

IMPROVEMENT OF RESPONSE RATES ON THE  
MANUFACTURERS' SHIPMENTS, INVENTORIES, AND  
ORDERS SURVEY WITH CATI

Douglas Bond and Richard Sigman

ABSTRACT

We used CATI follow-up for nonresponse to improve response rates to mailed questionnaires in a nine-month study of the Manufacturers' Shipments, Inventories, and Orders (M3) survey. Return rates increased 30 to 40 percentage points for a group of chronic nonrespondents, and reporting was more timely. In a group of good, but late, reporters, return rates improved an average of 20 and 40 percentage points in time for the advance and preliminary M3 reports, respectively, and imputation rates declined. Interviewers needed a little over three calls and 10 to 11 minutes logged on to the computer (including about six minutes of conversation) to complete a CATI case by the end of the study. Recommendations for CATI use and future research in the M3 survey are presented.

KEY WORDS: Manufacturing statistics, response rates, CATI

1. INTRODUCTION

Each month, the Census Bureau, through its Manufacturers' Shipments, Inventories, and Orders (M3) survey, obtains measures of current industrial activity and indicators of future production from U.S. manufacturers. Survey results are released in an advance report on a preset date that is three to four weeks after the end of the reference month. A preliminary report, released on a preset date about a week later, contains revised figures and more detail than the advance report. The Bureau issues final revised data for the reference month in the next month's reports.

The M3 survey is a principal Federal economic indicator. Government and business use M3 data to assess current and future economic conditions in the domestic manufacturing sector, develop economic policy, and estimate important components of the Gross Domestic Product. These data users need M3 results that are timely and accurate; therefore, the Bureau continually strives to reduce the size of revisions by obtaining responses from a high proportion of manufacturers in time for the advance and preliminary reports. In this report, we describe research using Computer Assisted Telephone Interviewing (CATI) for mail nonresponse follow-up, to improve reporting for two categories of M3 sample units: chronic nonrespondents, manufacturers who have not reported for at least 12 months; and late reporters, manufacturers who usually report data after the preliminary report has been released.

The Bureau selected CATI for this study for several reasons:

- CATI is a well-established method. Interviewers generally like it and response rates are not adversely affected by it (Weeks 1992).
- M3 analysts would not have to phone chronic nonrespondents and late reporters, because the Bureau's Hagerstown, Maryland Telephone Center (HTC) would do the CATI interviewing.
- It would eliminate key entry of data by M3 staff, reducing costs and speeding processing of data.
- CATI has the potential for improving timeliness and reducing data collection costs relative to traditional paper-and-pencil interviewing (PAPI) by telephone; however, studies have not conclusively shown that CATI is always the most cost-effective and timely option relative to PAPI (Nicholls and Groves 1986).
- Studies have shown that CATI can improve the quality of data relative to PAPI methods (Weeks 1992; Sigman et al. 1993).

CATI posed some possible disadvantages for the M3 survey: the personal contact between M3 analysts and respondents would be lost, which could hurt response rates and affect data quality, since CATI interviewers are not subject matter experts; and costs could go up, since the M3 Branch would have to pay HTC for data collection and overhead.

Our objectives in this study were to address the above concerns and determine the effects of introducing CATI as another mode of data collection for chronic nonrespondents and late reporters. Specifically, we wanted information on the following:

- Response rates and timeliness of responses.
- Reasons for refusals.
- Data quality.
- CATI calling characteristics - number of calls and length of interview per case.

In section 2 of this report, we describe the M3 survey more fully. Section 3 talks about the CATI system we used. In section 4, we describe the experimental design, and present results in section 5. Section 6 is a discussion, and section 7 presents recommendations.

## 2. THE M3 SURVEY

The M3 survey is a monthly nonprobability survey for which nearly all U.S. manufacturing companies, with \$500 million or more in annual shipments, voluntarily provide data. Smaller companies also report data to improve coverage in selected industry categories. The industry categories fall into one of two major groups: durable goods, such as motor vehicles, aircraft, computer equipment, and furniture; and nondurable goods, such as apparel and food products. A total of about 3,000 company units report each month. The reporting unit is usually all operations of a company that operates in only one industry category; but large, diversified companies often have multiple reporting units.

Respondents report five or seven items, depending on the industry category of the reporting unit. For most nondurable goods categories, there are five items: value of shipments, total inventory, and inventories by stage of fabrication (materials and supplies, work-in-process, and finished goods). For durable goods categories and a few nondurable goods categories, there are two more items: new and unfilled orders.

A link relative procedure estimates monthly shipments, unfilled orders, and total inventory. In this procedure, the universe level for the previous month is multiplied by the percent change in level from the previous to current month for reporting units; the result is the current month's estimate of universe level. The Bureau derives new orders by adding the previous to current month change in unfilled orders to the estimated shipments for the current month. The link relative procedure is not used because some companies do not completely report new orders every month. A modified link relative procedure is used to estimate inventories by stage of fabrication, at a higher level of aggregation than for other items. Then the estimates by stage are proportionally adjusted to sum to estimated total inventory. This modified procedure is necessary because some reporting units cannot break their inventories into stages.

The Bureau mails all questionnaires near the end of the reference month to the reporting units, who return them by mail or fax. M3 analysts phone all the largest nonrespondents and selected smaller reporting units who do not respond. Most units are followed up by phone if they do not respond for two consecutive months. Analysts also phone those units whose data look questionable.

M3 processing is in three cycles:

- Cycle 1 begins two to three weeks after the end of the reference month, and lasts about a week. The Bureau releases the advance report at 8:30 AM EST the day after cycle 1 ends; it only has estimates for durable goods industry categories.

- Cycle 2 ends about a week after cycle 1. The preliminary report is released at 10:00 AM EST the day after cycle 2 ends, and it contains estimates for durable and nondurable goods industry categories.
- Cycle 3 ends about two weeks later, when cycle 1 for the next reference month begins. There is no report after cycle 3.

The Bureau cannot use data received during cycle 3, or later (the "late period"), in the advance or preliminary reports for the reference month; however, these data are used in preparing estimates for the next reference month's reports.

More information on the M3 survey is available in the Bureau's M3 publication (U.S. Bureau of the Census 1993).

### 3. THE CATI SYSTEM

CATI interviewers called from the HTC, a centralized facility that uses a PC-based version of the Bureau's CATI system. The facility opened in 1985 and has about 60 interviewing stations. Interviewers used automated call scheduling and case management systems for the M3 survey, features which may improve the productivity of interviewers (Nicholls and Groves 1986).

The CATI instrument was form-based; that is, the table for entering shipments, inventories, and orders appeared on one screen. There were two on-line edit checks: current month data were compared with previous month data; and total inventory was compared with the sum of inventories by stage of fabrication. CATI data went through the same batch edits as M3 data collected by other modes.

The CATI system produced a history file which contained information on every call: the sample (case) ID; date; four times (logon to the case, respondent's phone ringing, hang up of respondent's phone, and logoff to case); and outcome codes (indicating completion, refusal by reason, callback scheduled, supervisor action, and so forth).

### 4. EXPERIMENTAL DESIGN

#### 4.1 Chronic Nonrespondents

We took several steps to allocate M3 units to control and treatment groups. We identified 342 chronic nonrespondents, units who had not reported for at least 12 months. We determined, on the basis of a power analysis (see the appendix), that the control group should have at least 52 units. We actually selected a random sample of 75 units, because we wanted to avoid conflicts with another experiment being conducted on the M3 survey that overlapped with this study. The remaining 267

units made up the treatment group. After adjusting to maintain consistency with the other experiment, and dropping several units who were special analyst contacts, we had 80 control units and 259 treatment units (a total of 339 units).

We began collecting data for this experiment for December 1991. The M3 Branch had been mailing questionnaires to the chronic nonrespondents every month, and continued to do so during the experiment. There was no follow-up for nonresponse in the control group. But in the treatment group, CATI interviewers followed up nonrespondents, beginning about three weeks after the end of the reference month, and attempted to contact all of them for 10 to 14 days. The CATI system assigned the highest priority to units whose data were needed for the advance report. We used December 1991 data collection as a feasibility test and did not load all the data into the database, so we could not compute response rates and evaluate data quality for that month. But the CATI system did generate a CATI history file, with which we computed lengths of interviews and tabulated reasons for refusals. We did have a complete data set beginning with collection of January 1992 data. We continued this experiment through collection of August 1992 data.

#### 4.2 Late Reporters

We identified 105 M3 units who were "late reporters," units who in the previous six months: (1) had reported at least four times; and (2) had reported at most three times before the release of the preliminary report. Most of these units reported during cycle 3 or later. Our power analysis indicated that the control and treatment groups should be the same size, so we randomly assigned 52 units to the control and 53 to the treatment. After reallocation to conform with our other experiment (as described in the previous section), we had 54 treatment units and 51 control units. Our treatment and control procedures were identical to those described for the chronic nonrespondents.

### 5. RESULTS

#### 5.1 Response Rates

To define our response rates, we constructed a framework of response and nonresponse components that is similar to the framework presented by Hidiroglou et al. (1992). In Figure 1, "Mailed" (box 2) includes out-of-scope and refusal units from previous months, to whom we did not actually mail a questionnaire (or contact them by CATI) in subsequent months. "Not mailed" units (box 3) were taken over by another unit or split into two or more units; we surveyed their successors. "Unresolved" units (box 5) could not be contacted, so we do not know if they are out-of-scope. "Out-of-scope" units (box 7) include out-of-

business, nonmanufacturing, and duplicate units. We assume "Refusals" (box 9) are in-scope. Units are "Responses" (boxes 10-13) when their data are entered in our database.

We computed two measures of response rates, by mode of data collection and survey cycle: the "return rate" and the "conservative response rate." The return rate is defined as

$$(\text{Responses (box 8)})/(\text{Mailed (box 2)}).$$

This rate includes the effect of the quality of the frame on response rates, because out-of-scope and unresolved units are included in the denominator.

The conservative response rate is

$$(\text{Responses (box 8)})/(\text{Estimated eligibles}),$$

where

$$\text{Estimated eligibles} = (\text{Mailed (box 2)}) - (\text{Out-of-scopes (box 7)}) - (1 - p) * (\text{Unresolved (box 5)}),$$

where

$$p = (\text{In-scopes (box 6)})/(\text{Resolved (box 4)}).$$

$p$  is the proportion of in-scope units among resolved units, so  $(1 - p) * (\text{box 5})$  is an estimate of the number of out-of-scope units among the unresolved units. The response rate in this case is conservative because the estimated number of eligibles is larger than if all unresolved units are assumed to be out-of-scope. We computed  $p$  only in the treatment groups, and assumed it was the same in the control groups, which (as we shall see) had lower response rates and less data with which to compute  $p$ .  $p$  was 86-88% for the chronic nonrespondents and 100% every month for the late reporters.

We compared overall (mail, fax, CATI, and analyst phone calls) response rates for the treatment and control groups with a one-tailed  $z$  test. This tests the hypothesis that the treatment response rate is greater than the control response rate.

5.1.1 Chronic Nonrespondents. Table 1 shows that CATI follow-up significantly improved the overall return rate, every month and every cycle. The improvement by the end of the late period was about 30 to 40 percentage points relative to the control rate, which rose gradually over the months and levelled off at about 5%. Most respondents provided data before the end of cycle 2, in time for the preliminary report. Mail and fax usually accounted for less than half of the overall rate.

CATI did a good job of identifying chronic nonrespondents who were refusals or out-of-scope, thereby "cleaning up" the sample units. Table 2 shows that by the end of the study, 47% of the CATI treatment units were classified as refusals, and 14% as out-of-scope, versus 2% and 4%, respectively, of the control units. Only 4% of the treatment group units were unresolved for August 1992 data, compared with 89% of the control group units. Section 5.2 presents a breakdown of reasons for refusals among the chronic nonrespondents.

Conservative response rates had the same patterns as the return rates (see Table 3). They were a little higher because of the removal of out-of-scope units from the denominator.

5.1.2 Late Reporters. Overall return rates were significantly higher for most cycles in the CATI treatment group of the late reporters, and they were over 90% every month in this group by the end of the late period (see Table 4). The average improvements were 20 percentage points by the end of cycle 1 and 40 percentage points by the end of cycle 2, relative to the control group. Therefore, far more data were available for the advance and preliminary reports because of CATI follow-up calls. Combined mail and fax usage by the end of cycle 2 had a slight upward trend in the treatment group, from an average of about 26% for January and February to an average of about 36% for July and August. So apparently, CATI follow-up prompted respondents to report a little earlier by these modes. The combined mail and fax return rate by the end of the late period did not show a definite upward trend; it averaged about 34% for January and February, and about 38% for July and August. Importantly, it did not decline during the experiment, indicating that respondents did not become dependent on a CATI call to report their data.

CATI follow-ups did not increase the refusal rate, and they left fewer units unresolved than in the control group; for example, 2% of the treatment units were unresolved for August, versus 16% of the control units (see Table 5). The conservative response rates were identical to the return rates in the treatment group, because there were no out-of-scope units in this group, and the relationship to the control group rates was similar to what occurred with return rates (see Table 6).

## 5.2 Reasons for Refusals

We tabulated reasons for refusals to CATI interviews by examining outcome codes in the CATI history files. There were five possible outcomes:

- No interview progress:
  - (1) Do not do voluntary surveys

- (2) Staff shortage
- (3) Doing less government reporting
- (4) Other
- (5) Refusal after interview started

Note that because we only looked at refusals to CATI interviewers, the refusal rates in this section are not comparable to refusal rates shown in Tables 2 and 5. In a few cases, CATI interviewers coded a case a refusal that an M3 analyst later decided was not a refusal. Also, refusals via mail, fax, or analyst phone call, though rare, were not included in the CATI history file.

Refusal rates to CATI calls were initially very high among chronic nonrespondents (31.5%), and declined each month as refusals were identified and eliminated from contact in subsequent months (see Table 7). Overall, 134 cases were refusals. The major reason was that companies did not do voluntary Census Bureau surveys (61.2% of all refusals). The next most common reasons were "other" (25.4% of all refusals), staff shortage (9.7%), and "doing less government reporting" (2.2%). Only 1.5% of the refusals occurred after the interview started.

The late reporters accounted for only six refusals over nine months of calling, so we do not show the breakdown by type of refusal for that experiment.

### 5.3 Data Quality

The return/response rates we presented in section 5.1 are indirect indicators of data quality. When more data are reported, fewer values have to be imputed, or less adjustment is necessary by reweighting estimates. We computed imputation rates to see how they were affected by the higher return/response rates seen with CATI follow-up for nonresponse.

In the M3 survey, not all missing values are replaced by imputed values; in general, imputation only occurs for items that had a value the previous month (reported or imputed). Values are computer-generated or selected by M3 analysts. The imputation rate is the number of unreported values that were imputed for the current and/or prior month, and used in the current month's estimate of month-to-month change, divided by the number of values that were used to compute the estimate. In most cases, the imputation rate was much lower in the CATI treatment group of the late reporters than in the control group, reflecting the higher return rates in the treatment group (see Table 8). We do not show imputation rates for the chronic nonrespondents because,

with the low return rates in the control group, the imputation rates are based on too little data to be of any analytical value.

#### 5.4 CATI Calling Characteristics

Data presented in this section are for cases that interviewers completed (obtained data), and exclude cases that were refusals, out-of-scope, inaccessible, and so forth. Interviewers took an average of 5.2 calls and 21 minutes to complete a chronic nonrespondent's case for the first month's data collection (see Table 9). By the end of the experiment, a case took an average of a little over three calls and 10 to 11 minutes to complete; about six minutes of this time was spent actually talking to someone on the phone.

Interviewers were able to complete cases for late reporters with fewer calls (an average of 3.6) and in less time (an average of 12 minutes) the first month than for the chronic nonrespondents (see Table 10). By the end of the experiment, calling characteristics were very similar to those of chronic nonrespondents, because interviewers had established contacts with reporters and were experienced with the survey.

## 6. DISCUSSION

CATI follow-up improved response rates and the timeliness of reporting for both longtime nonrespondents and good, but late, reporters. It did not increase refusal rates among the late reporters, and it was effective in resolving nonrespondent cases. Reporters did not seem to become dependent on CATI to report, since mail and fax rates did not decline in either experiment. Therefore the loss of contact between M3 analysts and respondents did not hurt response rates. Our only measures of data quality, response rates and imputation rates, indicated that adding CATI as another mode of data collection probably improved quality because more actual data were reported. Unfortunately, we were not able to measure data quality in other ways, such as edit failure rates, because data flags were not designed to capture the necessary information. These additional measures of data quality could have given us more information on the effects of CATI interviewing.

The improved reporting came at a price: HTC currently charges about \$12 per M3 case assigned to them. This covers interviewing and overhead costs and it does eliminate the costs (unknown, unfortunately) of an M3 analyst locating and contacting nonrespondents and M3 staff keying data.

CATI is a valuable tool for M3 data collection. More information is needed on its effects on data quality and its costs relative to other modes of data collection.

## 7. RECOMMENDATIONS

CATI should be used as another mode of data collection in the M3 survey. The Bureau should measure its effects on data quality by tracking edit failure rates and the frequency of inconsistent or blank data entries, for CATI and the other modes. The Bureau should also collect data to determine the costs of the different M3 data collection modes.

## REFERENCES

Hidioglou, M.A., Drew, J.D., and Gray, G.B. (1992), "Dealing with Nonresponse to Sample Surveys at Statistics Canada," Proceedings of the 1992 Annual Research Conference, Washington, DC: U.S. Bureau of the Census.

Nicholls II, W.L. and Groves, R.M. (1986), "The Status of Computer-Assisted Telephone Interviewing: Part I - Introduction and Impact on Cost and Timeliness of Survey Data," Journal of Official Statistics, 2, 93-115.

Sigman, R., Dorinski, S., and Tinari, B. (1993), "Comparison of Telephone Interviewing Methods in the Quarterly Apparel Survey," Proceedings of the International Conference on Establishment Surveys, Alexandria, VA: American Statistical Association (in press).

U.S. Bureau of the Census (1993), Manufacturers' Shipments, Inventories, and Orders: 1982-1992, Washington, DC: U.S. Government Printing Office.

Weeks, M.F. (1992), "Computer-Assisted Survey Information Collection: A Review of CASIC Methods and Their Implications for Survey Operations," Journal of Official Statistics, 8, 445-465.

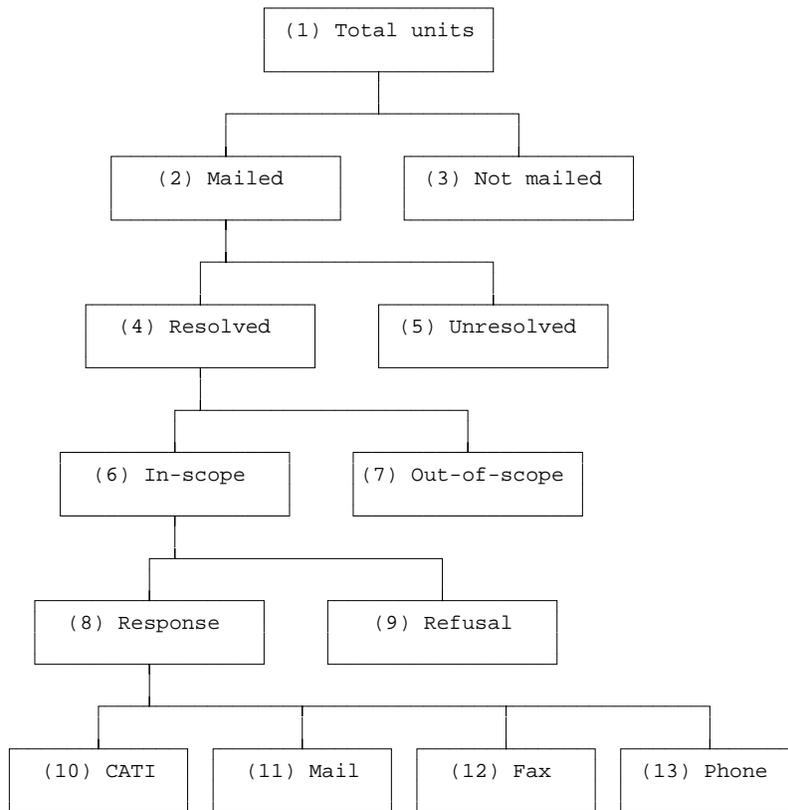


Figure 1: Components of Response and Nonresponse

Table 1: Return Rates (%) by Month for Which Data Were Collected, Chronic Nonrespondents

Month	Treatment	n <sup>1</sup>	Mode of data collection and cycle by which data were loaded into database																			
			Mail				Fax				CATI				Phone				All <sup>2</sup>			
			1	2	3	L	1	2	3	L	1	2	3	L	1	2	3	L	1	2	3	L
Jan 92	CATI	259	4	4	5	5	1	2	2	2	8	27	32	32	0	0	0	0	13*	32*	40*	40*
	Control	80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Feb 92	CATI	259	8	8	9	10	4	4	4	4	6	22	25	25	0	0	0	0	18*	34*	37*	39*
	Control	79	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
Mar 92	CATI	259	7	7	9	9	4	4	5	5	9	20	24	24	1	1	1	1	20*	32*	39*	39*
	Control	79	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
Apr 92	CATI	259	7	9	10	10	5	6	6	6	8	19	20	20	2	2	2	2	23*	36*	38*	39*
	Control	79	1	2	2	4	0	0	1	1	0	0	0	0	0	0	0	0	1	2	4	5
May 92	CATI	259	10	12	12	12	4	4	4	5	2	17	18	18	2	2	2	2	18*	36*	37*	37*
	Control	79	0	1	2	2	2	2	2	2	0	0	0	0	0	0	0	0	2	4	5	5
Jun 92	CATI	259	8	11	12	12	4	5	5	5	4	14	18	18	2	2	2	2	18*	32*	37*	37*
	Control	79	0	1	5	5	1	1	1	1	0	0	0	0	0	0	1	1	1	2	8	8
Jul 92	CATI	259	11	12	12	12	7	7	7	7	9	15	18	18	0	1	1	1	27*	34*	38*	38*
	Control	79	2	2	2	2	2	2	2	2	0	0	0	0	0	0	0	0	5	5	5	5
Aug 92	CATI	259	13	13	13	13	5	6	6	6	9	13	15	15	2	2	2	2	30*	34*	36*	36*
	Control	79	2	2	2	2	1	1	1	1	0	0	0	0	0	0	1	1	4	4	5	5

<sup>1</sup> n is the number "mailed out."

<sup>2</sup> \* indicates CATI percent greater than control (0.05 significance level).

Note: percents may not add to "all" due to rounding.

Table 2: Components of Response and Nonresponse (%) by Month for Which Data Were Collected, Chronic Nonrespondents

Month	Treatment	n <sup>1</sup>	Returned	Refusal	Out-of-scope	Unresolved
Jan 92	CATI	259	40	38	11	11
	Control	80	0	1	0	99
Feb 92	CATI	259	39	43	12	6
	Control	79	1	1	0	98
Mar 92	CATI	259	39	45	12	3
	Control	79	2	1	1	95
Apr 92	CATI	259	39	46	12	3
	Control	79	5	1	1	92
May 92	CATI	259	37	46	12	5
	Control	79	5	1	1	92
Jun 92	CATI	259	37	46	13	4
	Control	79	8	1	1	90
Jul 92	CATI	259	38	46	14	2
	Control	79	5	1	2	91
Aug 92	CATI	259	36	47	14	4
	Control	79	5	2	4	89

<sup>1</sup> n is the number "mailed out."

Note: percents may not add to 100 due to rounding.

Table 3: Conservative Response Rates (%) by Month for Which Data Were Collected, Chronic Nonrespondents

Month	Treatment	n <sup>1</sup>	Mode of data collection and cycle by which data were loaded into database																			
			Mail				Fax				CATI				Phone				All			
			1	2	3	L	1	2	3	L	1	2	3	L	1	2	3	L	1	2	3	L
Jan 92	CATI	228	4	4	6	6	1	2	2	2	10	30	37	37	0	0	0	0	15	37	46	46
	Control	70	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Feb 92	CATI	226	9	10	10	12	4	4	4	4	7	25	28	28	0	0	0	0	20	39	43	44
	Control	69	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
Mar 92	CATI	226	8	8	11	11	4	5	5	5	11	23	28	28	1	1	1	1	23	37	45	45
	Control	68	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
Apr 92	CATI	226	8	11	12	12	6	7	7	7	10	22	23	23	2	2	2	2	26	41	44	44
	Control	69	2	3	3	4	0	0	2	2	0	0	0	0	0	0	0	0	2	3	4	6
May 92	CATI	225	12	14	14	14	5	5	5	5	2	20	21	21	2	3	3	3	20	41	42	43
	Control	68	0	2	3	3	3	3	3	3	0	0	0	0	0	0	2	2	3	4	6	6
Jun 92	CATI	224	9	13	13	13	5	6	6	6	4	16	21	21	2	3	3	2	21	38	42	42
	Control	68	0	1	6	6	2	2	2	2	0	0	0	0	0	2	2	0	2	3	9	9
Jul 92	CATI	223	12	14	14	14	8	8	8	8	11	17	21	21	0	1	1	1	32	40	44	44
	Control	67	3	3	3	3	3	3	3	3	0	0	0	0	0	0	0	0	6	6	6	6
Aug 92	CATI	223	15	15	15	15	6	7	7	7	11	15	18	18	3	3	3	3	35	40	42	42
	Control	66	3	3	3	3	2	2	2	2	0	0	0	0	0	0	2	2	4	4	6	6

<sup>1</sup> n is the estimated number of eligible reporting units.  
 Note: percents may not add to "all" due to rounding.

Table 4: Return Rates (%) by Month for Which Data Were Collected, Late Reporters

Month	Treatment	n <sup>1</sup>	Mode of data collection and cycle by which data were loaded into database																			
			Mail				Fax				CATI				Phone				All <sup>2</sup>			
			1	2	3	L	1	2	3	L	1	2	3	L	1	2	3	L	1	2	3	L
Jan 92	CATI	54	11	18	22	28	6	7	9	9	26	50	56	56	0	0	0	0	43*	76*	87*	93*
	Control	51	18	24	55	57	2	4	6	8	0	0	0	0	0	0	0	0	20	29	71	74
Feb 92	CATI	54	15	20	22	22	4	7	9	9	22	54	63	63	0	0	0	0	41 <sup>n</sup>	82*	94*	94*
	Control	51	24	33	55	69	4	4	4	8	0	0	0	0	0	0	0	0	27	37	59	76
Mar 92	CATI	54	7	13	13	13	13	17	18	18	24	48	63	63	0	0	0	0	44*	78*	94*	94*
	Control	51	10	26	59	65	4	8	12	12	0	0	0	0	4	4	4	4	18	37	74	80
Apr 92	CATI	54	17	26	28	28	2	4	6	6	28	54	59	60	0	0	0	2	46*	83*	93*	94*
	Control	51	22	37	55	55	8	10	24	24	0	0	0	0	0	2	4	4	29	49	82	82
May 92	CATI	54	20	24	24	24	15	18	18	18	4	46	52	52	0	0	0	0	39 <sup>n</sup>	89*	94*	94*
	Control	51	27	45	61	65	2	4	8	10	0	0	0	0	0	0	0	0	29	49	69	74
Jun 92	CATI	54	17	24	24	24	7	13	15	15	9	43	52	52	0	0	0	0	33 <sup>n</sup>	80*	91*	91*
	Control	51	16	33	47	49	8	14	18	18	0	0	0	0	2	2	2	2	26	49	67	69
Jul 92	CATI	54	24	24	24	24	7	9	9	9	28	50	56	56	2	2	4	4	61*	85*	93*	93*
	Control	51	20	31	43	47	2	10	20	20	0	0	0	0	0	0	0	0	22	41	63	67
Aug 92	CATI	54	18	20	20	20	15	18	20	24	26	43	50	50	0	0	0	0	59*	82*	91*	94*
	Control	51	20	27	27	49	10	14	16	20	0	0	0	0	2	2	4	4	31	43	67	73

<sup>1</sup> n is the number "mailed out."  
<sup>2</sup> \* indicates CATI % greater than control (0.05 significance level); <sup>n</sup> indicates CATI % not greater than control.  
 Note: percents may not add to "all" due to rounding.

Table 5: Components of Response and Nonresponse (%) by Month for Which Data Were Collected, Late Reporters

Month	Treatment	n <sup>1</sup>	Returned	Refusal	Out-of-scope	Unresolved
Jan 92	CATI	54	93	2	0	6
	Control	51	74	2	2	22
Feb 92	CATI	54	94	4	0	2
	Control	51	76	4	2	18
Mar 92	CATI	54	94	4	0	2
	Control	51	80	6	2	12
Apr 92	CATI	54	94	4	0	2
	Control	51	82	6	2	10
May 92	CATI	54	94	4	0	2
	Control	51	74	6	4	16
Jun 92	CATI	54	91	4	0	6
	Control	51	69	6	4	22
Jul 92	CATI	54	93	4	0	4
	Control	51	67	6	6	22
Aug 92	CATI	54	94	4	0	2
	Control	51	73	6	6	16

<sup>1</sup> n is the number "mailed out."

Note: percents may not add to 100 due to rounding.

Table 6: Conservative Response Rates (%) by Month for Which Data Were Collected, Late Reporters

Month	Treatment	n <sup>1</sup>	Mode of data collection and cycle by which data were loaded into database																			
			Mail				Fax				CATI				Phone				All			
			1	2	3	L	1	2	3	L	1	2	3	L	1	2	3	L	1	2	3	L
Jan 92	CATI	54	11	18	22	28	6	7	9	9	26	50	56	56	0	0	0	0	43	76	87	93
	Control	50	18	24	56	58	2	4	6	8	0	0	0	0	0	2	10	10	20	30	72	76
Feb 92	CATI	54	15	20	22	22	4	7	9	9	22	54	63	63	0	0	0	0	41	82	94	94
	Control	50	24	34	56	70	4	4	4	8	0	0	0	0	0	0	0	0	28	38	60	78
Mar 92	CATI	54	7	13	13	13	13	17	18	18	24	48	63	63	0	0	0	0	44	78	94	94
	Control	50	10	26	60	66	4	8	12	12	0	0	0	0	4	4	4	4	18	38	76	82
Apr 92	CATI	54	17	26	28	28	18	4	6	6	28	54	59	59	0	0	0	2	46	83	93	94
	Control	50	22	38	56	56	8	10	24	24	0	0	0	0	0	2	4	4	30	50	84	84
May 92	CATI	54	20	24	24	24	15	18	18	18	4	46	52	52	0	0	0	0	39	89	94	94
	Control	49	29	47	63	67	2	4	8	10	0	0	0	0	0	0	0	0	31	51	71	78
Jun 92	CATI	54	17	24	24	24	7	13	15	15	9	43	52	52	0	0	0	0	33	80	91	91
	Control	49	16	35	49	51	8	14	18	18	0	0	0	0	2	2	2	2	26	51	69	71
Jul 92	CATI	54	24	24	24	24	7	9	9	9	28	50	56	56	2	2	4	4	61	85	93	93
	Control	48	21	33	46	50	2	10	21	21	0	0	0	0	0	0	0	0	23	44	67	71
Aug 92	CATI	54	18	20	20	20	15	18	20	24	26	43	50	50	0	0	0	0	59	82	91	94
	Control	48	21	29	50	52	10	15	17	21	0	0	0	0	2	2	4	4	33	46	71	77

<sup>1</sup> n is the estimated number of eligible reporting units.

Note: percents may not add to "all" due to rounding.

Table 7: Refusals to CATI Calls, Chronic Nonrespondents

Item	Month for which data were collected									Total
	Dec 91	Jan 92	Feb 92	Mar 92	Apr 92	May 92	Jun 92	Jul 92	Aug 92	
CATI cases attempted										
- Number	257	146	107	89	80	75	72	67	69	962
Refused cases										
- Number	81	28	13	6	4	1	0	0	1	134
- % of attempted	31.5	19.2	12.1	6.7	5.0	1.3	0.0	0.0	1.4	-
<u>Refusals by type</u>										
Do not do voluntary surveys										
- Number	51	15	9	6	1	0	0	0	0	82
- % of refused cases	63.0	53.6	69.2	100.0	25.0	0.0	-	-	0.0	61.2
Staff shortage										
- Number	8	3	2	0	0	0	0	0	0	13
- % of refused cases	9.9	10.7	15.4	0.0	0.0	0.0	-	-	0.0	9.7
Doing less govt reporting										
- Number	2	0	1	0	0	0	0	0	0	3
- % of refused cases	2.5	0.0	7.7	0.0	0.0	0.0	-	-	0.0	2.2
Other, no interview progress										
- Number	19	9	1	0	3	1	0	0	1	34
- % of refused cases	23.5	32.1	7.7	0.0	75.0	100.0	-	-	100.0	25.4
Refusal after interview started										
- Number	1	1	0	0	0	0	0	0	0	2
- % of refused cases	1.2	3.6	0.0	0.0	0.0	0.0	-	-	0.0	1.5

Table 8: Imputation Rates (%) by Month for Which Data Were Collected, Late Reporters

Month	Treatment	Item <sup>1</sup>						
		VS	UO	NO	TI	MI	WI	FI
Jan 92	CATI	4	0	15	4	3	3	3
	Control	12	12	24	16	14	20	13
Feb 92	CATI	2	4	19	2	2	3	6
	Control	16	17	29	18	19	27	17
Mar 92	CATI	4	0	18	4	5	5	6
	Control	12	12	18	13	11	12	12
Apr 92	CATI	0	0	15	0	0	0	0
	Control	2	0	11	3	3	4	3
May 92	CATI	0	0	12	0	0	0	0
	Control	19	19	28	20	14	15	12
Jun 92	CATI	2	0	12	2	2	3	3
	Control	18	12	22	19	18	16	19
Jul 92	CATI	4	4	16	4	5	5	6
	Control	32	29	38	34	29	33	27
Aug 92	CATI	0	0	8	0	0	0	0
	Control	19	21	27	21	19	23	13

<sup>1</sup> VS - value of shipments; UO - unfilled orders; NO - new orders;  
 TI - total inventory; MI - materials and supplies inventory;  
 WI - work-in-process inventory; FI - finished goods inventory.

Table 9: CATI Calling Characteristics for Completed Cases, by Month for Which Data Were Collected, Chronic Nonrespondents

Month	Number of cases	Number of calls		Case time (minutes) <sup>1</sup>		Talk time (minutes) <sup>2</sup>	
		Mean	Max	Mean	Max	Mean	Max
Dec 91	102	5.2	18	21	95	-	-
Jan 92	87	4.7	14	16	87	-	-
Feb 92	73	3.8	14	12	58	-	-
Mar 92	66	3.8	17	14	51	8	38
Apr 92	60	3.0	8	10	35	6	22
May 92	57	3.2	12	10	34	5	19
Jun 92	55	3.3	11	10	40	5	23
Jul 92	54	3.1	12	11	51	6	21
Aug 92	45	3.4	17	11	45	7	18

<sup>1</sup> Case time is the time logged on to the computer to interview a case.

<sup>2</sup> Talk time is the time actually talking to someone on the phone, and is included in the case time. Not available until Mar 92 data.

Table 10: CATI Calling Characteristics for Completed Cases, by Month for Which Data Were Collected, Late Reporters

Month	Number of cases	Number of calls		Case time (minutes) <sup>1</sup>		Talk time (minutes) <sup>2</sup>	
		Mean	Max	Mean	Max	Mean	Max
Dec 91	37	3.6	11	12	35	-	-
Jan 92	33	3.3	7	9	16	-	-
Feb 92	40	3.4	8	8	31	-	-
Mar 92	36	3.4	13	12	32	6	18
Apr 92	37	3.2	11	9	33	5	15
May 92	32	3.5	10	10	22	5	10
Jun 92	33	3.1	11	10	30	5	11
Jul 92	30	3.0	8	10	27	5	11
Aug 92	34	3.3	7	11	25	6	16

<sup>1</sup> Case time is the time logged on to the computer to interview a case.

<sup>2</sup> Talk time is the time actually talking to someone on the phone, and is included in the case time. Not available until Mar 92 data.

APPENDIX  
Power Analysis for Treatment-Group Sample Sizes

A power analysis determines the sizes of treatment and control groups needed to reliably detect a specified treatment effect. We performed two power analyses: one for the chronic nonrespondents and the other for the late reporters. For both of these experiments, the treatment was the use of CATI for nonresponse follow-up, and the control was no follow-up.

In the M3 survey, all reporters with the same address are called a "family." In both experiments we randomly assigned families - not reporters - to either treatment or control so that families would not be part treatment and part control.

The test statistic of interest is

$$z = (R_2 - R_1)/S,$$

where

$R_i$  = the response rate for group  $i$  (1 for control group and 2 for treatment group) averaged over  $H$  months, and

$S^2$  = estimated variance of  $R_2 - R_1$ .

For the power analyses, we assumed  $E(R_1R_2) = E(R_1)E(R_2)$ , so that  $S^2 = S_1^2 + S_2^2$ , where  $S_i^2$  = estimated variance of  $R_i$ . Let  $\sigma^2 = E(S^2) = E(S_1^2) + E(S_2^2) = \sigma_1^2 + \sigma_2^2$ .

We also assumed that for each group the number of mailed questionnaires was the same for all  $H$  months. This was denoted by

$n_i$  = number of questionnaires mailed to group  $i$

$$= \sum_{j=1}^{m_i} n_{ij} ,$$

where

$m_i$  = number of families assigned to group  $i$ , and

$n_{ij}$  = number of questionnaires mailed to family  $j$ , group  $i$ .

For given values of  $m_i$  and  $n_{ij}$ , the maximum value of  $\sigma_i^2$  occurs when reporters within the same family either all respond or all fail to respond, and this occurs with probability 0.5. This maximum  $\sigma_i^2$  is

$$\sigma_i^2 = (.25)a_i/(Hn_i^2), \tag{1}$$

where

$$a_i = \sum_{j=1}^{m_i} n_{ij}^2 .$$

Let  $A = a_1 + a_2$ . (The value of  $A$  is independent of which families are assigned to treatment and which to control.) Then  $a_i \approx n_i A/n$  where  $n = n_1 + n_2$ . Substituting this into (1) gives

$$\sigma_i^2 = (.25)A/(Hn n_i).$$

Hence the maximum value for  $\sigma^2$  is approximately

$$\sigma^2 = .25 A(n_1^{-1} + n_2^{-1})/(Hn). \quad (2)$$

The test's power also determines a maximum value for  $\sigma^2$ . In particular, for a one-tailed test with significance level  $\alpha$ , the test will have asymptotic power of  $1 - \beta$  if

$$z_\alpha \sigma \leq \delta - z_\beta \sigma ,$$

where  $z_p$  is the 100(1 - p) percentile of the standard normal distribution, and  $\delta$  is the absolute value of the smallest treatment effect that one is trying to detect.

Thus, the maximum value of  $\sigma^2$  is

$$\sigma^2 = \delta^2 / (z_\alpha + z_\beta)^2 . \quad (3)$$

Equating (2) and (3), and substituting  $n_2 = n - n_1$ , yields a quadratic equation in  $n_1$ , whose smallest root is

$$n_1 = \frac{n - \sqrt{n^2 - A(z_\alpha + z_\beta)^2 / (H\delta^2)}}{2} . \quad (4)$$

We used  $H = 6$ ,  $\alpha = \beta = 0.1$ , and  $\delta = 0.1$  (i.e., a minimum detectable difference in response rates of 10 percentage points). For the chronic nonrespondents,  $n = 342$  and  $A = 552$ , which when we evaluated (4) yielded  $n_1 = 52$ .

For the late reporters,  $n = A = 105$ . When we substituted these values into (4), however, we obtained a negative radicand. This indicated that in the maximum  $\sigma^2$  case the desired power cannot be achieved. Therefore, we set  $n_1 = 52$  and  $n_2 = 53$  because this yields maximum power.