
When They Don't Remember — A Multi-Method Approach to Analyzing Taxpayer Usage of Proposed Mailing Procedures

*Karen O'Connor and Fonda Sackett, Internal Revenue Service;
Deborah Stone, Bureau of Labor Statistics; and
David Cantor, Kerry Levin, and Dwayne Norris, Westat, Inc.*

The purpose of this paper is to discuss the advantages of mixing quantitative and qualitative methods in a multi-year iterative study of proposed mailing procedures. The Internal Revenue Service is exploring different methods for processing tax payments. In order to process payments in different ways, the mailing procedures must be changed so that payments can be handled separately from returns. The business vision of the IRS is to move 90 percent of payment processing to the financial service community through electronically transferred funds and/or direct mailing to a "lockbox" -- a financial institution which processes payments. Using an operational test, the Service began testing direct mailing methods with the 1992 tax packages. This paper describes that effort.

The project is presented in chronological order by tax year: tax year 1992 (TY92), TY93, and TY94. For reference, TY92 tax packages are mailed during calendar (or Processing Year 1993, PY93), which runs from January 1, 1993 through December 31, 1993. Whenever the type of year is not specified, such as: 1992 1040 Voucher Study, the year referred to is the tax year.

Within each year, the IRS operational test is presented and then the evaluation methodology is described and the results are discussed. The move to a different method of processing payments involves a tremendous change in the processing of tax returns. IRS is preparing for and testing the change through a series of operational tests, much like pretests. As the Service gains experience, the number of taxpayers involved in the tests increases. In the first year, only 115,532 taxpayers were part of the test. For tax year 1995, all taxpayers will be asked to use a voucher if they are mailing in a payment with their return.

This project has used many different methods for collecting information and each has given us useful information. In the first year, 10-percent of the taxpayers followed the new mailing procedure; in the second year, 34 percent. For the third year, a preliminary estimate gives 74 percent of the taxpayers following the new mailing procedures. Every year the process has been improved and more taxpayers are using the new mailing procedures.

■ TY92 — Operational Test

Test

For Tax Year 1992 the Service sent a test tax package with two envelopes to about 115,532 taxpayers in the Southeast Region. The package instructed the taxpayers to put their return in a large envelope. If they were sending in a payment, then they were to send their check and a payment voucher in a small envelope. The voucher is similar to the payment coupon that is used with credit card payments. The TY92 evaluation consisted of: a customer satisfaction survey, two focus groups with tax practitioners, and laboratory interviews with taxpayers.

Customer Satisfaction Survey

Methodology -- We began with a Customer Satisfaction Survey. The methodology of the Customer Satisfaction Survey is described in more detail by Norris et al. (1995). Some of us naively thought that the survey would answer all of our questions. We selected 2,801 taxpayers at random from the following strata:

- taxpayers who sent in their payments separately from their returns;

- taxpayers who sent in their payments with their returns;
- taxpayers who requested an extension;
- taxpayers who received a refund; and
- taxpayers who sent in a return, but either no payment was made or the payment was not recognized as part of the test.

The population consisted of taxpayers in the Southeast Region who received a test tax package. IRS administrative information, which included the name, address, accounting codes indicating payments, and document locator numbers, was used to select and administer the survey. (There were a large number of taxpayers requesting an extension for TY92, because taxpayers were allowed to delay payment without a penalty when they requested an extension. This special extension was only for processing year 1993.)

Results -- The customer satisfaction survey was intended to identify taxpayers' concerns with and use of the new mailing procedures. The response rate for the overall survey was 70.0 percent. The survey identified a major problem in collecting this information: many taxpayers did not remember the test tax package! **All of the taxpayers received the test tax package**, but only 57 percent said that they had some knowledge of it. Of those who followed the new procedure correctly, only 78.2 percent indicated that they had some knowledge of the test tax package.

Part of the problem may have been due to the fact that the Customer Satisfaction Survey was conducted several months after most returns were filed. An obvious solution is to conduct the interview a week after the taxpayer mails his/her payment. However, this is not feasible, because it takes several months to obtain administrative records with filing information on taxpayers. Administrative records identify taxpayers with payments. Since only about 20 percent of all taxpayers send in payments, it is necessary to obtain administrative records in order to construct an efficient sample design.

With about half of the respondents not remembering the test tax package, we decided to analyze the opinions of taxpayers on the new mailing procedures. Table 1 summarizes our findings from the survey.

Table 1.-- Opinions on the Tax Form, TY 1992 (in percent)

Opinion	Strongly Agree	Strongly Disagree	Don't Know
New procedure would not be confusing ¹	62.8	13.9	23.3
Would not like using additional postage ²	58.2	31.3	10.5
Would feel uneasy about mailing check and return separately ³	57.3	35.5	7.2

¹ Answered by all mail and telephone respondents who did not recall mailing procedure.

² Answered by all mail and telephone survey respondents.

³ Answered by all telephone and mail survey respondents who recalled mailing procedure.

Conclusions -- Unfortunately, opinions don't necessarily translate into actions. The new procedures may not have seemed confusing over the phone in the abstract, but according to IRS accounting records, only about 10 percent of the taxpayers making payments followed the new instructions. Since the survey data were not consistent with the data from accounting records, we decided to use qualitative methods to explore the inconsistencies. The survey data were used to look at variables which correlated with incorrect and correct use of the mailing procedure.

Use of a tax preparer correlated with incorrect use of the mailing procedures. We thought that the

way that tax preparers handled taxpayers' forms might be the reason for the correlation, so we decided to do two focus groups with tax preparers.

Whether the tax form was kept and used when taxes were filed was related to correct use of the mailing procedures. We decided to do some laboratory interviews with taxpayers, so we could see what taxpayers do with their tax packages. We also added this topic to the focus groups.

Because age and income levels were also related to usage, it was evident that we needed to talk with taxpayers of different ages and income levels in the focus groups and laboratories.

TY92 Focus Groups for Tax Preparers

When we conducted focus groups for tax preparers, we found that tax preparers do not use tax packages. They generate their own forms, using various software, or keep their own supply of forms. Even if clients bring in their own forms, the tax preparer will not use them or will use only the labels. Since the voucher can not be printed by the preparers, following the new procedure would require tax preparers to file and save the labels, envelopes, and vouchers for each client. From the feedback we received, we expect that the tax preparers will probably resist using the vouchers until the vouchers can be generated through their software.

TY92 Laboratory Interviews for Taxpayers

The purpose of the laboratory interviews of taxpayers was to better understand the responses to the survey questions. We wanted to know if the respondents would send their checks in separately from their return, even if they felt uneasy; if they would pay the additional postage; and why they weren't using the new procedures.

Seven interviews were conducted using the TY92 tax package to address these questions. Seven in-

terviews were also conducted on the TY93 preliminary tax packages, which had incorporated some changes into the voucher procedure -- information about the new procedure was placed on the front cover of the tax instructions and two mailing labels were provided, instead of two envelopes. Because of the long lead time required to print and distribute millions of tax packages, we had to explore the problems of the TY92 package while preparing the TY93 tax packages.

The purpose of the interviews on the TY93 packages was to improve the TY93 tax packages before they were mailed. For all of the laboratory interviews, the following procedure was followed.

The interviewer explained to the respondent that he/she would be asked to complete and prepare for mailing a mock tax return. It was further explained that for the purposes of the study, the respondent should "think aloud" as much as possible during the task, as this would help the interviewer observe reactions as directly as possible. The respondent then completed and prepared the mock tax return for mailing.

Much of the information in the Form 1040 had been inserted in advance by the researchers, so that it would be clear to respondents that they "owed taxes." Thus, a mock check (representing a tax payment to be "mailed" to the IRS) was filled out by the respondent, so that reactions to the voucher and separate envelope procedures could be observed. Upon completion of the task, the interviewer asked the respondent a series of prepared probes regarding his/her understanding of and reactions to the TY92 tax package.

Results of the laboratory interviews on the TY92 tax packages were:

- None of the subjects read the "What's New" section of the tax instructions, either prior to or while completing the Form 1040. The What's New section provided taxpayers with information about the change in procedure.

- Only 2 out of 7 respondents located the voucher. They both used the voucher. Only 1 of them used the two envelopes correctly.
- The voucher and envelopes were difficult to find.
- The directions were confusing.

Results of the laboratory interviews on the preliminary TY93 tax packages were:

- The directions were confusing and contained contradictions.
- None of the participants read the front cover of the tax instructions either prior to or while completing the Form 1040.
- It seems to be easier to use the one envelope and two labels mailing procedure than the two envelope mailing procedure. Only 1 out of the four participants using the two envelope mailing procedure used the materials correctly. Both of the participants using the one envelope/two label procedure used the materials correctly.

To recap, in the first year, a customer satisfaction survey, laboratory interviews, and focus groups of tax preparers were conducted. This information

was used to modify the forms and mailing procedures for the second year. The qualitative and quantitative methods were combined to provide a justification for changing the forms and instructions, and the information on how to change the forms and instructions.

■ TY93 – Operational Test

Test Package Alternatives

For Tax Year 1993 (the second year) the Service tested three test tax packages using different mailing procedures. Each test package was tested at one test site. Logistic constraints prevented us from testing each procedure at all three sites. Table 2, below, shows site, method, and volume:

The **Ogden Test Package** included one large envelope with two color-coded preaddressed labels (**two-label procedure**). If taxpayers had a balance due, they were instructed to mail the voucher, return, and check in the envelope using the appropriate label. If the taxpayer had no balance due, they were instructed to mail their return in the envelope using the other label. Both labels were addressed to different IRS post office boxes.

Table 2.--Operational Test: Site, Method, and Volume

Summary of Laboratory Performance					
Summary of Laboratory Performance	One Envelope	Two Envelopes		Total	
	Ogden (n=20)	Brookhaven (n=13)	Cincinnati (n=13)		
Read the Front Cover	1	0	2	3/46	(6.5)*
Used the Voucher	8	3	5	16/46	(34.7)
Used Two Envelopes	N/A	3	4	7/26	(26.9)
Used Correct Label	11	N/A	N/A	11/20	(55.0)
Followed Procedures Correctly	6	3	4	13/46	(28.3)

*Numbers in parentheses represent the percentage of sample in specified cell.

The **Brookhaven Test Package** included two envelopes -- one large and one small (**two envelope procedure**). If taxpayers had a balance due, they were instructed to mail their return in the large envelope and their payment voucher and check in the small envelope. If taxpayers had no balance due, they were instructed to mail their return in the large envelope. Both envelopes were addressed to different IRS post office boxes.

The **Cincinnati Test Package** was the same as the Brookhaven test, except that the small envelope was postage paid (business reply mail) (**two envelope -- one postage paid procedure**).

We planned to do a customer satisfaction survey for the 1993 Tax Year, but we could not get a file of the taxpayers who received the test tax packages until the end of October. Based on our experience with the TY92 Customer Satisfaction Survey, we figured that, by October, the taxpayers would have forgotten how they had mailed their payments. Therefore, there was no value to be gained in conducting a customer satisfaction survey so long after most of the taxpayers had mailed in their returns.

Instead of a customer satisfaction survey and as a result of the successful use of qualitative methods in the TY92 evaluation, a series of laboratory interviews and focus groups, telephone interviews with state personnel, and expert analysis were conducted. Administrative records were also obtained for the taxpayers who received the test tax packages, to measure the percentage of taxpayers who followed the new procedures. They showed that the percentage of taxpayers who followed the mailing instructions went from 10 percent in the first year to 34 percent across all the packages in the second year.

We used quantitative and qualitative methods to look at the TY93 test tax packages. The quantitative methods -- laboratory interviews and administrative data -- will be discussed first. They are both used to estimate the percentage of the popu-

lation which followed the instructions in the test tax packages. Then, the qualitative methods are discussed.

Evaluation Methodology and Results

Laboratory Interviews -- Forty six one-on-one "think aloud" laboratory interviews were conducted with taxpayers in five geographic locations (Rockville, MD; Baltimore, MD; Phoenix, AZ; Los Angeles, CA; and St. Louis, MO), to examine the usability of the test tax packages. The taxpayers were presented with one of the test tax packages and asked to complete the package. All of the information needed to complete the return was provided. It was the same procedure as in TY92, except that there were three tax packages being tested. In addition, in TY93, the taxpayers were asked to complete a small questionnaire specific to the test tax package which the respondent filled out. Once the questionnaire was completed, the facilitator introduced the remaining tax packages to the respondent. Respondents were then asked to discuss their thoughts about the revised procedures, indicate whether they had a preference among the three procedures, and provide any recommendations for improvements in the clarity of the instructions and/or procedures.

Administrative Data -- Based on administrative records, such as the Individual Master File and summary record counts from the lockboxes and the mail out of the test tax packages, Table 3 presents the percent payments processed by lockbox banks for the test bed service centers.

Table 3.--Percent Test Tax Package Taxpayers Sending Payments to Lockbox Banks for TY93

Lockbox	Timely Filers	Extentions and Late Filers	All Filers
Ogden	30%	NA	46%
Brookhaven	31%	10%	23%
Cincinnati	32%	20%	27%

The IRS identified the taxpayers to receive the test tax packages based on the time that they filed their return in the previous year. The first group of taxpayers -- timely filers -- were primarily taxpayers who filed by April 15, 1993. Because of the large volume of returns and payments in April, it took until July to process all of the information.

The second group -- extensions -- were primarily taxpayers who filed an extension and filed their return in August; their returns were processed by the end of October 1993. This group was unusually large in processing year 1993, because that year taxpayers were allowed to delay their payment until August without a penalty when they filed an extension, as mentioned with regard to the TY92 Customer Satisfaction Survey.

The third group -- late filers -- was taxpayers who filed by the first week in December 1993. These taxpayers filed their returns very late in the filing year [1].

The higher percentage in Ogden is probably the result of Ogden's publicity of the lockbox program. Cincinnati and Brookhaven did not publicize the lockbox program. Because Ogden expanded the number of taxpayers who knew about the lockbox program beyond those that received the test tax package, it is difficult to estimate the percentage of taxpayers that used the lockbox based on the tax package instructions. If all of the missing taxpayers in the third group for Ogden sent their payments to the lockbox, the total percent processed by the lockbox would have been 41 percent. If none of them had sent their payments to the lockbox, the percent would have been 29 percent.

It is interesting to see that the laboratory interviews gave similar results to the administrative data. And, it seems quite clear that the TY93 test tax packages were easier to use than the 1992 test tax packages.

TY93 Qualitative Data Methodology -- The following five methods of collecting qualitative data were used to obtain recommendations for re-

visions to the test tax packages: tax preparer focus groups, taxpayer focus groups, taxpayer laboratory interviews (after completing the test tax packages), interviews with state revenue offices, and expert analysis. The methodology of the focus groups is essentially the same as in TY92. The format of the laboratory interviews was described above.

Personnel from the Department of Revenue in Pennsylvania, South Carolina, and Virginia were interviewed. They were asked to describe their experiences using mailing procedures, similar to those being tested, involving one mailing envelope and multiple address labels.

The primary purpose of the expert analysis was to analyze the payment voucher and the instructions in the individual income tax package and *Your Federal Income Tax*, a guide book for filing published by the IRS, in order to identify problems that might produce taxpayer confusion.

Results -- The recommendations of these groups were compared, touching on all aspects of the process including the front cover, voucher instructions, voucher reference and location, envelopes, lockbox, and other comments. In all, there were 52 recommendations. Some were very specific, such as use a bright color (not blue) as the headline designation for the new procedure. Many of the groups gave the same suggestions. Presser and Blair (1994) reached similar conclusions in their comparison of the comments from a conventional pretest, behavior coding, cognitive interviews, and expert panels.

The taxpayers in the cognitive lab had some unique comments, such as use thicker paper for the front cover and change the wording to reflect that the new procedures are not optional. These taxpayers were the only participants of the groups to actually complete the new mailing procedures. Once again, there seems to be a difference between thinking about the new mailing procedure and doing the procedure. This feedback proved very valu-

able in making changes to the procedure for the TY94 operational test.

■ TY94 Operational Test

For TY94, the third year, IRS sent out about three million test tax packages nationwide. These were focused on self-employed taxpayers, fishermen and farmers who were expected to have a balance due when they filed their return [2].

IRS tested the two-label procedure, pretested in Ogden, and the two-envelope procedure, pretested in Brookhaven. The method with postage paid envelopes was dropped for TY94. The differences between the TY93 and TY94 packages are in the appearance of the forms and the instructions, which were improved between TY93 and TY94, based on the qualitative recommendations received.

Because it was decided late in the planning to test the two-envelope method, the only option for including a two-envelope test tax package in the TY94 was to switch one of the packages from the two-label method to the two-envelope method. So, it was decided that the two-envelope package would be sent to the less complex self-employed sample of taxpayers.

■ TY94 Evaluation Methodology and Results

For the third year, a sample of envelopes was drawn and examined to determine the way that taxpayers were following the mailing procedures. Later this year, the sample data will be merged with administrative data to estimate the characteristics of filing. At this point, based on crude administrative data, it appears that 74 percent of the taxpayers sent their payments to the lockbox -- a vast improvement over the earlier tests.

■ Conclusion

The advantage of mixing quantitative methods with qualitative methods is that the two comple-

ment each other. Quantitative methods identify problems and measure the size of the problem. Qualitative methods explain why the problem occurred and what can be done to correct the problem. They convince people to take action.

The survey data let us know that we had a serious problem. It took the laboratory interviews and the focus groups to know what was causing the problem and some of the solutions. The results from survey data may not always be apparent. On the other hand, there is tremendous impact from watching as taxpayers overlook the front cover and incorrectly complete a test tax package or watching taxpayers touch, read, and ignore the voucher. It was very difficult for the IRS program personnel to believe that people did not read the material that was sent to them. They felt that a statement on the cover and in the Commissioner's message alerting taxpayers to the new mailing procedure was all that was needed. The laboratory interviews and focus groups brought IRS personnel into close contact with taxpayers, allowing them to see where the difficulties lay in the use of vouchers.

As a result of all of the evidence, the IRS moved to include vouchers in all tax packages by TY96. The number of taxpayers receiving vouchers was also increased in TY94 and TY95. In addition, various changes were made in the Form 1040 instructions, mentioning the voucher, starting in TY94.

■ Footnotes

- [1] Approximately, 28 percent of the mailing records for this study were accidentally lost when some of the tapes containing the administrative information were scratched. We have administrative records for 43 percent of the Brookhaven taxpayers, 55 percent of the Cincinnati taxpayers, and 86 percent of the Ogden taxpayers. The preceding table is based on a combination of the administrative records and processing totals.

Synthetic estimation was used to fill the cells where administrative data were missing. We assumed that the percentage of taxpayers sending in payments is 47 percent in all of the empty cells. Available data estimated the percentage of the test taxpayers who made payments to be between 47 percent and 52 percent. Conservatively, 47 percent was used as the estimate. If 52 percent were used the percentages in the ALL FILERS column would have been Brookhaven, 20 percent; Cincinnati, 24 percent; and Ogden, 42 percent.

- [2] Every year the IRS sends out different packages of tax forms to individual from 1040 taxpayers, based on the tax forms they filed the prior year. There are altogether ten packages with different sets of tax forms. Packages 1040-5, 1040-6, and 1040-10 have higher percentages of taxpayers with a balance due when they file their return. Package 1040-5 goes to the self-employed; 1040-6 is sent to fishermen and farmers. Package 1040-10 also goes to self-employed, but Package 10 has fewer forms than Package 5. All of the test tax pack-

ages in the TY92 and TY93 studies were 1040-5 packages. Test tax packages with vouchers were produced for packages 1040-5, 1040-6 and 1040-10. The test tax packages are abbreviated as 1040-5V, 1040-6V, and 1040-10V. The test tax packages went to taxpayers who would have received tax packages 1040-5, 1040-6, or 1040-10 and had a balance due the prior year.

■ References

Norris, Dwayne; Levin, Kerry; Cantor, David; Stone, Deborah; and O'Connor, Karen (1995). A Mixed-method Approach to Evaluating New Tax Payment Procedures, Presentation at the 1995 Annual Meeting of the American Association of Public Opinion Research.

Presser, Stanley and Blair, Johnny (1994). Survey Pretesting: Do Different Methods Produce Different Results? *Sociological Methodology*, P. V. Marsden (ed.) Oxford: Blackwell Publishers. ■