NONRESPONSE UNDER MANDATORY VS. VOLUNTARY REPORTING IN THE 1989 SURVEY OF POLLUTION ABATEMENT COSTS AND EXPENDITURES (PACE)

Daniel R. Tulp, Jr., C. Easley Hoy, Gary L. Kusch, and Stacey J. Cole, U.S. Bureau of the Census
C. Easley Hoy, Bureau of the Census, Rm. 2017-4, Washington, DC 20233

KEY WORDS: Split panel, establishment survey, conditioning

BACKGROUND

The Economic Area of the Census Bureau conducts economic censuses every 5 years and economic surveys annually, quarterly, and monthly. Many of these surveys are conducted under mandatory reporting authority by law, but some are conducted under voluntary reporting. The nonresponse rates to date for the Census Bureau's voluntary economic surveys are higher than those for other surveys, and there is great concern about the possible causes and remedies, as well as quality implications. Therefore, there was a need to design and implement an experiment to measure nonresponse under mandatory vs. voluntary reporting without the confounding effects from other reasons for nonresponse.

When such a test was being considered in 1989, there was no new survey to serve as the test vehicle for mandatory vs. voluntary reporting. However, in 1989 there was an ongoing annual mandatory survey, the Survey of Pollution Abatement Costs and Expenditures (PACE), being reselected based on the 1987 economic census of manufactures and the 1989 Annual Survey of Manufactures (ASM). The PACE collects pollution abatement expenditures and operating costs data. The actual test consisted of selecting a panel from the newly selected 1989 PACE sample, collecting this panel's data on a voluntary basis, and comparing the response rates for this voluntary panel with those from the mandatory panel.

SURVEY DESIGN

Assumptions

The main goal of this study was to test for significant differences in response rates between mandatory and voluntary reporting, and this goal, along with other concerns, such as the effect of the study on the PACE estimates, dictated the design of the experiment. At the outset of this study, it was assumed, based upon the experience noted for other economic surveys at the Census Bureau, that the voluntary rate of response would likely be less than the mandatory rate. As a result of this basic assumption, we decided that we wanted to be able to detect response rate differences of at least 10 percent for any of the mandatory vs. voluntary comparisons.

A second assumption made in considering the design of this study was that prior conditioning would bias response rates upwards. Under this assumption, we decided it was important to not only compare overall mandatory and voluntary response rates, but also response rates for establishments with previous exposure to the PACE survey ('old' establishments) and response rates for establishments with no previous exposure to the PACE ('new' establishments).

Finally, our basic assumption presumed that a certain percentage of establishments would no longer provide data once converted to a voluntary status. For the large plants of certainty companies (certainty companies being large companies which have all of their plants included in the ASM), this loss of vital reported data could possibly affect the quality of the PACE estimates. In order to avoid this potential problem, we decided to exclude all establishments of certainty companies from any chance of selection for the voluntary segment of the study sample.

Sample Design

As mentioned above, a new sample was selected for the 1989 PACE survey. The PACE survey is an establishment-based survey and is resampled every 5 years as a subsample of the ASM. The PACE is sampled using a probability proportionate to size (PPS) design. The measure of size used to determine sampling probabilities for each establishment is weighted ASM total value of shipments (TVS).

Once the PACE survey is mailed, the following efforts are made to obtain data for delinquent plants: (1) after 60 days, reminder letter mailed, (2) after 90 days, questionnaire remailed with follow-up letter, and (3) after 120 days, another reminder letter mailed. These procedures were followed for both the mandatory and the voluntary panels.

The sampling frame for the voluntary portion of the 1989 PACE survey was created from four basic sources: the new 1989 PACE sample, the sample of 1988 ASM births to be added to the new sample, the 1988 PACE sample (selected in 1984), and the 1989 ASM company certainty cases. Once established, the frame, on a company basis, was further partitioned by 'new' and 'old' strata. 'New' refers to companies which had no previous exposure to the PACE survey, and 'old' to those which had previous exposure. The classification of 'old' multiunit establishments was not necessarily clean because multiunit companies classified as 'old' may have had establishments which were new to the PACE survey. These 'new' establishments were not included in the 'old' comparison, and since we assumed that a company's previous exposure would indirectly affect plants of the company not previously in the survey, it was also decided not to include these plants in the 'new' comparison. However, they are included in the overall response rate comparison. After some final frame adjustments due to various coverage actions, (deletes for deaths, etc.), a final sampling frame consisting of 11,204 establishments resulted.

The voluntary sample was drawn independently for each stratum ('new' and 'old') using Poisson sampling with equal probabilities. Although sampling was performed at the company level, probabilities were assigned for each stratum so that the proportionate numbers of 'new' and 'old' plants were nearly the same as the universe proportions. A final voluntary sample of 661 'new' establishments and 365 'old' establishments was drawn from the universe of 11,204 plants. All establishments not selected formed the mandatory panel.

Response Rate and Variance Estimation

All of the results from this study and the interpretations of these results are based on the various response rates and the standard errors calculated on these response rates. Since in our analysis we were interested in response rates relative to the selected PACE panel, simple unweighted rates were calculated for each possible partition i (i.e. mandatory overall, voluntary 'new', mandatory 'old', etc.) as follows:

$$R_i' = X_i / n_i$$

where, X_i = total number of sample respondents in partition i
n_i = total number of sample establishments in partition i

The use of this "unweighted" estimator for the overall mandatory and overall voluntary partitions was justified by the fact that the sampling fractions for the 'old' and 'new' panels were virtually the same. Thus, the simple estimate is equivalent to a weighted estimate.

Variances of these estimated response rates were calculated taking into account all three stages of the sample selection. These three stages were: (1) ASM sample selection from the manufacturing universe, (2) PACE sample selection from the ASM, and (3) voluntary (mandatory) sample selection from the PACE. As we preferred in our analysis not to extrapolate the response rates beyond the selected PACE panel (i.e. to effectively ignore the first and second stages), these variance estimates represent overstatements. They were calculated as follows:

$$\sigma^2_{R_i'} \approx \frac{1}{n^2} \left(\sigma^2_{X_i} + R^2 \sigma^2_{n_i} - 2R\sigma_{X_i n_i} \right)$$

$$\approx \frac{1}{n^2} \left[(1 - 2R_i') \sum_{j=1}^{n_i} X_{ij}(1 - p_{ij}) + (R_i')^2 \sum_{j=1}^{n_i} (1 - p_{ij}) \right]$$

where, n = total number of sample establishments in study panel
X_{ij} = 1 if establishment j is respondent
= 0 otherwise
p_{ij} = probability of establishment j being in partition i

and n_i and R_i' are defined as before.

Some results of this study can be seen in tabular form and graphical form. This section provides brief descriptions of each table and each graph.

Table 1 deals with final response rates for the three mandatory vs. voluntary comparisons. This table shows final mandatory and voluntary response rates for the overall, the 'new', and the 'old' comparisons. Along with the response rates themselves, the components of each response rate (i.e., the number of establishments in each stratum and the number of respondents in each stratum) also are shown. Standard errors on each response rate are provided as a means to test whether observed response rate differences are significant. One thing should be noted when looking at this table. Totals for the 'new' and 'old' comparisons do not add to the overall comparison totals. This is because 'new' plants of 'old' companies were not included in the 'old' response rate comparisons, but were included in the overall comparison. This ensured that the 'old' comparison would truly include only plants which had previous exposure to the PACE survey.

Graph 1 relates to the timeliness of response for mandatory vs. voluntary reporting. This timeline displays response rates for the overall mandatory vs. voluntary comparison from the time of the first follow-up through closeout. Using this timeline, response rates can be compared at each weekly interval (each data transmission date). The three follow-ups are denoted on the timeline so that response behavior after each follow-up can be clearly observed.

ANALYSIS AND EVALUATION

Interpretations

Results obtained from this study indicate that mandatory reporting is more effective in obtaining higher response rates than voluntary reporting. Mandatory rates are higher than voluntary rates for all three comparisons of interest.

The first response rate comparison was the overall mandatory vs. voluntary comparison. The final response rate for mandatory reporting was 82.5 percent, which was 21 percentage points higher than the corresponding final voluntary rate of 61.5 percent. Based upon the standard errors on these response rates, this difference of 21 percentage points is statistically significant. This result indicates that when a new sample is selected for an existing survey concurrently with a shift from mandatory to voluntary reporting, response deteriorates.

The second comparison was the comparison between mandatory and voluntary response rates for establishments with no previous exposure to the PACE survey (i.e., 'new' establishments). This comparison enabled us to most closely simulate a new survey. As with the overall mandatory vs. voluntary comparison, the final response rate obtained for mandatory reporting was significantly higher than the voluntary rate. The mandatory response rate of 81.2 percent was almost 25 percentage points higher than the voluntary

273

rate of 56.4 percent. This result indicates that higher response can be obtained if a new survey is mandatory rather than voluntary.

The last comparison was between mandatory and voluntary response rates among establishments with previous exposure to the PACE survey (i.e., 'old' establishments). This comparison allows us to examine the effects of prior conditioning on response. Once again, the final mandatory response rate was higher than the corresponding voluntary rate. However, the mandatory rate of 85.4 percent was only about 11 percentage points higher than the voluntary rate of 74.1 percent. Although this smaller difference is still statistically significant, this result seems to indicate that prior conditioning has some effect on response, though not enough to obtain the same response rate with voluntary reporting as with mandatory reporting. Furthermore, we can speculate, but are not able to show from this study, that over time this conditioning effect would dissipate, resulting in higher nonresponse. The results of these three comparisons can be seen in Table 1 of the appendix to this paper.

In addition to comparing final mandatory and voluntary response rates, we also compared the timeliness of response for mandatory and voluntary reporting. This was done by monitoring response at weekly time intervals. Graph 1 (also in the appendix) shows mandatory and voluntary rates for the overall comparison at each weekly data transmission from the first follow-up to closeout. This timeline shows that response under mandatory reporting was markedly higher (about 15 percentage points) even before follow-up began. After the first follow-up, mandatory response increased at a higher rate than voluntary response. After each succeeding follow-up, mandatory and voluntary response rates increased at similar rates. From these results, we conclude that mandatory reporting also improves the timeliness of response.

In order to more closely examine the relationship between establishment size and response, we partitioned the study panel into three size categories. This enabled us to see whether or not we would obtain similar mandatory vs. voluntary response rate differences among size groups as we observed for the same comparisons over the entire panel. The response rate differences were quite similar. The differences exhibited for the 'large' establishments were slightly lower, reflecting the possibility that more intensive effort, such as analyst phone calls, might have been expended to collect data from larger establishments whose data are considered vital to the estimates. The following summarizes these differences for all comparisons:

	OVERALL	'NEW'	'OLD'
ENTIRE PANEL	21.0%	24.8%	11.3%
'SMALL'	20.4%	22.9%	9.0%
'MEDIUM'	24.2%	29.0%	13.8%
'LARGE'	12.8%	17.6%	8.4%

All of these results seem to show that regardless of establishment size, mandatory reporting yields higher response rates than voluntary reporting.

This further strengthens the idea that mandatory reporting provides higher response than voluntary reporting.

The last analysis was to examine the impact of nonresponse on the PACE estimates for operating costs and expenditures. In other words, how would these estimates be affected if the PACE survey were conducted on a voluntary basis? This analysis was done on the mandatory respondents in the study for both the 'new' stratum and the 'old' stratum. Delinquent imputation which would have been used had each plant not reported was replicated. This entailed using the appropriate imputation ratios for operating costs and imputing zero for expenditures. Once this was done, a probability of response under voluntary reporting was assigned for both the 'new' and the 'old' portions (probabilities based on respective voluntary response rates obtained in this study). Expected values were then calculated for each plant and both actual and expected simple weighted estimates were computed for both the 'new' and the 'old' strata. The actual and expected weighted estimates were then compared to determine the impact on the estimates under the assumption that the PACE survey was conducted on a voluntary basis.

Results from this analysis are quite interesting. For operating costs, the expected final weighted estimate for the 'new' comparison was 9.5 percentage points lower assuming that the PACE was voluntary. The expected final weighted estimate for the 'old' comparison was 3.6 percentage points lower than the actual weighted estimate for the stratum. Both of these differences were considered trivial when compared with the actual weighted estimate for operating costs for the entire mandatory portion of the study panel.

Results obtained from the analysis on expenditures estimates showed that nonresponse has a larger effect on these estimates. As mentioned earlier, zero is imputed for nonresponse, which as our results show, may not be a good assumption. The 'new' comparison showed that the expected weighted estimate was 43.6 percentage points lower than the actual estimate for that portion of the panel, while the 'old' comparison showed the expected estimate to be 25.9 percentage points lower than the actual weighted estimate. These results seem to show that imputing zero for nonresponse for expenditures may not be a good idea, and that further research may be needed to develop an imputation algorithm for expenditures.

Limitations of the Results

There were several factors involved in this study which may have affected the results and/or the interpretations of these results. The first contributing factor was the limited scope of the study. The PACE survey excludes all establishments with total employment (TE) less than 20, therefore, the study did not include any of these establishments. Also, plants of ASM certainty companies were excluded from any chance of selection for the voluntary portion of the study panel, so they were not included in the study. These plants were excluded to prevent the potential loss of vital data.

Since the PACE survey was used as the test vehicle for this study of mandatory vs. voluntary reporting, all of the operating procedures normally followed for the PACE survey were executed as usual. For instance, regular PACE follow-up procedures were followed for nonresponse. Additional follow-up methods could have been tested as a means for trying to obtain voluntary response rates comparable to the higher mandatory rates. This idea is precisely the basis for the follow-up study being done using the 1990 PACE survey. This follow-up study is described briefly later in this paper. Therefore, it should be realized that the results obtained from this test of mandatory vs. voluntary reporting apply to the PACE survey itself, and care should be exercised in extrapolating these results to other surveys.

One other factor which affected the results and/or the interpretations of these results was sampling error. All of the study sample estimates are subject to sampling error. All significance testing for this experiment was done at the 95 percent confidence level.

RECOMMENDATIONS

Current Research

Based upon the results obtained from this study, mandatory reporting seems to yield higher response than voluntary reporting. Mandatory response rates for **every** response rate comparison are higher than the corresponding voluntary rates. These results, and any conclusions based upon these results, may only be applicable to the PACE survey. Care should be used in extrapolating results to other surveys.

Despite the suggestion that results only be extrapolated to the PACE survey itself, this is a well-controlled study which provides some concrete evidence that response under mandatory reporting is higher than response under voluntary reporting. This is only the beginning for research in the area of nonresponse under mandatory vs. voluntary reporting. Further research needs to be done to follow-up on the results obtained in this study. The 1990 PACE follow-up study, which is currently in progress, is designed to provide further information on the issue of nonresponse under mandatory vs. voluntary reporting.

Future Research

As a result of the 1989 PACE split panel test, other research questions arose. What effects would telephone prompting, telephone follow-up, or certified letter follow-up have on the voluntary response rates? Would such data collection techniques narrow the difference between response rates for mandatory vs. voluntary reporting? Therefore, another test was developed. In this test, the voluntary portion of the 1990 PACE survey (both 1989 respondents and nonrespondents) was randomly split into five panels. For each panel, a different data collection technique is tested to improve response.

Panel I - For this panel, a premail telephone call was made to each sample unit 30 days **prior** to

mailout to inform the respondents to expect a PACE questionnaire in the mail soon. This panel tests the effects of a prenotification telephone call on voluntary response rates.

Panel II - For this panel, a post-mail telephone call was made to each sample unit 30 days **after** mailout to confirm receipt of the PACE questionnaire and to remind the respondent of the questionnaire's due date. This panel tests the effects of a reminder telephone call on voluntary response rates.

Panel III - For this panel, residual nonrespondents received follow-up telephone calls to obtain the requested data at the 90-day due date and the 120-day reminder date. This panel tests the effects of follow-up telephone calls on voluntary response rates.

Panel IV (control group) - For this panel, the normal PACE follow-up procedures referred to earlier were followed.

Panel V - For this panel, residual nonrespondents received certified mail follow-up letters to obtain the requested data at the 90-day due date and the 120-day reminder date. This panel tests the effects of certified mail follow-up letters on voluntary response rates.

The original allocation of the sample was altered as a result of deletions and shifting between panels. Ghosted establishments were allowed to shift panels to ensure that all establishments of a particular company received the same treatment. The following summarizes the final establishment counts and response rates for each panel as of the end of September:

	# estab	# resp	resp. rate
mandatory	9965	8562	85.9%
Panel I	104	54	51.9%
Panel II	102	71	69.6%
Panel III	105	66	62.9%
Panel IV	538	313	58.2%
Panel V	147	100	68.0%

Comparing these response rates shows that the mandatory response rate is still the highest. The lowest response rate is for the voluntary panel in which phone calls were made 30 days prior to mailout (Panel I). Perhaps to be more effective, these premail phone calls should have been made closer to mailout.

While the premail phone calls do not seem to have improved response, the **post**-mail phone calls seemed to have a positive effect on improving response rates. The response rate for the voluntary panel employing post-mail phone calls (Panel II) is the highest rate among all voluntary panels. With respect to follow-up, certified mail (Panel V) seems to have been the most effective follow-up method for improving response (increase of about 29 percent since follow-up began).

Finally, the Industry Division of the Bureau of the Census is planning to conduct additional data collection research in other surveys in the near future to identify more data collection

improvements. With these improvements, we hope to improve our survey response rates and, hence, the quality of data for our users.
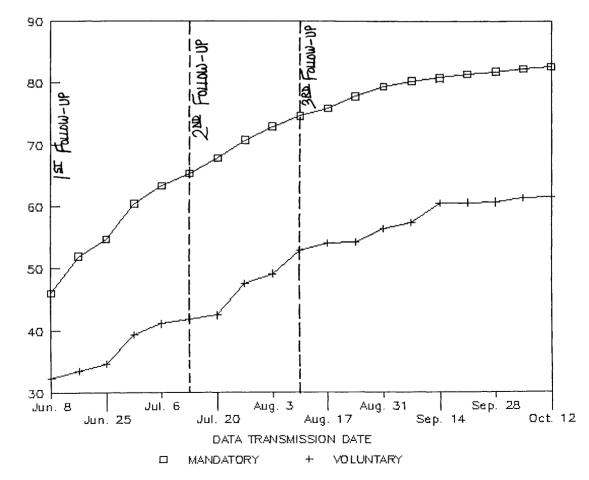
REFERENCES

1989 Publication for the Survey of Pollution Abatement Costs and Expenditures (PACE), Industry Division, Bureau of the Census.

Garrett, Joseph K., Konschnik, Carl A., and Raglin, David A. (1988), "Some Indications of the Effects of Voluntary Versus Mandatory Reporting in the Retail Trade Surveys," Business Division, Bureau of the Census.

Ogus, Jack L. and Clark, Donald F. (1971), Technical Paper No. 24, The Annual Survey of Manufactures: A Report on Methodology, Industry Division, Bureau of the Census.

Tulp Jr., Daniel R. (1990), Internal Documentation, "Summary of the Sample Selection Operations for the Voluntary Portion of the 1989 Survey of Pollution Abatement Costs and Expenditures (PACE) - MA200(V)," Industry Division, Bureau of the Census.

Table 1. FINAL MANDATORY VS. VOLUNTARY RESPONSE RATES

	Overall				'New'				'Old'			
	# Cases in Stratum	# Resp. in Stratum	Resp. Rate (%)	S.E. (%)	# Cases in Stratum	# Resp. in Stratum	Resp. Rate (%)	S.E. (%)	# Cases in Stratum	# Resp. in Stratum	Resp. Rate (%)	S.E. (%)
MANDATORY	10178	8401	82.5	0.3	6753	5483	81.2	0.4	1909	1631	85.4	0.5
VOLUNTARY	1026	631	61.5	1.8	661	373	56.4	2.1	224	166	74.1	3.3
DIFFERENCE (%) =	21.0				24.8				11.3			

Graph 1. TIMELINESS OF RESPONSE