviewers are trained, and sometimes re-trained in techniques for gaining cooperation from reluctant respondents. In addition, interviewers attempt to reduce the potential for bias from nonresponse by revisiting sampled households, and attempt to persuade reluctant respondents to cooperate. In some cases, supervisors may assign highly skilled interviewers to persuade the most reluctant respondents to cooperate.

The amount of effort that interviewers can apply to achieve high response rates is limited by time and resource constraints. That "effort" may be regarded as a resource to be allocated in an efficient manner in order to reduce the potential for bias from nonresponse while simultaneously maintaining a dedicated and productive field staff. Indiscriminate application of extensive efforts to gain cooperation may not reduce the potential of bias from nonresponse in an effective manner. In principle, effort could be differentially applied to cases based on the potential for reducing bias from nonresponse (either by lowering the nonresponse rate or targeting cases that might be especially different from current respondents). In order to manage field staff in this way, survey organizations require information on the likelihood of successfully completing an interview, especially with regard to factors that are potentially under management control such as how field interviewers manage their time and organize their assignments. In addition, these efforts should be informed by an understanding of how interviewers allocate their efforts and whether the level of effort is related to survey measures of interest.

In this paper, we examine the use of administrative call record data from the National Survey on Drug Use and Health (NSDUH) in order to address these needs. We begin by describing the NSDUH and the available call record data. In the second section, we examine how the calling strategies of interviewers can affect contact and cooperation rates. In the third section, we conduct analyses that examine the relationships between the volume of call attempts and survey estimates and in turn, the potential for bias due to nonresponse.

Background

The NSDUH is an in-person cross-sectional study of nearly 67,500 persons in all 50 states and the District of Columbia each year. The survey is administered primarily through an audio computer-assisted self-interview (ACASI) via a laptop computer and is designed to measure the prevalence and correlates of drug use in the United States population age 12 and older. Respondents completing the one-hour interview receive a $30 cash incentive. The NSDUH is sponsored by the Substance Abuse and Mental Health Services Administration (SAMHSA). Data collection is conducted by RTI International.

The survey sample consists of a 50-state design with an independent, multistage area probability sample for each of the 50 states and the District of Columbia. The sample design features eight large sample states (California, Florida,

Illinois, Michigan, New York, Ohio, Pennsylvania and Texas) that have sample sizes large enough to allow for design-based state level estimates. For other states and the District of Columbia, sample sizes are adequate for the production of state level estimates using small area estimation (SAE) methods. The design also oversamples children (ages 12 to 17) and young adults (ages 18 to 25). The survey is designed to produce estimates for the civilian, noninstitutional population aged 12 years or older. The NSDUH universe includes residents of noninstitutional group quarters (e.g. shelters, rooming houses, dormitories), residents of Alaska and Hawaii, and civilians living on military bases. Persons not covered by the survey include those with no fixed household address (e.g. homeless persons not living shelters), and residents of institutional group quarters, such as jails and hospitals. The NSDUH is fielded on a quarterly basis; roughly one-fourth of the sample is fielded each calendar quarter and interviewing is conducted within that quarter.

The NSDUH interviewing process consists of a screener interview with sampled dwelling units in order to determine if any household members are eligible for the survey. The screener interview is conducted using a handheld computer (HP IPAQ), which gathers information on household size, age, and race and ethnic background to determine eligibility of the household and then selects zero, one or two household members for the survey. About 170,000 households are screened each year. Interviews are conducted primarily using ACASI but some items (mostly demographic) are gathered through Computer Assisted Personal Interviewing (CAPI).

Table 1 shows the screening dispositions for the 2004 NSDUH. Consistent with NSDUH data from prior years, nonresponse on the NSDUH screener is dominated by refusals. About 69 percent of the unweighted screener nonresponse is made up of refusals with another 18 percent consisting of "not at home" cases. *Table 2* shows the Final interview dispositions for the 2004 NSDUH. Similar to the screener, about 65 percent of the interview nonresponse is due to refusals.

Table 1   Distribution of 2004 NSDUH Screening Result Codes

	Sample Size	Percent of Total
Total Sample	169,514	100.00
Ineligible cases	26,902	15.87
Eligible cases	142,612	84.13
Ineligibles	26,902	15.87
Vacant	15,204	56.52
Not a primary residence	4,122	15.32
Not a dwelling unit	2,062	7.66
All military personnel	282	1.05
Other, ineligible	5,232	19.45
Eligible Cases	142,612	84.13
Screening complete	130,130	91.25
No one selected	73,732	56.66
One selected	30,499	23.44
Two selected	25,899	19.90
Screening not complete	12,482	8.75
No one home	2,207	17.68
Respondent unavailable	259	2.07
Physically or mentally incompetent	265	2.12
Language barrier—Hispanic	51	0.41
Language barrier—Other	391	3.13
Refusal	8,588	68.80
Other, access denied	660	5.29
Other, eligible	10	0.08
Other, nonresponse	51	0.41

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004.

Table 2   Distribution of 2004 NSDUH Interview Result Codes

Final Interview Code	Persons Aged 12 or Older		Persons Aged 12 to 17		Persons Aged 18 or Older	
	Sample Size	%	Sample Size	%	Sample Size	%
Total	81,973	100.00	25,141	100.00	56,832	100.00
Interview Complete	67,760	82.66	22,309	88.74	45,451	79.97
No One at Dwelling Unit	1,156	1.41	147	0.58	1,009	1.78
Respondent Unavailable	1,762	2.15	302	1.20	1,460	2.57
Break-Off	46	0.06	7	0.03	39	0.07
Physically/Mentally Incompetent	699	0.85	143	0.57	556	0.98
Language Barrier – Hispanic	131	0.16	12	0.05	119	0.21
Language Barrier – Other	398	0.49	27	0.11	371	0.65
Refusal	7,871	9.60	583	2.32	7,288	12.82
Parental Refusal	1,491	1.82	1,491	5.93	0	0.00
Other	659	0.80	120	0.48	539	0.95

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004.

Upon each visit to a sampled dwelling unit, interviewers enter case status information into the "record of calls" (ROC) using the handheld computer. Collected data elements include the outcome of the call (noncontact, refusal, completed screening, completed interview, etc.) and an open-ended notes field for recording information to help with case management and scheduling. Other elements, including the time and day of the call, are automatically recorded in the handheld computer. These results are transmitted back to a central database daily and can be reviewed by an interviewer's supervisors in a web accessible case management system (CMS). The supervisor can view attempts by day and time and can suggest alternative dates and times to attempt a case that is difficult to contact.

Frequencies of result codes for all screener attempts to a dwelling unit from the ROC data are shown in *Table 3*. Result codes are grouped by pending vs. final status and among finalized cases, by whether or not the sampled dwelling unit was eligible for the survey. If the sampled dwelling unit was eligible for the survey, frequencies are shown for whether or not the screener was completed.

We should note that the number of records in the ROC data can reflect entries other than those that represent actual visits to a dwelling unit. For example, final result codes for screener nonresponse cases are entered into the ROC data as separate entries by the interviewer after review and approval of the final code by a field supervisor. The time and date that is recorded for these entries reflect the point at which this decision is made rather than the last actual visit. As such, these entries are not considered "attempts" and were removed from the data for our analyses.

In their analysis of ROC data from the 2001 and 2002 NSDUH surveys, Painter, Chromy, Meyer, Granger and Clarke (2003) note that interviewers are instructed to work in a segment at least four hours out of the day they are in the segment and that interviewers can make multiple visits to the same dwelling unit within the same day. As such, the number of attempts may overstate the level of effort on the part of the interviewer, particularly if the more crucial decision is whether the interviewer should go back out to the segment at all. That is, if a significant portion of the field cost is due to interviewer travel to segments, then analysis of interviewer effort should be based on call days (the number of different days that an interviewer visits a sampled dwelling unit) rather the number of individual attempts. In some of our analyses, we use the concept of the call attempt while for others we use the call day concept. To provide more detail on variation in effort by time of day, we divide call days into two "call slots"

(before 4PM and after 4PM) where appropriate.

Table 3   Frequencies of Result Codes for All Screening Attempts During 2004 NSDUH

Code	Description	Frequency	Percent
Pending Screening Codes			
	Total Pending	490,049	100.00
01	No One at DU	377,029	57.28
02	Screening Respondent Unavailable	43,674	6.63
03	Neighbor Indicates Occupancy	14,495	2.20
04	Physically/Mentally Incompetent	469	0.07
05	Language Barrier (Spanish)	1,352	0.21
06	Language Barrier (Other)	1,123	0.17
07	Refusal to Screening Questions	29,089	4.42
08	Unable to Locate SDU	2,107	0.32
09	Other	20,711	3.15
Final Screening Codes			
Screener Ineligibles			
	Total Screener Ineligibles	26,464	
10	Vacant	15,029	2.28
13	Not a Primary Residence	4,055	0.62
18	Not a Dwelling Unit	1,970	0.30
19	GQU Listed as HU	38	0.01
20	HU listed as GQU	1	0.00
22	DU Contains Only Military Personnel	280	0.04
25	No Eligible SDU Members	28	0.00
26	In DU less than ½ of the Quarter	4,353	0.66
29	Listing Error	710	0.11
Screener Nonresponse			
	Total Screener Nonresponse	11,709	
11	No One at DU after Repeated Visits	2,145	0.33
12	SR Unavailable after Repeated Visits	250	0.04
14	Physically/Mentally Incompetent	259	0.04
15	Language Barrier (Spanish)	49	0.01
16	Language Barrier (Other)	387	0.06
17	Refusal	8,439	1.28
23	Other	180	0.03
Screener Completes			
	Total Screener Completes	129,446	
30	No One Selected for Interview	73,385	11.15
31	One Selected for Interview	30,313	4.61
32	Two Selected for Interview	25,748	3.91

Abbreviations: SDU (Sampled Dwelling Unit); GQU (Group Quarters Unit); DU (Dwelling Unit); HU (Household Unit); SR (Screener Respondent)
Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004.

Table 4 presents frequencies on all call attempts from the ROC for Interview cases. Result codes are grouped by pending interview codes, interview nonresponse codes, and interview complete. Percentages are calculated as the percentage of all interview attempts that the code represents. The number of final completes is somewhat higher than the number of completes reported in *Table 2*. The ROC file is unedited and as such contains all attempts made in the field. As questionnaire data are edited and evaluated after collection, a number of cases are removed from the final data set as not usable as a complete. As we are studying effort used in data collection, we include in our

analyses all call attempts made in the field regardless of whether it is retained in the final data.

Table 4   Frequencies of Result Codes for All Interview Attempts During 2004 NSDUH

Code	Description	Frequency	Percent
Pending Interview Codes			
	Total Pending Interview	227,119	
50	Appointment For Interview	48,895	15.48
51	No One At Dwelling Unit	84,186	26.66
52	Respondent Unavailable	69,573	22.03
53	Breakoff (Partial Interview)	469	0.15
54	Physically/mentally Incompetent	734	0.23
55	Language barrier (Spanish)	1,145	0.36
56	Language barrier (Other)	475	0.15
57	Refusal	17,514	5.55
58	Parental Refusal for 12-17 Year Old	4,128	1.31
59	Other	5,881	1.86
Final Interview Codes			
Interview Nonresponse			
	Total Interview Nonresponse	14,719	
71	No One At Home After Repeated Visits	1,188	0.38
72	Respondent Unavailable after Repeat Visits	1,725	0.55
73	Break Off (Partial Interview)	56	0.02
74	Physically/mentally Incompetent	700	0.22
75	Language Barrier (Spanish)	145	0.05
76	Language Barrier (Other)	394	0.12
77	Refusal	7,669	2.43
78	Parental Refusal for 12-17 Year Old	1,778	0.56
79	Other	1,064	0.34
Interview Completes			
70	Interview Complete	68,072	21.56

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004.

Analysis of Interviewer Call Strategies

One aspect of interviewer effort concerns call strategy rather than simply call volume. The quantity of calls may reflect a great deal of effort, but that effort is wasted if the calls were made at times respondents were unavailable. If an interviewer made 20 calls to a household by visiting every Friday afternoon at between 11AM and 3PM for four weeks in a row and never found anyone home, the volume of effort would be great, but the strategy would be poor. The more efficient a call strategy is, the fewer the number of calls needed to finalize a case. The use of more efficient calling strategies could result in more resources available for returning to households that are difficult to screen or interview. For these reasons, we examine the effort interviewers apply to initial call attempts. By identifying potentially ineffective call efforts, we can identify areas for reorganization of effort to maximize resources.

Table 5 presents data from first calls made by interviewers in Quarters 3 and 4 of the 2004 NSDUH. We divided the week into 14 "call slots" where each day of the week has two slots – before 4PM and after 4PM.[1] Because Quarter 3 began on a Thursday and Quarter 4 on a Friday, it is not surprising to see that a relatively large percentage of first calls were made on those days. This is consistent with examination of these data from additional quarters. NSDUH interviewers devote a considerable amount of effort to interviewing on Saturdays. In additional tabulations of call record data for 2003 and all of 2004, we found that interviewers made about 20 to 24 percent of their screening

---

[1] We selected 4PM as a defining point since guidelines for NSDUH interviewers specifically mention after 4PM as a productive time to make call attempts.

attempts on Saturdays. Interviewer guidelines for the NSDUH specifically mention Saturdays as especially productive days for interviewing. It may also be the case that interviewers view Saturday as a particularly efficient day for obtaining interviews.

Interviewers also tend to make more attempts during daytime hours rather than evening hours, which is consistent with findings from other surveys. Based on analyses of data from the Survey of Income and Program Participation Methods Panel, Nancy Bates (2003: 10) writes, "It appears that current interviewer practices are over-emphasizing weekday afternoons and underutilizing weekends and weekday nights". In their analysis of calling patterns in the National Survey of Health and Stress (NSHS), Groves and Couper (1998) note that the preponderance of daytime vs. evening call attempts may simply be an artifact of more hours in the daytime than in the evening, although our classification of evening as "after 4PM" allows for a more even distribution of available hours between daytime and evening calls. But they also go on to note that first calls may take place disproportionately in the daytime if interviewers are "precanvassing" their assignments. That is, in the initial visit, interviewers may assess the environment for personal safety concerns as well as "evidence of any locked gates or entrances, levels of visible activity in the area, places to park the car within observational distance of sample units, etc." This practice is generally discouraged in the NSDUH but it is also less necessary. NSDUH uses many experienced field interviewers and these interviewers are usually assigned to segments that they have worked in the past and as such, are generally familiar with the characteristics of their assignments. Some leeway is given to novice interviewers or interviewers assigned to areas they have not worked before to conduct "precanvassing".

Table 5   First Call Statistics by Slot for Quarters 3 and 4, 2004 NSDUH

Call Slot	First Calls	Percent of First Calls	Percent Contacted	Percent Complete (Given Contact Made)	Response Rate
Sunday before 4PM	2,335	3.06%	41.07%	71.53%	29.38%
Sunday after 4PM	1,601	2.10%	44.72%	73.18%	32.73%
Monday before 4PM	4,549	5.96%	38.49%	74.36%	28.62%
Monday after 4PM	3,308	4.33%	46.58%	74.30%	34.61%
Tuesday before 4PM	4,812	6.30%	37.47%	70.72%	26.50%
Tuesday after 4PM	3,712	4.86%	47.90%	73.68%	35.29%
Wednesday before 4PM	4,595	6.02%	35.58%	75.23%	26.77%
Wednesday after 4PM	3,462	4.53%	46.04%	73.71%	33.94%
Thursday before 4PM	8,532	11.17%	36.53%	72.60%	26.52%
Thursday after 4PM	6,463	8.46%	43.90%	72.01%	31.61%
Friday before 4PM	10,210	13.37%	35.61%	72.72%	25.90%
Friday after 4PM	6,824	8.94%	43.30%	70.73%	30.63%
Saturday before 4PM	11,795	15.45%	42.50%	71.21%	30.26%
Saturday after 4PM	4,159	5.45%	42.10%	71.10%	29.93%
Total	*76,357*	*100.00%*	*40.71%*	*72.39%*	*29.47%*

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004.

For the data in *Table 5*, interviewers typically make about 1.3 to 1.5 attempts during daytime hours for each attempt that is made in the evening hours. The exception to this is Saturday, in which the daytime to evening attempt ratio is 2.8. We investigated whether fewer calls were made in the evening because evening calls are more likely to be interview calls lasting an hour or more. A review of the data suggests that a greater proportion of evening calls represent interviews than day calls, but the difference does not fully account for the lower rate of interviewer effort expended during evenings.

Table 5 also provides contact rates and cooperation rates (conditional upon contact) for each call slot. Among the

pending screening codes, No One at DU (01), Neighbor Indicates Occupancy (03), Unable to Locate Sampled Dwelling Unit (08) and Other (09) were considered noncontacts and all other pending codes were considered contacts. Cases that were finalized as ineligible on the screener were excluded from the entire calculation. The results suggest that day slots may not result in the highest rates of contact with household members – a prerequisite for obtaining completed screening interviews. For most days of the week, contact rates are higher for evening attempts than for the corresponding daytime attempts. For example, in the case of Tuesdays, the difference in the contact rate is about 10 percentage points (47.9 percent for Tuesday evening as opposed to 37.5 percent for Tuesday afternoons). The notable exception to this is Saturday in which the contact rates are similar between evening and daytime calls. This makes sense since it is common knowledge that people work or go to school during the week rather than on the weekend.

In order to look more closely at the question of the most productive times of day to make a first call to a case, we plotted the contact rate for first calls by hour of the day for Quarters 3 and 4 of the 2004 NSDUH. This was done separately for weekdays (Monday through Friday) and weekend days (Saturday and Sunday) since the hours that people are at home likely differ by the two categories and as noted earlier, the distribution of calls before and after 4PM is quite different on Saturdays than for weekdays. We also plotted the percent of first calls made by interviewers by hour of the day to get an idea of whether they concentrated their effort at the most productive times of the day.

Figure 1 shows that the contact rate for first calls made during the week was generally around 35% before 3:00PM. The contact rate increased throughout the morning and into the early afternoon. Between 3PM and 6PM, the contact rate increased sharply from a little over 35% to a little over 45%. The contact rate remained above or near 45% into the 8PM hour. The plot of calls made by interviewers followed a more normal distribution, with the high point coming in the 3PM hour.

Figure 1   **Contact Rate and Proportions of First Calls Made by Hour for Weekdays, Quarters 3 and 4, 2004 NSDUH**
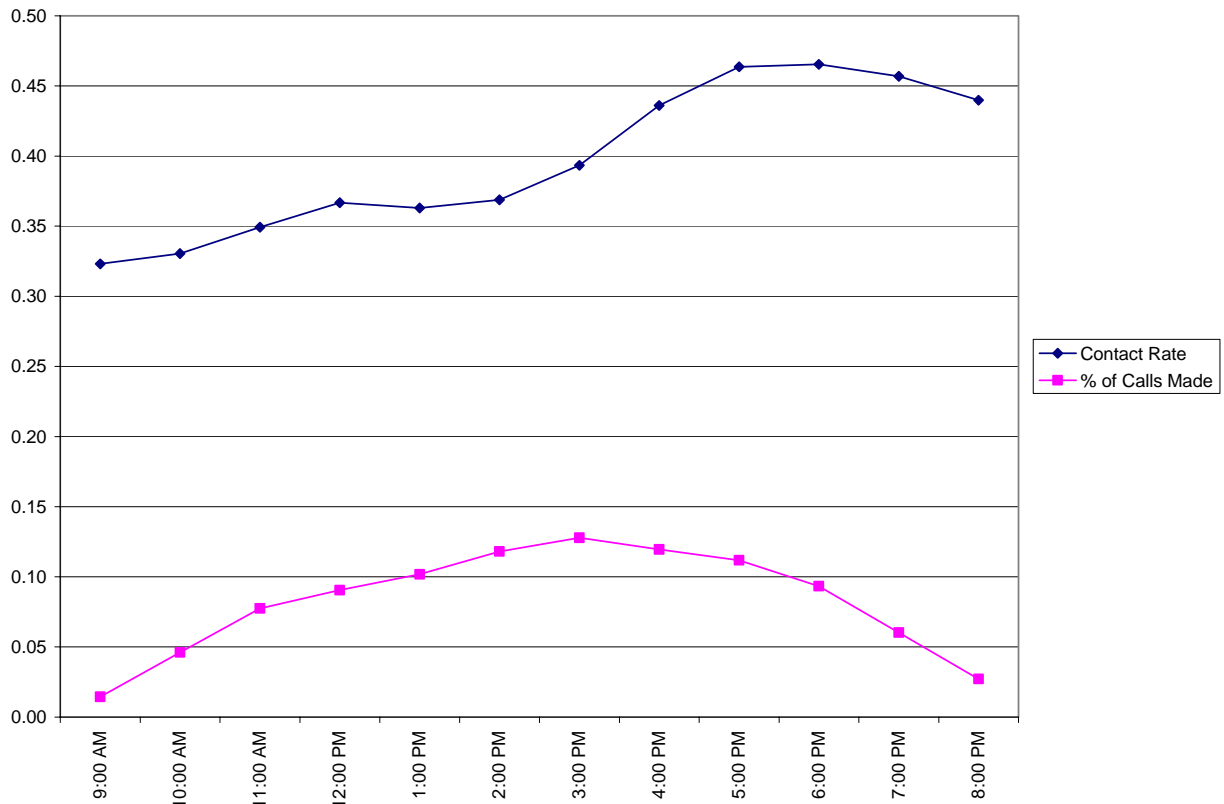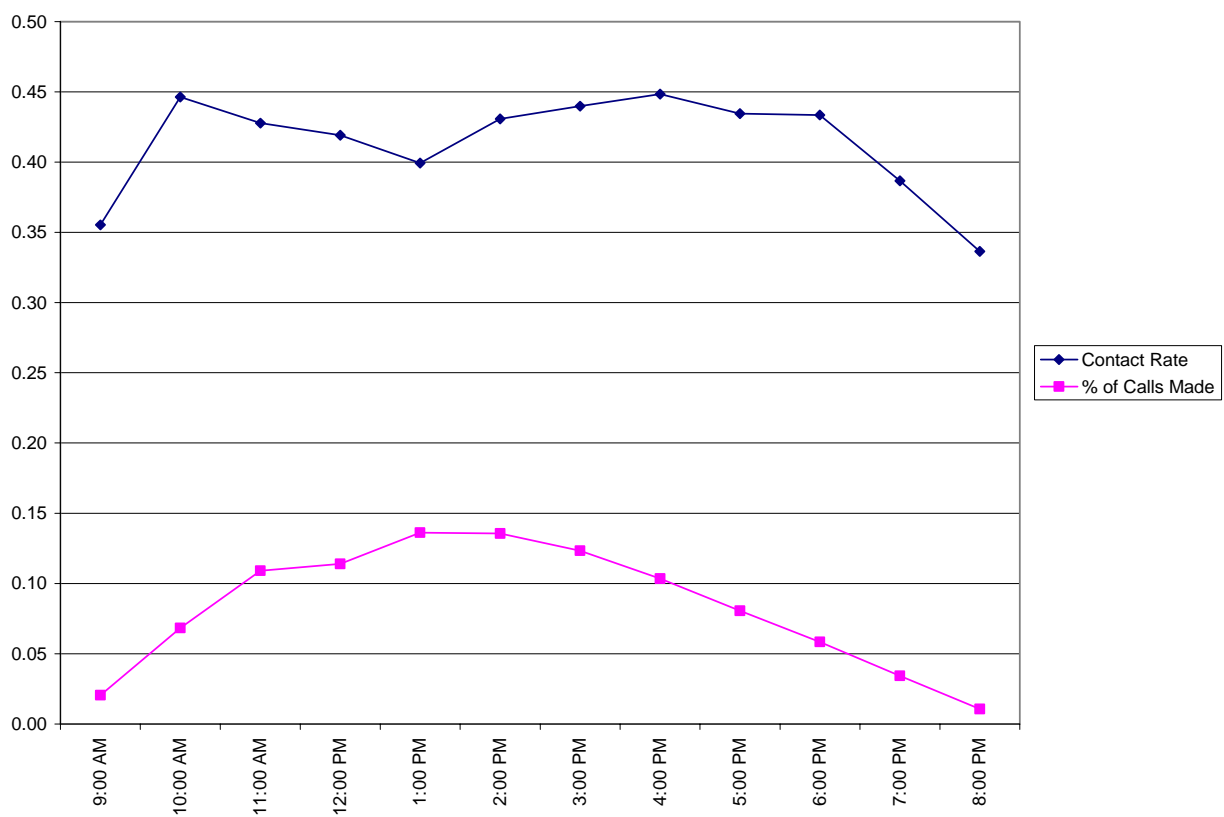
Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004.

Figure 2 presents the contact rate and distribution of calls by hour of the day for first calls made on the weekend. Compared to weekdays, the contact rate was generally high during the first part of the day. The contact rate remained between 40% and 45% from 10AM through the 6PM hour. The contact rate declined beginning at 7PM. The distribution of first calls again followed a normal curve, but the modal time of 1PM was a little earlier than during the week. Interestingly, this high point in interviewer effort corresponds with a slight dip in the contact rate seen around 1PM for weekend first calls.

Figure 2 **Contact Rate and Proportions of First Calls Made by Hour for Weekends, Quarters 3 and 4, 2004 NSDUH**

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004.

To test whether the effect of time of slot could be "explained away" by the inclusion of other variables found to correlate with contact, we ran a logistic regression model predicting contact after one call. The variables used as predictors included: whether or not the first call was made during the evening (before or after 4PM); whether or not the first call was made on the weekend (Saturday or Sunday); whether or not the selected unit had controlled access features that might restrict access (gate, guard, etc.), whether the unit was in a building with 2-9 units, 10-49 units, 50+ units, or was not a house or apartment (single unit buildings were the reference category); the number of years of experience of the interviewer; the percentage of the segment in which the unit was located that was age 50 or older; the percentage in the segment who were black; the percentage who were Hispanic; several variables indicating population density; and the median housing value in the segment. The results of this model are provided in *Table 6*.

Table 6   Logistic Regression Model Predicting Contact Based on First Call Characteristics, Quarters 3 and 4, 2004 NSDUH

Variable	Coefficient	Standard Error	*P* Value	Odds Ratio	Predictive Marginal
Intercept	-0.2087	0.0306	0.0000	0.8117	0.4071
Time of Call					
Evening – After 4PM	0.2986	0.0155	0.0000	1.3480	0.4509
Afternoon - Before 4PM (RC)	0.0000	0.0000	.	1.0000	0.3797
Day of Call					
Weekday	-0.1528	0.0171	0.0000	0.8583	0.3977
Weekend (RC)	0.0000	0.0000	.	1.0000	0.4341
Controlled Access Barrier?					
Yes	-0.3901	0.0249	0.0000	0.6770	0.3315
No (RC)	0.0000	0.0000	.	1.0000	0.4216
Building Type					
Single Unit (RC)	0.0000	0.0000	.	1.0000	0.4306
Apartment/Condo 2-9 units	-0.4142	0.0244	0.0000	0.6609	0.3341
Apartment/Condo 10-49 units	-0.4423	0.0332	0.0000	0.6426	0.3280
Apartment/Condo 50+ units	-0.3264	0.0351	0.0000	0.7215	0.3537
Other Housing type	-0.0152	0.0528	0.7730	0.9849	0.4269
Population Density					
≥1 million (MSA) (RC)	0.0000	0.0000	.	1.0000	0.4056
250,000 to 1 million (MSA)	0.0130	0.0199	0.5147	1.0131	0.4087
< 250,000 (MSA)	0.0406	0.0271	0.1340	1.0414	0.4153
Non-rural, Non-MSA	0.0539	0.0257	0.0355	1.0554	0.4185
Rural, Non-MSA	-0.0496	0.0241	0.0398	0.9517	0.3940
Interviewer Experience (Years)	-0.0150	0.0032	0.0000	0.9852	.
Percent 50 and older in Segment	0.0094	0.0129	0.4675	1.0094	.
Percent Black, Non-Hispanic in Segment, 12 and older	0.0008	0.0003	0.0242	1.0008	.
Percent Hispanic in Segment, 12 and older	0.0026	0.0005	0.0000	1.0026	.
Median Housing Value	0.0000	0.0000	0.4354	1.0000	.

(RC): Reference Category
Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004.

Even accounting for the control variables, the model shows that evenings were 1.35 times better for obtaining a contact on a first call than afternoons (significant at p<.05), or about 7 percentage points higher in terms of predictive margins. Weekdays were less productive in terms of contact than weekends (only .86 times as productive – also significant at p<.05). As expected, certain building characteristics were also associated with lower contact rates (namely the presence of controlled access and non-single unit structure status).[2] We also find evidence of higher contact rates on the first call for segments with larger proportions of minorities (Black Non-Hispanics and Hispanics). Somewhat surprisingly, we found that when controlling for other variables, more experienced interviewers had lower rates of contact than those with less experience, although the effect is not especially large. For each additional year of NSDUH experience, the probability of contact is only reduced by 0.375 percent for individuals with a predicted probability of 50 percent.

---

[2] Data collection on controlled access features in the NSDUH began with the 2004 survey. Some evidence suggests that some interviewers may have recorded the presence or absence of controlled access in a way that would lead to bias in the estimated effects of controlled access on contact. If interviewers were more likely to record the presence of controlled access features when access was actually blocked and less likely to report it when they were not blocked, the effects of controlled access on cooperation would be overstated. The other regression coefficients would be affected as well, by the correlations between the other variables and the controlled access measure.

It is important to note that our analyses are not based on a randomized design in which the values of independent variables are randomly assigned to cases. Interviewers are not randomly assigned to segments and interviewers do not randomly determine their call times. For example, our estimate of the effect of calling during the evening rather than the daytime on contact would be biased if other factors strongly associated with calling time are also strongly related to contact and not included in our regression model. While we feel that we have controlled for a number of variables that are associated with both calling times and the likelihood of contact, there is always the possibility that we have omitted relevant variables that could confound the relationship between calling time and contact.

Our findings on the effects of calling time (evening vs. daytime and weekend vs. weekday) on the likelihood of contact are generally consistent with results from other studies. Using data from the Family Resources Survey, an in-person survey of households in Great Britain, Purdon, Campanelli and Sturgies (1999) report that contact rates were the highest for weekday evening attempts (defined as calling after 5PM) and lowest for weekday morning and afternoons. Similarly, Groves and Couper (1998) report higher contact rates for evening call attempts (defined as after 6PM) than for daytime call attempts using data from the National Survey of Health and Stress. Bates (2003) reports the highest contact rates for weekend attempts with similar rates for weekday evening attempts (defined as after 5PM) and lower rates of contact for weekday mornings and afternoons.

The finding that the presence of controlled access barriers reduces the probability of contact on the first call is consistent findings from other studies as well. Groves and Couper (1998) carried out logistic regression analyses predicting contact on the first call attempt using data from the NSHS and report that the presence of an access impediment has a negative effect on contact. Using data from Cycle 6 of the National Survey of Family Growth (conducted in 2002), Groves and Ziniel (2003) report a contact rate on the first call of 50 percent for households without access impediments and 42 percent for those with such impediments.

The differences in contact rates between evening and daytime calling, raise the possibility that the allocation of attempts between daytime and evening times could be due to nonrandom factors, which in turn could be correlated with survey items of interest (in the case of NSDUH, substance use estimates). For example, safety concerns may lead some interviewers to prefer to make their first attempts in the daytime rather than the evening. If these concerns were also strongly correlated with the prevalence of substance use in their segments, the allocation of calling times could lead to bias in survey estimates of substance use. In their analysis of data from the NSHS, Groves and Couper (1998) found little evidence that interviewers were less likely to make evening calls, controlling for environmental characteristics such as urban status, population density, crime rates (in the county), percentage minority and units in disrepair in the neighborhood. As a first step, we estimated a logistic regression model predicting whether the first call attempt was made in the evening or not (after 4PM), using the same predictors as in the model predicting contact. The results are shown in *Table 7*.

Table 7   **Logistic Regression Model of Evening Attempt (After 4PM) for First Call Characteristics, Quarters 3 and 4, 2004 NSDUH**

Variable	Coefficient	Standard Error	P Value	Odds Ratio	Predictive Marginal
Intercept	-0.3716	0.0273	0.0000	0.6896	0.3867
Controlled Access Barrier?					
Yes	0.2928	0.0368	0.0000	1.3402	0.4423
No	0.0000	0.0000	.	1.0000	0.3774
Building Type					
Single Unit (RC)	0.0000	0.0000	.	1.0000	0.3959
Apartment/Condo 2-9 units	-0.0803	0.0269	0.0028	0.9228	0.3737
Apartment/Condo 10-49 units	-0.1754	0.0422	0.0000	0.8391	0.3558
Apartment/Condo 50+ units	-0.1513	0.0515	0.0033	0.8596	0.3613
Other Housing type	0.1327	0.0920	0.1491	1.1419	0.4089
Controlled Access by Building Type					
Controlled Access, Single Unit	0.0000	0.0000	.	1.0000	0.4542
Controlled Access, Apt/Condo 2-9 units	-0.0874	0.0600	0.1455	0.9163	0.4137
Controlled Access, Apt/Condo 2-9 units	0.0050	0.0683	0.9412	1.0050	0.4130
Controlled Access, Apt/Condo 2-9 units	0.0054	0.0719	0.9402	1.0054	0.4189
Controlled Access, Other housing type	-0.4584	0.1149	0.0001	0.6323	0.3765
No Controlled Access, Single Unit	0.0000	0.0000	.	1.0000	0.3841
No Controlled Access, Apt/Condo 2-9 units	0.0000	0.0000	.	1.0000	0.3656
No Controlled Access, Apt/Condo 2-9 units	0.0000	0.0000	.	1.0000	0.3441
No Controlled Access, Apt/Condo 2-9 units	0.0000	0.0000	.	1.0000	0.3495
No Controlled Access, Other housing type	0.0000	0.0000	.	1.0000	0.4155
Population Density					
≥1 million (MSA) (RC)	0.0000	0.0000	.	1.0000	0.4142
250,000 to 1 million (MSA)	-0.0926	0.0199	0.0000	0.9115	0.3921
< 250,000 (MSA)	0.0609	0.0269	0.0238	1.0628	0.4289
Non-rural, Non-MSA	-0.4681	0.0267	0.0000	0.6262	0.3077
Rural, Non-MSA	-0.2758	0.0246	0.0000	0.7590	0.3498
Interviewer Experience (Years)	-0.0268	0.0032	0.0000	0.9735	.
Percent 50 and older in Segment	0.0070	0.0120	0.5617	1.0070	.
Percent Black, Non-Hispanic in Segment, 12 and older	0.0054	0.0003	0.0000	1.0054	.
Percent Hispanic in Segment, 12 and older	0.0032	0.0005	0.0000	1.0032	.
Median Housing Value	0.0000	0.0000	0.0041	1.0000	.

(RC): Reference Category
Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004.

Interviewers are more likely to make their first call attempts after 4PM than before 4PM in smaller MSAs (less than 250,000 persons (relative to MSAs with populations over 1,000,000) and in segments with higher percentages of Black, non-Hispanic and Hispanic populations. They are less likely to make such attempts in non-MSAs and MSAs with between 250,000 and 1 million persons than in MSAs with over 1,000,000 persons. More experienced interviewers (as measured by years of NSDUH experience) are less likely to make evening attempts than less experienced interviewers, although this may be an artifact of more experienced interviewers having to travel greater distances more frequently.

We also find that the relationship between controlled access and making first attempts after 4PM depends upon the type of dwelling unit. For single housing units, interviewers were more likely to make their first attempt after 4PM if there was an access barrier than if there were no such barriers. For apartment/condo type dwelling units, interviewers were less likely to make attempts during the evening when there were no access barriers present than

when there were barriers.[3] There is also evidence of differential treatment of other housing types (e.g. dormitories, group quarters, nursing homes,) depending on whether or not there are access barriers. For these types of housing, interviewers were somewhat less likely to make first attempts during the evening if there were access barriers and more likely to have made evening attempts when there were no such barriers. Overall however, the effects of these variables on the likelihood of making the first call attempt after 4PM as opposed to before 4PM appear fairly small. The predictive margins associated with the variables in the model are not all that different from each other.[4]

Analyses of Interviewer Effort and Survey Estimates

At this point, we have evidence indicating that evening call attempts are more productive than daytime call attempts on the first attempt and that calling times are not randomly determined. We are long way from concluding that interviewers are not efficiently allocating their time in making their first call attempts. But suppose such inefficiencies exist. Do these inefficiencies lead to higher levels of nonresponse which in turn lead to increased bias in survey estimates? If there are no differences between respondents and nonrespondents on the survey items of interest, inefficiencies in gathering data would have implications for costs but not for the accuracy of survey estimates. In this section we begin to address this question by using measures of interviewer effort in terms of the numbers of call attempts and call days, along with other information from the survey.

Different interviewers and supervisors have different work habits and patterns that can lead to variation in the expenditure of effort required to complete screenings and interviews. While the ideal is to complete screenings and interviews with each sampled dwelling unit and respondent in the most efficient manner possible, there are often barriers to achieving pre-determined response rate goals. Because completed screenings and interviews cannot always be obtained, the question becomes whether all cases and segments are worked to the point of "adequate finalization." That is, is there a set of disposition patterns that indicate a case or segment has been worked to the point where further effort would not change the final outcome?

In analyses of data from the Survey of Consumer Finances, Arthur Kennickell (2003, 2004) raises the possibility that interviewers may sequence their work assignments in such a way as to maximize response rates while keeping costs low. Given a fixed amount of time to complete assignments, interviewers may schedule their work in such a way that cases expected to be "easy" to complete are worked first with more difficult cases postponed until later during the field period. In short, cases are ordered by the probability of response and cases with the highest probability of response are worked first or given more attention than cases with lower probabilities of response. If this practice takes place on a large scale, it could be problematic if survey items of interest are also related to the propensity to respond.

We might expect that if interviewers are able to effectively implement this strategy of maximizing response rates while minimizing effort (and keeping costs low), easy to interview cases will be interviewed in early attempts while later attempts would be characterized by increasing proportions of final refusal cases. *Figure 3* shows how screener cases that were unresolved at a given call attempt ultimately wound up being resolved as a complete, a refusal, as other nonresponse (of which two-thirds are final noncontacts) and as ineligible. While the proportion of unresolved cases that wind up as final refusals increases sharply between the first and sixth attempts, the rate of refusal at that point flattens out around 16 percent. In contrast, there is a much sharper tendency for unresolved cases to wind up as other types of nonresponse as the number of call attempts increase. By about the 15th screener attempt, the chances that a case will be resolved as a completed screener are about equal to the chances that it will resolve as some form of nonresponse.

---

[3] The finding that interviewers are more likely to make evening attempts for single dwelling units when access barriers are present than when absent is consistent with a finding by Groves and Ziniel (2003) that cases with access impediments were more likely to be called in the evening on the first call than cases without such impediments.

[4] We do not wish to imply that interviewers make decisions on a dwelling unit basis on whether to make the first attempt during the evening or not. In principle, interviewers cannot know about the building type or presence of controlled access barriers prior to making their first attempts. But as noted earlier, experienced interviewers are often assigned to segments they have worked before and as such, are likely to have at least a rough idea of the types of buildings and prevalence of access barriers in their segments.

The nature of how the probability of refusal first increases in the early call attempts and then flattens out after the 10th call attempt could reflect a combination of respondent willingness to participate in the survey as well as interviewer willingness to exert effort. It may be that hostile refusals are being resolved in the earlier call attempts while soft refusals and other cases in which the interviewer feels that it is worth the effort to follow up are resolved only after many more call attempts. However, before reaching such a conclusion, we would need to carry out further analyses that account for whether or not contact ever occurred and the first call attempt at which contact occurred.

Figure 3   Percentages of Active Screener Cases Ultimately Resolving as Complete, Refusal, Other Nonresponse, or Ineligible by Call Attempt
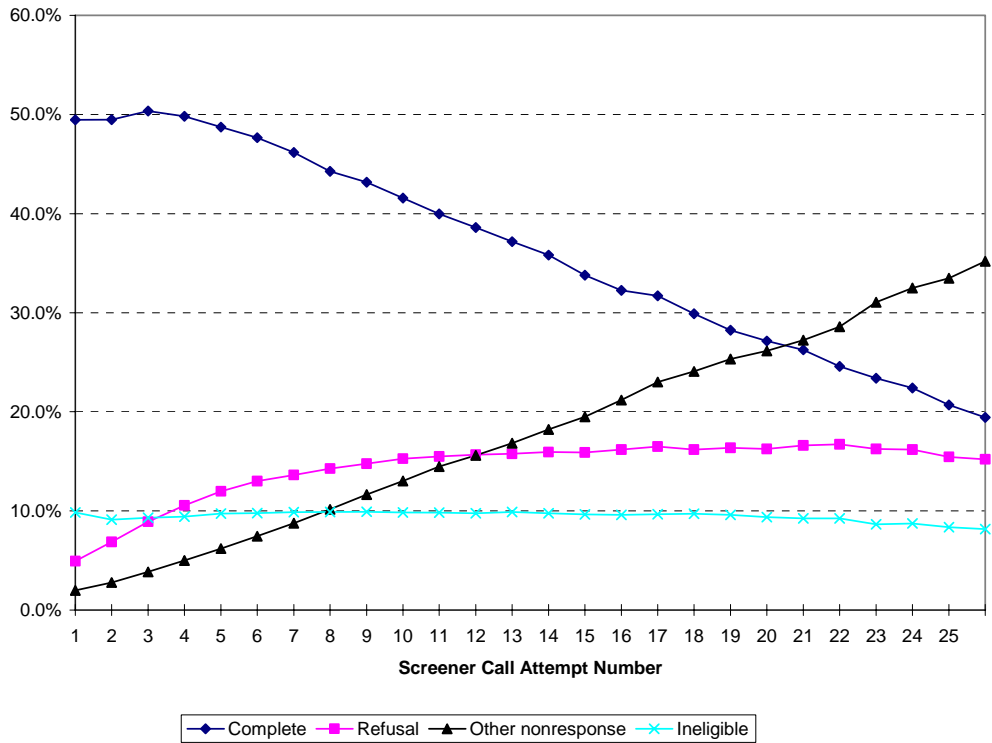
Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004.

Figure 4 presents similar information for interview attempts (but with only three categories of final result codes, complete, refusal and other nonresponse). In the case of interviewing, there is an immediate decline in the proportion of unresolved cases that will eventually be completed after the first call attempt. In fact, by about the fifth call attempt, the chances of resolving as a completed case are lower than resolving as nonresponse. Despite the different nature of the screener and interview requests, the percentages of cases that wind up as refusals for unresolved interview cases is similar in shape to that for screener refusals (a sharp increase through the first six or so call attempts followed by a constant rate thereafter). Also, as with the screener refusals, the proportion of interview cases that resolve as some other type of nonresponse increases steadily as more call attempts are applied.

Figure 4 **Percentages of Active Interview Cases Ultimately Resolving as Complete, Refusal, or Other Nonresponse by Call Attempt**
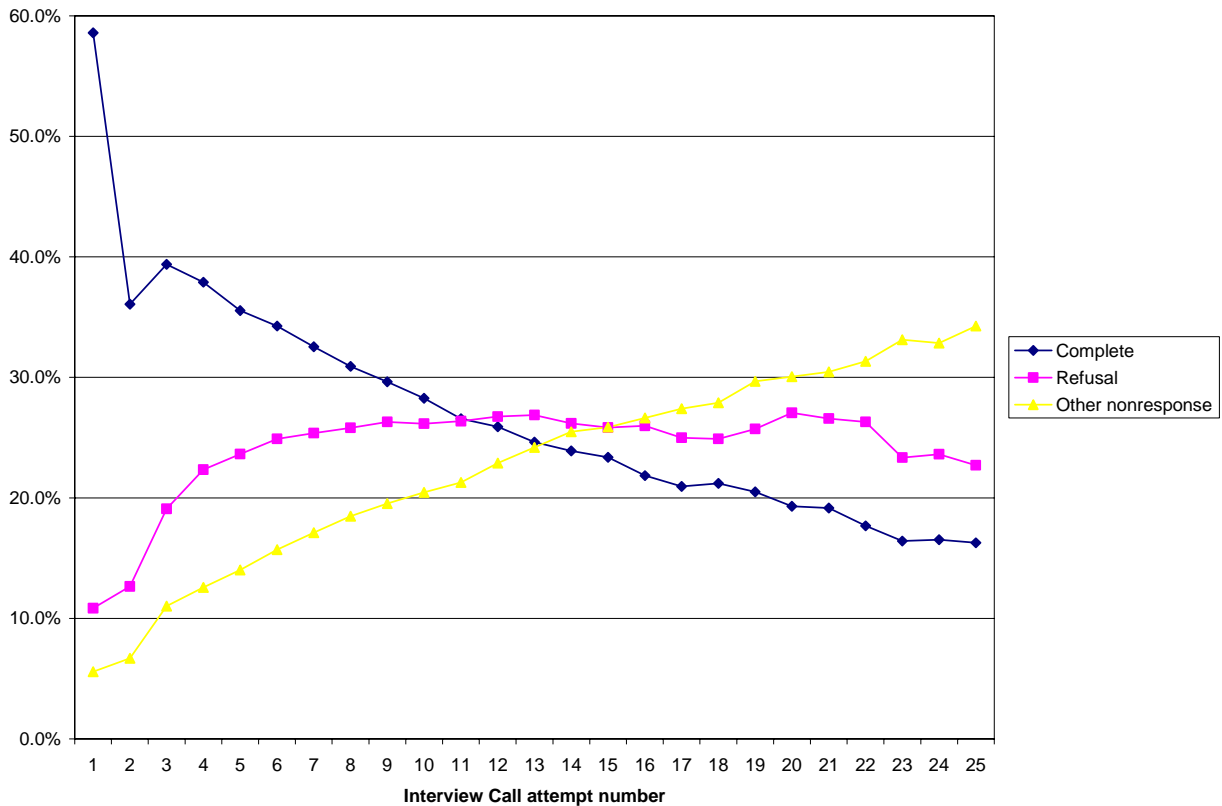
Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004.

In short we find that as the number of screening and interviewing call attempts increases, unresolved cases are increasingly likely to resolve as other types of nonresponse rather than refusals. Since a substantial proportion of the other nonresponse category consists of final noncontact cases, the probability of resolving as other nonresponse should increase as the number of call attempts increase. The number of call attempts or call days can be viewed as a proxy indicator of respondent unavailability to be interviewed due to time constraints (e.g. respondents who are generally not at home should require higher numbers of call attempts than those who are usually at home).

For further evidence on the nature of the refusal and other nonresponse components for the NSDUH, we produced estimates for some key survey items based on two indicators of difficulty of obtaining an interview which could reflect respondent unwillingness to be interviewed, difficulty in contacting respondents or interviewer willingness to exert effort to gain the interview. Ideally, we would like to have measures that allow us to separate out the portions of these measures that are due to respondent resistance and interviewer persistence. We created measures of respondent resistance based on 1) number of call days required to complete the screener interview, 2) number of additional call days required to complete the interview and 3) whether or not the respondent ever refused to do the interview. Estimates in these tables use selection weights; adjustments for nonresponse and population weighting adjustments (poststratification) are excluded in the computation of these weights.

Tables 8, 9, 10 and 11 present estimates for some key NSDUH survey items, lifetime, past year and past month use of any illicit drugs and use of marijuana for children and adults, respectively, by numbers of call days required for screening and interviewing. In general, those who did not require any additional interview attempts were less likely to report substance use than those who required additional interview attempts, more so for the lifetime measures than the past year and past month measures.

Table 8   **Estimates of Illicit Drug and Marijuana Use Among Adults, 18 and older by Number of Screener Call Days Required to Complete Screener**

| Measure | Screener Call Days Required to Complete Screener | | | | | | |
	1	2	3	4	5 to 9	10+	Total
Any Illicit Drug Use: Lifetime	46.30	49.46	49.28	51.94	51.63	49.61	48.36
	0.57	*0.75*	*1.07*	*1.52*	*1.23*	*2.84*	*0.40*
Any Illicit Drug Use: Past Year	13.70	13.94	13.67	14.46	15.09	16.01	13.97
	0.33	*0.45*	*0.56*	*0.88*	*0.78*	*1.84*	*0.23*
Any Illicit Drug Use: Past Month	7.72	7.81	8.20	7.24	8.04	8.07	7.81
	0.26	*0.33*	*0.48*	*0.55*	*0.51*	*1.23*	*0.18*
Marijuana Use: Lifetime	41.38	44.21	44.25	46.89	46.73	45.00	43.35
	0.56	*0.75*	*1.00*	*1.53*	*1.21*	*2.80*	*0.39*
Marijuana Use: Past Year	10.02	10.11	10.33	10.42	11.15	11.24	10.24
	0.28	*0.36*	*0.47*	*0.67*	*0.67*	*1.48*	*0.19*
Marijuana Use: Past Month	5.90	5.87	6.43	5.58	6.25	6.37	5.98
	0.22	*0.29*	*0.40*	*0.48*	*0.43*	*1.07*	*0.15*

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004. Estimates are weighted using selection weights. Standard errors are in italics and were calculated using linearization methods.

Table 9   **Estimates of Illicit Drug and Marijuana Use Among Children, 12 to 17, by Number of Screener Call Days Required to Complete Screener**

| Measure | Screener Call Days Required to Complete Screener | | | | | | |
	1	2	3	4	5 to 9	10+	Total
Any Illicit Drug Use: Lifetime	30.49	28.90	30.04	30.68	29.41	30.44	29.95
	0.55	*0.77*	*1.08*	*1.44*	*1.11*	*3.62*	*0.35*
Any Illicit Drug Use: Past Year	21.24	20.55	20.56	22.06	19.96	19.65	20.90
	0.51	*0.67*	*0.98*	*1.35*	*0.98*	*3.24*	*0.32*
Any Illicit Drug Use: Past Month	10.35	10.47	10.39	10.23	11.04	9.90	10.43
	0.37	*0.51*	*0.73*	*0.90*	*0.79*	*2.61*	*0.24*
Marijuana Use: Lifetime	19.56	18.27	18.04	18.80	19.53	21.20	19.01
	0.50	*0.67*	*0.90*	*1.24*	*1.05*	*3.23*	*0.32*
Marijuana Use - Past Year	14.62	13.80	13.57	15.12	14.56	15.47	14.32
	0.44	*0.57*	*0.82*	*1.15*	*0.94*	*2.88*	*0.28*
Marijuana Use - Past Month	7.56	7.29	7.02	7.09	7.73	9.27	7.43
	0.34	*0.42*	*0.55*	*0.77*	*0.73*	*2.51*	*0.20*

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004. Estimates are weighted using selection weights. Standard errors are in italics and were calculated using linearization methods.

Table 10   **Estimates of Illicit Drug and Marijuana Use Among Adults, 18 and older by Number of Additional Call Days Required for Interview**

	Additional Call Days Required to Complete Interview							
Measure	**0**	**1**	**2**	**3**	**4**	**5 to 9**	**10+**	**Total**
Any Illicit Drug Use: Lifetime	45.53	51.53	50.53	49.02	48.39	49.72	50.47	48.36
	0.52	*0.67*	*1.23*	*1.74*	*2.28*	*2.56*	*1.62*	*0.40*
Any Illicit Drug Use: Past Year	13.64	14.06	14.96	13.41	15.45	15.80	13.63	13.97
	0.33	*0.39*	*0.71*	*1.00*	*1.32*	*1.64*	*1.00*	*0.23*
Any Illicit Drug Use: Past Month	7.73	7.67	8.45	8.01	7.84	8.46	7.74	7.81
	0.25	*0.31*	*0.55*	*0.72*	*0.94*	*1.16*	*0.67*	*0.18*
Marijuana Use: Lifetime	40.70	46.15	45.83	43.87	43.68	43.52	45.86	43.35
	0.51	*0.67*	*1.17*	*1.69*	*2.18*	*2.53*	*1.63*	*0.39*
Marijuana Use: Past Year	10.14	10.12	10.77	10.14	10.10	11.46	10.48	10.24
	0.27	*0.33*	*0.58*	*0.86*	*0.98*	*1.35*	*0.89*	*0.19*
Marijuana Use: Past Month	5.85	5.85	6.67	6.39	5.73	6.67	6.20	5.98
	0.20	*0.25*	*0.48*	*0.65*	*0.71*	*1.03*	*0.62*	*0.15*

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004. Estimates are weighted using selection weights. Standard errors are in italics and were calculated using linearization methods.

Table 11   **Estimates of Illicit Drug and Marijuana Use Among Children, 12 to 17, by Number of Additional Call Days Required for Interview**

|	Additional Call Days Required to Complete Interview							
Measure	**0**	**1**	**2**	**3**	**4**	**5 to 9**	**10+**	**Total**
Any Illicit Drug Use: Lifetime	28.37	31.38	33.64	37.11	33.13	36.34	33.87	29.95
	0.45	*0.70*	*1.49*	*2.28*	*2.60*	*3.74*	*2.72*	*0.35*
Any Illicit Drug Use: Past Year	19.52	22.45	24.39	25.60	21.66	23.25	25.35	20.90
	0.39	*0.63*	*1.37*	*2.07*	*2.47*	*3.20*	*2.55*	*0.32*
Any Illicit Drug Use: Past Month	9.65	11.34	11.63	11.87	12.74	14.68	14.12	10.43
	0.30	*0.49*	*1.01*	*1.41*	*1.96*	*2.79*	*1.99*	*0.24*
Marijuana Use: Lifetime	17.11	21.13	23.59	23.19	22.62	28.74	23.68	19.01
	0.41	*0.60*	*1.34*	*2.08*	*2.51*	*3.51*	*2.44*	*0.32*
Marijuana Use: Past Year	12.85	16.08	17.90	16.90	17.22	19.27	18.76	14.32
	0.34	*0.56*	*1.27*	*1.82*	*2.29*	*2.98*	*2.26*	*0.28*
Marijuana Use: Past Month	6.64	8.26	9.40	7.83	10.53	11.43	10.16	7.43
	0.26	*0.42*	*0.93*	*1.12*	*1.82*	*2.51*	*1.72*	*0.20*

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004. Estimates are weighted using selection weights. Standard errors are in italics and were calculated using linearization methods.

Table 12 provides estimates for these same items by whether or not there was ever a pending refusal recorded in that interviews call history. In contrast to the estimates based on additional call days, adults who temporarily refused to do the interview were generally less likely to report substance use than those who never refused to do the interview. For children, there were no significant differences on these substance use measures between those with an interim refusal in their call history and those without an interim refusal. We should note however that those who temporarily refuse to do the interview but eventually go on to complete the survey are a very small fraction of all completed interviews. In 2004, only about 2.4 percent of all completed interviews had temporarily refused to do the interview. Thus, we should be cautious in inferring that interim refusals could be proxies for those who ultimately refuse to do the survey.

Table 12   **Estimates of Illicit Drug and Marijuana Use Among Adults (18 and older), by Ever Having Refused the Interview**

	Adults			Children		
	Ever Refused Interview?			Ever Refused Interview?		
Measure	No	Yes	Total	No	Yes	Total
Any Illicit Drug Use: Lifetime	48.61	42.49	48.36	29.96	29.42	29.95
	0.40	*2.04*	*0.40*	*0.35*	*3.78*	*0.35*
Any Illicit Drug Use: Past Year	14.07	11.82	13.97	20.91	20.65	20.90
	0.23	*1.12*	*0.23*	*0.32*	*3.16*	*0.32*
Any Illicit Drug Use: Past Month	7.87	6.42	7.81	10.44	10.04	10.43
	0.18	*0.82*	*0.18*	*0.25*	*2.36*	*0.24*
Marijuana Use: Lifetime	43.56	38.38	43.35	19.01	18.58	19.01
	0.39	*2.01*	*0.39*	*0.32*	*3.08*	*0.32*
Marijuana Use: Past Year	10.33	8.02	10.24	14.30	17.31	14.32
	0.20	*0.88*	*0.19*	*0.28*	*3.03*	*0.28*
Marijuana Use: Past Month	6.04	4.51	5.98	7.41	9.43	7.43
	0.15	*0.68*	*0.15*	*0.21*	*2.32*	*0.20*

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004. Estimates are weighted using selection weights. Standard errors are in italics and were calculated using linearization methods.

Finally, we produced selection weighted estimates for the lifetime, past year and past month measures of any illicit drug use that cumulated over interviewed cases based on the number of screener and interview call days. In these estimates, we examine the degree to which overall survey estimates change as a function of the number of screener and interview call days required to complete screening and interviewing. These are plotted in *Figures 5 – 10* for additional interview call days. For the upper and lower bounds, we simply multiplied the design based standard error by two. In *Figure 5*, we see that with just one additional interview call day, the overall estimate for any illicit lifetime drug use among adults is quite similar to the final estimate based on the full sample. That is, the estimate based on only on interviews completed with no more than one additional call day is within the sampling error of the final estimate based on all cases.

In *Figure 6* however, we find that the estimate of any past year illicit drug use among adults, based on those who completed the interview immediately (without requiring any additional call days) is within sampling error of the final selection weighted estimate. This result is repeated in *Figure 7* for past month use of any illicit drug. *Figures 8 – 10* show similar results for child estimates for these same measures.

In short, it appears that only one additional interview call day is needed to obtain estimates for these items that are statistically similar to those obtained after much more extensive calling efforts. Naturally, there are a number of caveats worth noting. First, these estimates are only selection weighted estimates. In principle, to get the "true" estimates based on different numbers of completed interviews based on different levels of effort, we would have to reweight the retained cases (including nonresponse and poststratification adjustments) to examine the full impact of restricting the sample based on different levels of interviewing. We note however that the selection weighted estimates based on the full sample are very similar to estimates in which the weights use nonresponse and population weighting adjustments.

Figure 5   **Cumulative Percentages of Adults (18+) Reporting Any Illicit Lifetime Drug Use By Additional Interview Call Days**
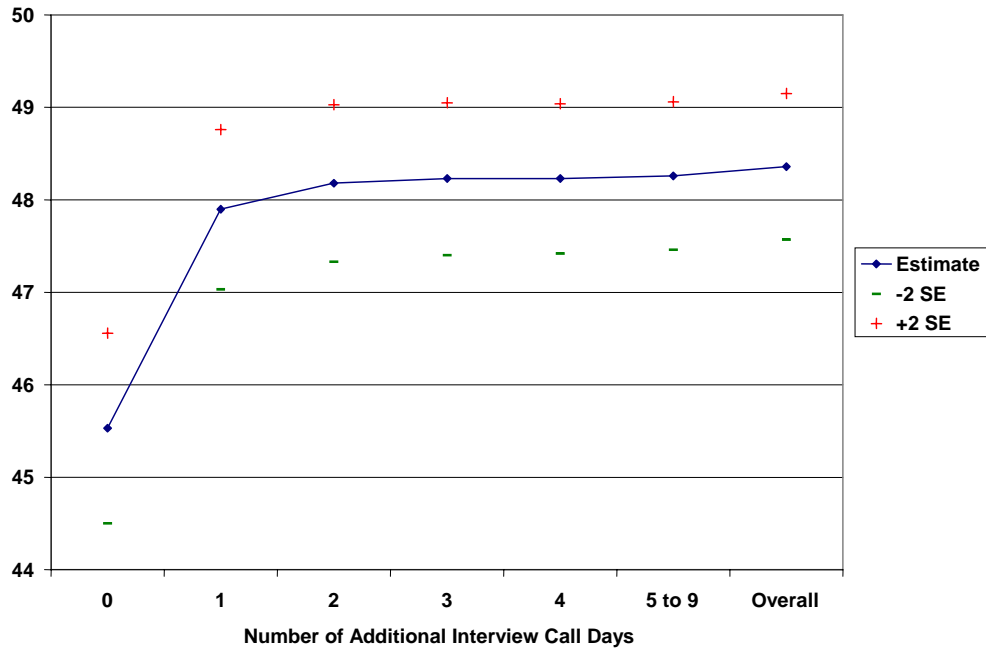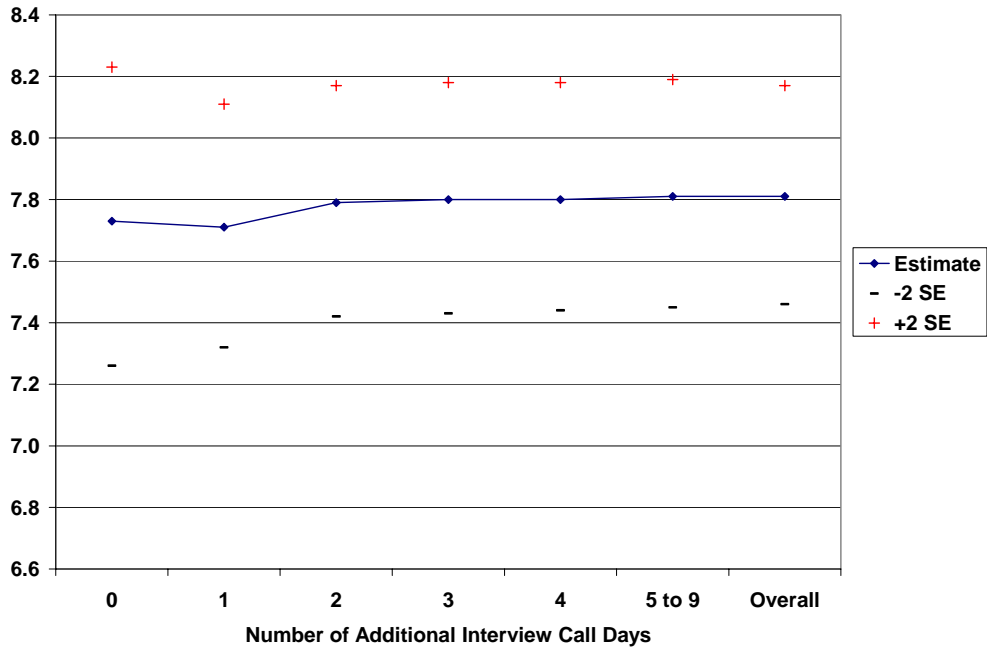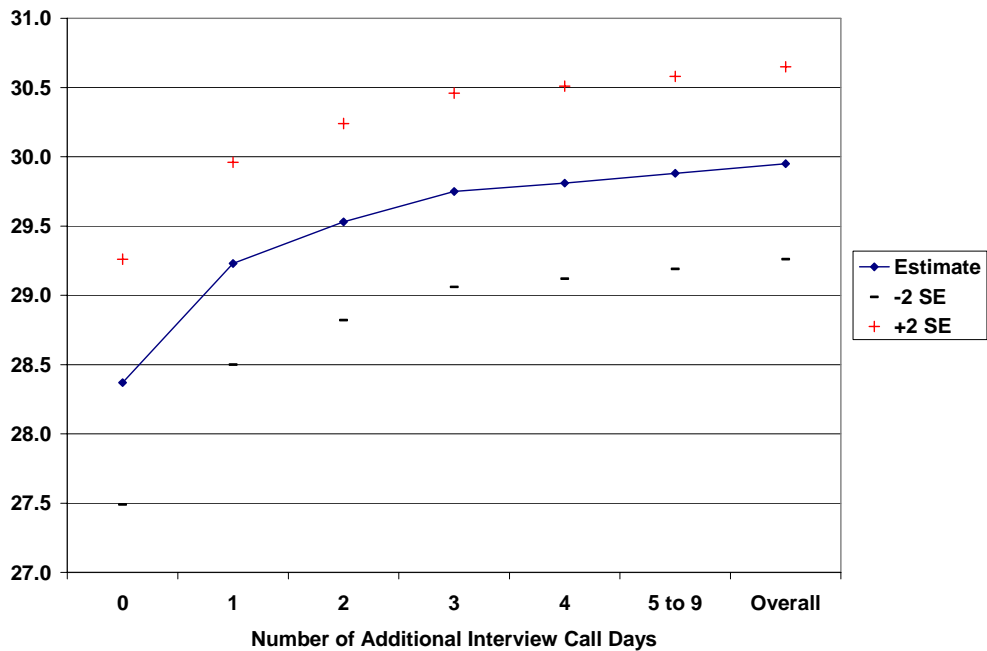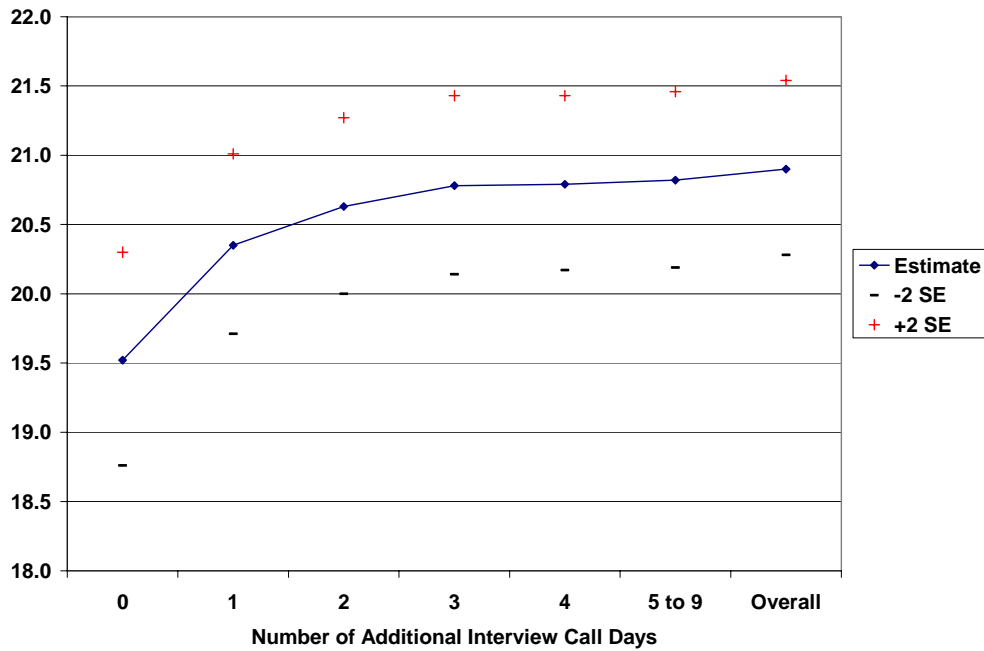

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004. Estimates are weighted using selection weights.

Figure 6   **Cumulative Percentages of Adults (18+) Reporting Any Illicit Past Year Drug Use By Additional Interview Call Days**
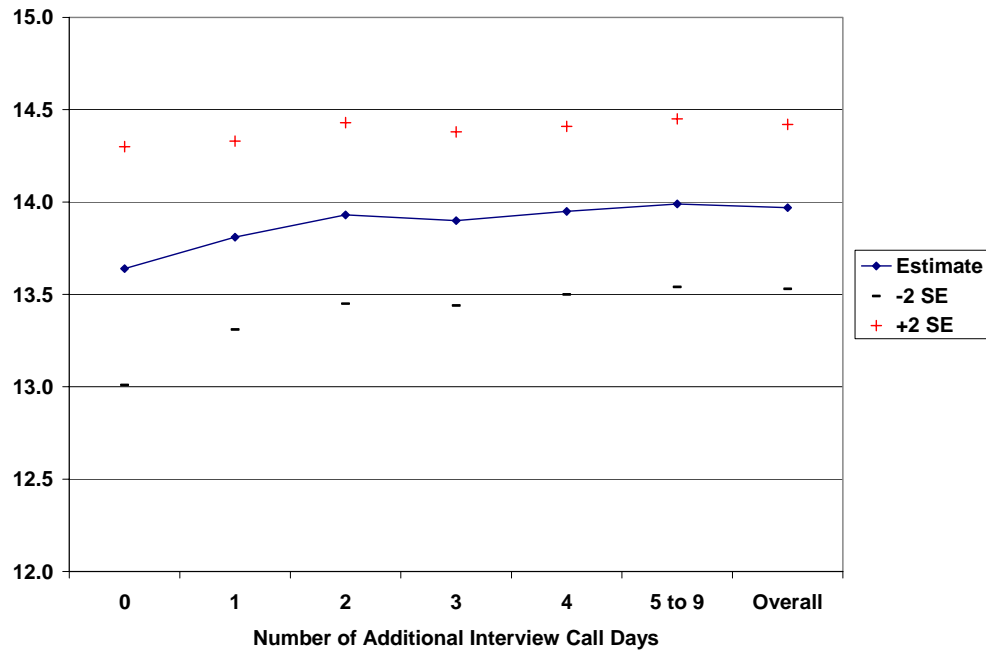
Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004. Estimates are weighted using selection weights.

Figure 7   **Cumulative Percentages of Adults (18+) Reporting Any Illicit Past Month Drug Use By Additional Interview Call Days**
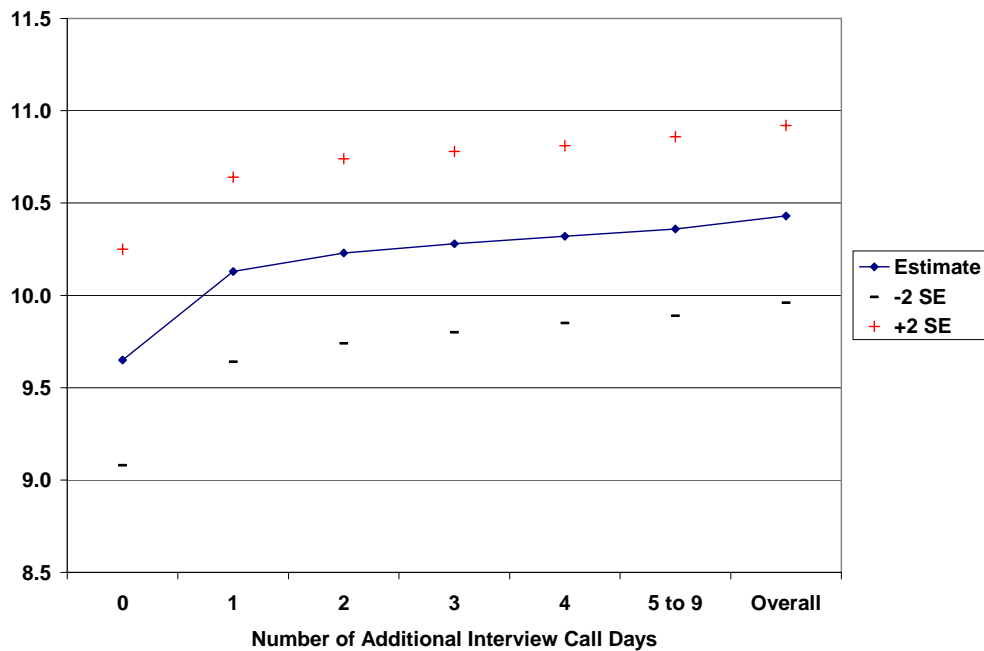

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004. Estimates are weighted using selection weights.

Figure 8   **Cumulative Percentages of Children (12-17) Reporting Any Illicit Lifetime Drug Use By Additional Interview Call Days**

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004. Estimates are weighted using selection weights.

Figure 9   **Cumulative Percentages of Children (12-17) Reporting Any Illicit Past Year Drug Use By Additional Interview Call Days**

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004. Estimates are weighted using selection weights.

Figure 10   **Cumulative Percentages of Children (12-17) Reporting Any Illicit Past Month Drug Use By Additional Interview Call Days**

Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004. Estimates are weighted using selection weights.

With this preliminary evidence that some estimates do not change substantially with extensive interviewing efforts, we examined the relationship between the numbers of call attempts and response rates on NSDUH interviewing. Response rates at the interview level (unweighted), and cumulative percentages of all interviewing call attempts and cumulative percentages of completed interviews are plotted against the number of call attempts *Figure 11*.

Figure 11   2004 NSDUH Response Rates and Cumulative Percentages for Call Attempts, Completed Interviews
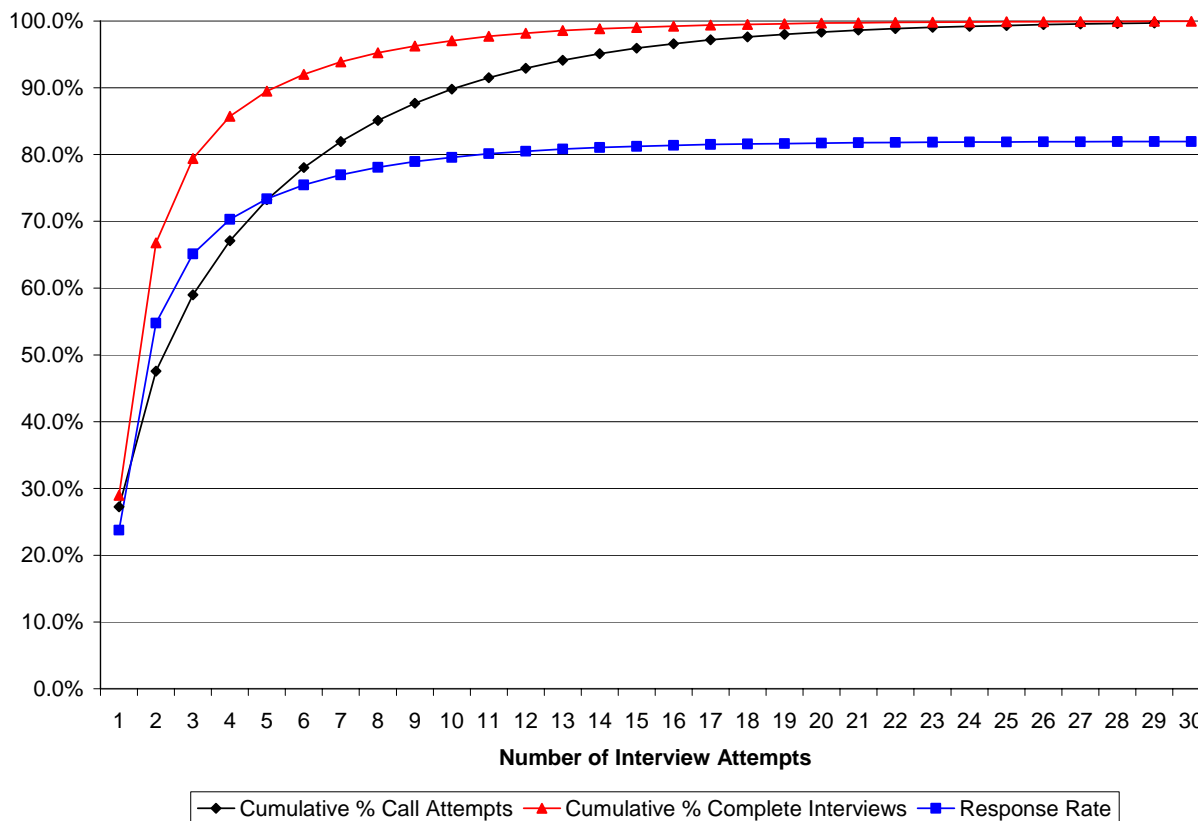
Source: SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, 2004.

Figure 11 shows that a considerable number of attempts are taking place in order to achieve relatively small increases in completed interviews and response rates. Overall, almost 305,000 call attempts were made in the interviewing process for the 2004 NSDUH. If interviewing were stopped after the sixth call attempt, the result would be about 62,000 interviews, or about 92 percent of actual number of interviews completed. The unweighted response rate would be about 76 percent (compared to the final response rate of 82 percent). About 238,000 interview attempts would be needed, about 22 percent less than the total of 305,000 interview attempts. Cutting off cases after the ninth interview attempt would produce about 96 percent of the completed interviews (about 65,000 interviews) with a 79 percent response rate. About 267,000 interview attempts would be required, which is about 12 percent less than the total number of interviewing call attempts made during the 2004 NSDUH.

Conclusion

In this paper, we have taken the first steps in analyzing NSDUH process data to address questions regarding interviewer efforts and effects on response rates and survey estimates. We have found that calling times, defined by the time of the call (before or after 4PM) and the day of the week (weekday vs. weekend) are related to contact on the first attempt for the screener. We suspect that the same is true for cooperation on the screener but we have not carried out this analysis yet. We also found evidence that using less intensive follow up efforts would not

necessarily lead to survey estimates that differ appreciably from estimates obtained with greater effort. Furthermore, reduction of interviewing effort on a per case basis, at least on the surface, could lead to reductions in data collection costs.

While we could limit the number of call attempts, we have to be wary of changing survey procedures such that interviewers change their work habits in unanticipated ways. For example, if interviewers were actually instructed not to follow up as much with cases or make far fewer attempts, this might convey a message that response rates are not important and lead to even lower response rates than we would be willing to accept. We should also note that the results of the effort analysis do not imply that there is no bias due to nonresponse; only that for the measures examined in this paper, we get about the same substantive results from a lower amount of effort (allowing for just a few call days) than exerting a lot of effort. Nonrespondents to the survey could still have characteristics quite different from respondents. Still, we think it is worth exploring the implementation of a system to track key survey estimates as field work progresses in order to provide managers with information about when cases or segments could be considered adequately finalized. Heeringa and Groves (2004) discuss the application of such a system in the National Survey of Family Growth (NSFG).

We understand that there are certain limits on the degree to which interviewers can allocate their time to evening as opposed to daytime attempts. Interviewers have access to information, sometimes at the individual dwelling unit level that we cannot observe, that they use to determine opportune times to make their call attempts. In addition, there a number of personal factors regarding work habits that also affect when interviewers decide to make attempts that may be beyond the control of supervisors and field managers. Finally, there are constraints that interviewers must work within that limit the degree to which they can work in the evenings. For example, interviewers that need to travel greater distances to reach their segments may not be able work very much after 6PM unless they stay overnight in the segment, which in turn increases costs. In future analyses, our goal is to take into account more of the factors in the field that can confound relationships between calling times and outcomes such as contact and cooperation rates.

References

Bates, N., (2003) Contact Histories in Personal Visit Surveys: The Survey of Income and Program Participation (SIPP) Methods Panel. *Proceedings of the Survey Research Methods Section, American Statistical Association*, Alexandria, VA, pp. 7-14.

Groves, R.M., & Couper, M.P. (1998). *Nonresponse in household interview surveys*. New York: Wiley.

Groves, R.M. & Ziniel, S.I. (2003) Using Interviewer Observations as Predictors of Contactability in U.S. Face to Face Surveys. Paper presented at the 2003 International Workshop on Household Survey Nonresponse.

Heeringa, S.G. & Groves, R.M. (2004) Responsive Design for Household Surveys. *Proceedings of the Survey Research Methods Section, American Statistical Association*, Alexandria, VA, pp. 3644-3651.

Kennickell, A.B. (2003) Reordering the Darkness: Application of Effort and Nonresponse Bias in the Survey of Consumer Finances. *Proceedings of the Survey Research Methods Section, American Statistical Association*, Alexandria, VA, pp. 2119-2126.

Kennickell, A.B. (2004) Action at a Distance: Interviewer Effort and Nonresponse in the SCF. *Proceedings of the Survey Research Methods Section*, *American Statistical Association*, Alexandria, VA, pp. 3774-3781.

Painter, D., Chromy, J.R., Meyer, M., Granger, R.A. & Clarke, A. (2003) Effects of Incentives on Data Collection: A Record of Calls Analysis of the National Survey on Drug Use and Health. *Proceedings of the Survey Research Methods Section, American Statistical Association*, Alexandria, VA, pp. 170-176.

Purdon, S., Campanelli, P. & Sturgis, P. (1999) Interviewers' Calling Strategies on Face-to-Face Interview Surveys. *Journal of Official Statistics,* 15: 199-216.