

Mode Effects in a Survey of Consumer Expenditures October 2008

Adam Safir and Karen L. Goldenberg

U.S. Bureau of Labor Statistics, 2 Massachusetts Avenue, Washington, DC 20212

Abstract

In an effort to increase response rates and control survey costs, survey designers have come to increasingly consider mixed mode survey designs. The literature suggests that mode equivalence can be maximized through attention to instrument design and careful implementation of survey protocol (e.g., Martin, et al, 2007). However, the impact of differential administration by mode is less clear for ongoing data collection efforts that are primarily designed for a single mode of administration, where alternative modes have emerged over time upon respondent request (e.g., by telephone instead of personal visit). This is the case for the Consumer Expenditures (CE) Quarterly Interview Survey, a personal visit household survey with a non-ignorable percentage of cases interviewed by telephone. This paper draws on four years of data to assess variation in mode of administration and evaluate the size of mode effects in key survey estimates. The results of the study suggest that mode itself matters less than does the respondent behavior typically associated with mode. While telephone interviewing may impact the quality of CE data collected relative to that of personal visit interviewing, we find clear steps that can be undertaken to maximize mode equivalence and minimize mode effects.

Key Words: Mode equivalence; Mixed-mode designs; Survey costs; Data collection procedures

1. Introduction

1.1 Background

Mixed-mode survey designs represent an appealing option for survey designers interested in increasing response rates and controlling survey costs. A well-planned mixed-mode design incorporating personal visit and telephone interviewing, for example, has the potential to benefit from the superior data quality associated with personal visit interviewing for non-sensitive topics, the lower costs and greater operational efficiency of telephone interviewing, and the higher overall response rates generally found when allowing for more than one mode of response (Rogers, 1976). However, achieving this optimal outcome can be quite challenging. When evaluating a possible transition to a mixed-mode design, researchers from the National Election Study (NES) concluded that key estimates from personal visit and telephone NES interviews were not equivalent, even after controlling for demographic differences. The NES researchers expressed significant concerns about adopting a mixed-mode design for their survey (Wessel, et al., 2000). Similarly, staff from the Survey of Income and Program Participation expended years of research before successfully moving to a mixed-mode design (Allen, 1993; Huggins, 1994; King, 1995; Nelson, et al., 2001).

Ultimately, the optimization of response rates and cost efficiencies in a mixed-mode design must be balanced against the overall impact on the quality of key survey estimates. Evidence in the literature suggests that mode effects, or the systematic response pattern differences attributable to respondents being interviewed in different modes, can be minimized through careful attention to instrument design and strict adherence to survey protocol (e.g., see Martin, et al., 2007; Wessel, et al., 2000). Nonetheless, the lower response rates, higher coverage error, faster-paced interviews, and larger measurement error¹ of telephone interviewing relative to personal visit interviewing can be a cause for concern (Groves & Kahn, 1979; Groves, 1989; Hochstim, 1967).

¹ A notable exception is for questions about sensitive or socially undesirable topics (Aquilino, 1994), in which increased respondent anonymity in telephone interviewing has been associated with higher reporting. However, findings from meta-analyses are mixed, with other studies suggesting that telephone interviews are less effective at eliciting sensitive information (e.g., see de Leeuw & van der Zouwen, 1988).

1.2 Emergent Mixed Modes

The impact of mixed modes is even less clear for ongoing data collection efforts that have been primarily designed for a single interviewing mode, but in which alternatives have emerged over time upon respondent request (e.g., to be interviewed by telephone instead of personal visit). This is the case for the Consumer Expenditures (CE) Interview Survey, conducted by the Census Bureau on behalf of the Bureau of Labor Statistics (BLS). The CE Interview Survey is a personal visit household survey, with a non-ignorable percentage of cases interviewed by telephone.

Contrary to mixed-mode surveys with an established protocol for interviewing in more than one mode, telephone interviewing in the CE Interview Survey does not occur by design. Rather, interviewers are provided with a basic level of guidance for conducting a telephone interview in lieu of a personal visit interview². Over the years, the percentage of cases completed by telephone has fluctuated, most recently stabilizing at approximately 35 percent – a level still quite high for a survey designed to be administered in a personal visit format. In response, survey management has sought to better understand the characteristics and consequences of the survey's emergent, and unplanned, mixed-mode administration, as well as the costs and benefits of implementing a more detailed and rigidly-followed protocol for telephone interviewing.

Of particular concern for CE Interview Survey management is that several critical components of proper CE survey administration are largely dependent on the physical presence of an interviewer. Such aspects of survey administration include: (a) the survey's preference for respondents to consult financial records and purchase receipts in their reporting of expenditures, (b) the interviewer role in aiding respondents in retrieving correct information from their records, and (c) the effective use of an Information Booklet³ provided during the interview. Although the manner and degree to which interviewers successfully accomplish these behaviors during personal visit interviews may vary, by definition, the physical absence of an interviewer during telephone interviews limits the ability to adhere to protocol, and is therefore thought to increase the potential for measurement error.

2. Consumer Expenditure Surveys Program

2.1 Survey Overview

One of the principal objectives of the CE Surveys Program is to obtain information on the buying habits of America's consumers. The Census Bureau collects data for the survey under a Memorandum of Understanding with BLS. Twelve Census Bureau regional offices (ROs) oversee field operations. Differences in geographic coverage across ROs are substantial. For example, while the New York RO covers just 19 *counties*, the Denver RO includes 10 *states*.

The CE Surveys Program consists of two surveys. The CE Diary Survey is designed to collect small, detailed expenditures that would be difficult for respondents to recall over an extended period of time. The CE Interview Survey is designed to collect less frequent, more expensive, memorable purchases, such as expenditure data that respondents can reasonably recall for a period of three months. The analysis described in this paper is based on data from the CE Interview Survey only.

2.2 Survey Methods

The CE Interview Survey sample is selected on a rotating panel basis, with approximately 7,000 consumer units⁴ (CUs) interviewed per calendar quarter. Sampled CUs are interviewed once per quarter, for five consecutive quarters. Individual quarterly interviews – also called “waves” for methodological purposes – for a given CU are treated as independent observations. The data collected in Interview Wave 1 are used primarily to create an inventory of major household items, and for bounding purposes; that is, to prevent the reporting of expenditures from an indefinite past period. The resulting expenditure data are not used in published estimates. The second through fifth interview waves

² For more information about telephone interviewing protocol, see the 2007 Field Representative's (FR) Manual, Part A, Chapter 1, Topic 5 (Page 1-12)

³ The Information Booklet is a visual aid intended for use by respondents as they complete the interview. The booklet consists of a series of show cards presenting lists of response options and examples of the types of expenditures that fall under each option.

⁴ Consumer units, the unit of analysis in CE, consist of people living in a household related by blood or marriage, or unrelated people who share household expenditures. Each household consists of one or more consumer units. For most housing units, the household and the consumer unit are the same.

collect expenditure information on a fixed set of topics for the previous three months. Questions about income, employment, assets, and liabilities are asked in the second and fifth interview waves only⁵.

The CE Interview Survey is administered by computer-assisted personal interviewing⁶ (CAPI), either by personal visit or telephone. While the survey is based on a personal visit design, including a preference for the use of physical recall aids such as financial records, purchase receipts, and the Information Booklet, telephone interviewing is officially only permitted as a “last resort.” Telephone interviewing is decentralized, usually conducted from the interviewer’s home, RO, or some other location, using the same CAPI instrument as used for the personal visit interview. However, as described in greater detail below, the available data suggest that telephone interviewing is being conducted more often than what might be reasonably expected from “last resort” situations.

3. Analysis

3.1 Objectives

The analysis for this paper extends earlier work by McGrath (2005) and Verlander (2006) to address three basic areas of inquiry: (a) the degree to which the characteristics of telephone interview CUs differ from those of personal visit interview CUs, (b) the effects of telephone interviewing on data quality and outcome measures, and (c) whether current survey protocol or materials should be changed to accommodate the relatively high level of telephone interviewing. Data for this research were collected during the period of April, 2003, through March, 2007. The overall CE Interview Survey response rate was 78.6 percent in 2003, 76.0 percent in 2004, 74.5 percent in 2005⁷, and 76.6 percent in 2006⁸ (Krieger, 2007). It should be noted that the findings described below are selected from a broader series of analyses investigating mode effects in the CE Interview Survey. Although the fully detailed series of analyses could not be included for space reasons, the summary information presented here is representative of the overall conclusions reached in the broader research.

3.2 Data & Methods

The analysis file was restricted to Interview Waves 2 through 5 due to the limited availability of processed data from Interview Wave 1⁹. The analysis file was also limited to complete interviews with an assigned interviewing mode (interviewers assign mode by answering a post-survey assessment question). The sample size for the final analysis file was 118,698 interviews, including 73,723 personal visit interviews and 44,975 telephone interviews.

We used contingency table analysis and regression modeling to evaluate telephone administration rates and the relationship between mode and sample characteristics, data quality indicators, and outcome measures. Interviewing mode refers to the mode in which all, or most, of the interview was completed, as reported by the interviewer. Detailed findings by analysis domain are presented below.

4. Findings & Discussion

4.1 Telephone Administration Rates

In reviewing telephone administration rates, we found that the level of telephone administration for Interview Waves 2 through 5 decreased from 39.2 percent in 2006 Quarter 2 to 36.3 percent in 2007 Quarter 1 (see Figure 1). Additionally, overall telephone administration rates appear to vary substantially across ROs – possibly due to differing training, operational, or intrinsic socio-geographic factors associated with primary sampling units (PSUs), although we don’t have evidence to point in any one direction (see Figure 2).

⁵ For more detailed information about the survey, see the 2005 Consumer Expenditures Interview Survey Microdata Documentation.

⁶ Data for the CE Interview Survey were collected using a paper-and-pencil interviewing form through March, 2003. Data reported here include interviewers’ first experience with CAPI, as well as its use as interviewers became more familiar with the electronic instrument.

⁷ Note that there was a sample design change in 2005. The Interview Survey response rate tabulated for 2005 is for “2000-based” PSUs, and there were no cases from Interview Wave 2 through 5 in January 2005 in those PSUs.

⁸ There was a revision to the CE sample design in 2006 which involved the elimination of 11 PSUs. The sample reduction was effective in January for the Diary Survey and April for the Interview Survey.

⁹ CE does not perform routine processing on Wave 1 variables, because the data are not used for estimation.

4.2 Sample Characteristics

We examined 12 socio-demographic characteristics and found five that showed meaningful differences¹⁰ of greater than 3 percentage points and 5 percent (see Table 1). Compared to personal visit CUs, telephone interview CUs were more likely to be homeowners, with a reference person¹¹ who is female and has a college education. We also found that telephone interview CUs were less likely to have a reference person of Hispanic origin or to be below the poverty threshold. Five of the 12 characteristics showed differences that, although sizable in percent difference, had absolute differences of fewer than two percentage points: telephone interview CUs were observed to have more earners, higher income, and to be more urban, while having fewer adults over age 64 and children under 18 in the household. Two of the 12 characteristics resulted in differences of less than one percent (reference person age and CU size).

4.3 Data Quality Indicators

In bivariate analyses of the effect of telephone administration on estimates, we consistently found a meaningful relationship between telephone administration and data quality. Telephone interview cases tended to exhibit lower rates for “positive” data quality indicators, such as (a) use of recall aids, including financial records, purchase receipts, and the Information Booklet, (b) positive responses to screener questions¹², and (c) interview length, while exhibiting higher rates for “negative” data quality indicators such as (d) reporting of expenditures at an inadequate level of detail, (e) item nonresponse, (f) data editing, and (g) reporting of rounded values.

The results showed a consistent association between telephone interviews and lower data quality relative to personal visit interviews. Interestingly, subsequent multivariate analyses of interview length, controlling for sample characteristics and use of recall aids, suggest that the use of recall aids is the more powerful predictor of data quality differences, rather than mode alone (see Table 2). This finding with respect to the importance of recall aids is further supported in the discussion of outcome measures in Section 4.4.

4.4 Outcome Measures

4.4.1 Regressing Expenditure Amounts on Mode

To examine the association between telephone administration and total expenditures, while controlling for other variables associated with the dependant variable, we specified an ordinary least squares (OLS) model regressing weighted total expenditures on telephone administration and 10 control variables. In Table 3, we present coefficient estimates from four models regressing total expenditures on various combinations of the explanatory variables¹³. The first model – which excludes variables for Information Booklet use and record use – shows telephone administration to have a statistically significant and negative association with total expenditures.

In the second model, we introduce the Information Booklet variable as a control, and observe that the size of the telephone administration coefficient decreases in magnitude dramatically (from -\$253 to -\$15), the telephone administration coefficient loses its statistical significance, and the Information Booklet coefficient shows statistical significance. In the third model, we see similar results after replacing the Information Booklet use with record use. The telephone administration coefficient decreases sharply (from -\$253 to -\$86), the telephone administration coefficient loses its statistical significance, and the record use coefficient shows statistical significance.

Finally, in the fourth model, we include both Information Booklet use and record use as control variables. In this “full” model, not only is the telephone administration coefficient much lower (from -\$253 to \$59), but the sign changes from negative to positive. Further, the telephone administration coefficient again loses its statistical significance, and both

¹⁰ Given the large, pooled sample sizes from the four-year analysis file (n=118,698), estimates of statistical significance were not calculated for the contingency table findings presented in this paper. Instead, we focused on a discussion of “meaningful” differences, defined as differences of at least 3 percentage points and at least 5 percent.

¹¹ In the CE Interview Survey, the reference person is the first member mentioned by the respondent when asked to “Start with the name of the person or one of the persons who owns or rents the home.” It is with respect to this person that the relationship of the other consumer unit members is determined.

¹² In the CE Interview Survey, screener questions generally precede lists of similar expenditure types.

¹³ The variables presented in the regression results are grouped by type: mode (1); recall aids (2, 3); variables associated with expenditure levels (4-6); and variables associated with differences in sample characteristics by mode (7-11).

the Information Booklet use and record use coefficients show statistical significance. Additional models specified to regress individual components of total expenditures on these same predictor variables resulted in similar findings.

The implications of these results are striking. In assessing the impact of mode after controlling for respondent behaviors that typically differ dramatically between the two modes, we find a pronounced shift in the size, direction, and significance of the mode coefficient. The mode coefficient has a large, negative, and statistically significant impact on expenditures in the model that excludes measures of respondent behavior. However, after controlling for Information Booklet use and record use, the mode coefficient shrinks, becomes positive, and loses its statistical significance.

In terms of impact on expenditure levels, the results suggest that mode itself matters less than does the respondent behavior typically associated with mode. To the extent that telephone respondents tend to use recall aids at a lower rate than do personal visit respondents, the expenditure levels of telephone respondents decrease accordingly. When respondents do use recall aids, expenditures increase in comparability. This implies that successfully encouraging the use of recall aids during telephone interviews does much to promote mode equivalence and reduce mode bias.

Additionally, the difference in the magnitude of the Information Booklet use and record use coefficients is notable. The results indicate that record use has a much greater proportionate impact on expenditures than does Information Booklet use (\$1,135 to \$259). This has important implications for a mixed-mode survey redesign, given that designing an Information Booklet to mail to or drop off with respondents for use in a telephone interview may be easier to implement than encouraging record use (a more interviewer-dependent behavior) in telephone interviews.

4.4.2 Comparing Expenditure Amounts by Mode

Comparisons of aggregate expenditures by mode showed relatively few differences. However, while total aggregate expenditures appeared relatively similar across mode, further analyses of the number of individual expenditure items reported and the costs associated with each item showed that telephone interviews resulted in *fewer* reported expenditure items and *higher* reported expenditure amounts. The causal factors driving the underreporting of individual expenditure items relative to personal visit respondents and the over-reporting of higher average amounts for those same items are unclear (i.e., more costly items may simply be more salient, or telephone interview respondents may be reporting higher amounts for the same types of purchased items). We hypothesize that the combined effect of the underreporting of expenditure items and the over-reporting of expenditure amounts result in fewer differences observed at the aggregate level. The importance of this finding is that the relatively “smooth surface” of aggregate expenditures across mode may in fact mask considerable and systematic underlying differences by type of expenditure.

4.4.3 Comparing Expenditure Amounts by Questionnaire Section and Mode

In a final look “below the surface” of aggregate reported expenditures, we compared reported expenditure amounts by mode within questionnaire sections. We found that expenditure item reporting appears to cluster by questionnaire section; some questionnaire sections and expenditure items appeared disproportionately sensitive to mode, with large differences, while others were comparatively robust, showing few differences. This finding is important within the larger discussion of design changes that might be implemented to accommodate mixed-mode interviewing. For example, were the survey to be administered in a modular, rotating panel design, certain sections might be considered for telephone interviewing, while others might be limited to personal visit interview administration only.

5. Conclusion

5.1 Summary of Key Findings

Five key findings from the analysis have important implications for the future direction of the CE Interview Survey:

- a. A persistent negative relationship was observed between telephone interviewing and seven indirect indicators of data quality.
- b. Lower rates of recall aid use were reported for telephone interview respondents.
- c. Recall aid use was found to improve reporting overall but had more of an impact in personal visit interviews.
- d. Record and receipt use appeared to have more of an impact than Information Booklet use on both data quality indicators and outcome measures.

- e. Telephone interviewing itself was found to have less of an impact on driving differences in key survey estimates than whether recall aids were used during the interview.

Overall, these findings suggest that while telephone interviewing may impact the quality of data collected relative to that of personal visit interviewing, there are clear steps that survey management can undertake to maximize mode equivalence and minimize mode effects.

5.2 Recommendations

Based on these findings, we recommend exploring means of encouraging higher use of recall aids in both modes – for example, through a mailable version of the Information Booklet, or a more user-friendly checklist for records and receipts. We also recommend improving measurement of recall aid use through more explicit coding guidelines, and developing new recall aid measures with greater specificity. One limitation of the analysis presented in this paper is the unknown level of variation in coding of recall aid use by interviewers. For the future, we also recommend assessing sensitivity to mode on a question-by-question basis, rather than at an aggregate survey level. The finding that different types of questions are affected by mode in different ways supports the notion that mode effects are a question-by-question issue, *not* a survey-level issue.

Acknowledgements

The views expressed in this paper are those of the authors and do not necessarily reflect the policies of the U.S. Bureau of Labor Statistics. The authors would like to thank Jeanette Davis, Jennifer Edgar, James Esposito, Scott Fricker, Jeffrey Gonzalez, Sioux Groves, Bill Mockovak, Carolyn Pickering, Jay Ryan, Roberta Sangster, and Lucilla Tan for helpful discussions of the Consumer Expenditure Interview Survey and earlier versions of this paper.

References

- Allen, T. M. (1993). "Analysis of Cost and Performance Changes since the Introduction of Telephone Interviewing." Internal Census Bureau memorandum, February 26th.
- Aquilino, William (1994). "Interview Mode Effects in Surveys of Drug and Alcohol Use: A Field Experiment." *The Public Opinion Quarterly*, Vol. 58, No. 2. (Summer, 1994), pp. 210-240.
- Consumer Expenditure Interview Survey Public Use Microdata Documentation (2005). U.S. Department of Labor.
- de Leeuw E.D., van der Zouwen J. (1988). "Data quality in telephone and face to face surveys: a comparative meta-analysis." In: Groves RM, Biemer PP, Lyberg LE, Massey JT, Nicholls WL II, Waksberg J, eds. *Telephone Survey Methodology*. New York: Wiley: 273:99.
- Groves, R. (1989). *Survey Errors and Survey Costs*. New York: John Wiley.
- Groves, R. and Kahn, R. (1979). *Surveys by Telephone: A National Comparison with Personal Interviews* (Academic Press, 1979)
- Hochstim, J.R. (1967). "A critical comparison of three strategies of collecting data from households." *Journal of the American Statistical Association*. 1967;62:319:976-989.
- Huggins, V. J. (1994). "Effect of Telephone Interviewing on Type A and Record Usage Rates." Internal Census Bureau memorandum to S. L. Durant, September 20, 1994.
- King, K. E. (1995). "SIPP 1986-1993: Monitoring the Effects of Telephone Interviewing - for Analysis of Type A and Type D Noninterview Rates." Internal Census Bureau memorandum, February 16, 1995.
- Krieger, S. (2007). "2006 Response Rates for the Diary and CE Interview Surveys." Internal Bureau of Labor Statistics memorandum to D. Swanson, June 19, 2007.
- Martin, E., J. Childs, T. DeMaio, J. Hill, C. Reiser, E. Gerber, K. Styles, and D. Dillman. (2007). "Guidelines for Designing Questionnaires for Administration in Different Modes." U.S. Census Bureau, Washington, DC 20233.
- McGrath, D. (2005). "Comparison of Data Obtained by Telephone versus Personal Visit Response in the U.S. Consumer Expenditures Survey." Paper presented at the Joint Statistical Meetings of the American Statistical Association.
- Nelson, C. (2001). "Survey of Income and Program Participation Users' Guide (Supplement to the Technical Documentation)." Westat. 3rd Edition Washington, D.C.
- Rogers, Theresa (1976). "Interviews by Telephone and in Person: Quality of Responses and Field Performance." *The Public Opinion Quarterly*, Vol. 40, No. 1. (Spring, 1976), pp. 51-65.

Verlander, R. (2006). "Survey Mode Effect on Data Quality: Telephone vs. Personal Visit." Internal BLS Analysis.
 Wessel, Christina, Wendy Rahn, and Tom Rudolph (2000). "An Analysis of the 1998 NES Mixed-Mode Design."
 University of Minnesota. Prepared for NES Board of Overseers meeting, La Jolla, CA, February 12, 2000.

Figures & Tables

Figure 1: Telephone Administration Rate by Interview and Quarter

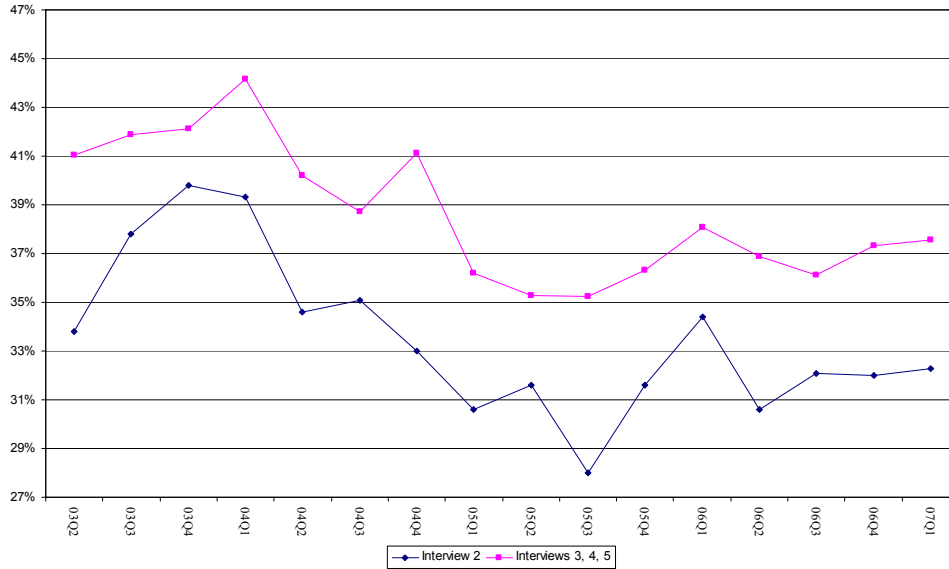
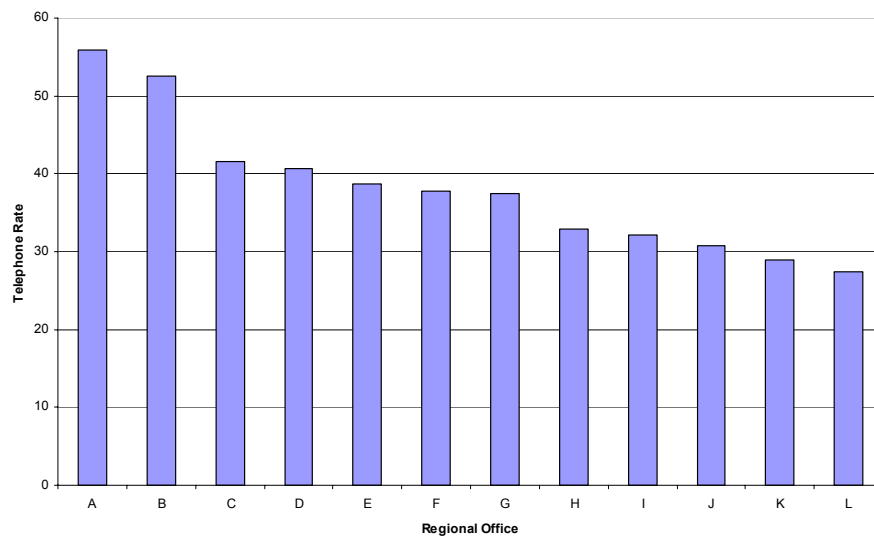


Figure 2: Telephone Administration Rates by Regional Office¹⁴



¹⁴ Regional Office locations have been replaced with letter references for disclosure reasons.

Table 1: Sample Characteristics by Mode

Characteristic	Personal Visit Estimate	Telephone Estimate	Point Difference	% Difference
Mean Consumer Unit Characteristics				
Age of Reference Person	49.2	48.8	-0.4	-0.8
Average Number in Consumer Unit				
Persons	2.50	2.50	0.0	0.0
Children under 18	0.67	0.65	0.0	-3.0
Persons 65 and over	0.34	0.29	-0.1	-14.7
Earners	1.30	1.40	0.1	7.7
Income Before Taxes	46,076	47,732	1,656	3.6
Percent Distribution				
Urban	75.6	76.8	1.2	1.6
Homeowner	66.1	71.2	5.1	7.7
Below Poverty Threshold	16.2	13.1	-3.1	-19.1
Reference Person Characteristics				
Hispanic Origin	12.1	9.0	-3.1	-25.6
Female	51.0	54.2	3.2	6.3
Bachelors Degree or Higher	25.1	30.4	5.3	21.1

Table 2: Use of Recall Aids by Mode

Use of Recall Aids	Obs	Personal Visit Interview			Telephone Interview			
		% of Total Cases	Median Interview Length (min.)	% Increase over 'None'	Obs	% of Total Cases	Median Interview Length (min.)	% Increase over 'None'
None	17,693	25.0	50.1	-	27,174	63.6	47.1	-
Information Booklet Only	25,626	36.2	60.4	20.5	4,684	11.0	51.2	8.5
Record Use Only	6,756	9.6	66.6	33.0	8,383	19.6	61.2	29.7
Information Book & Record Use	20,641	29.2	77.4	54.4	2,479	5.8	66.9	41.9
<i>Overall</i>	<i>70,716</i>	<i>100.0</i>	<i>63.3</i>	<i>-</i>	<i>42,720</i>	<i>100.0</i>	<i>51.5</i>	<i>-</i>

Table 3: Multiple Regression Results for Total Expenditures

Explanatory Variables	Model 1: Restricted (β)	Model 2: Info Book (β)	Model 3: Record Use (β)	Model 4: Full Model (β)
(Intercept)	4,033	3,794	3,872	3,742
Telephone Administration	-253	[-15]	[-86]	[59]
Info Booklet Use		414		259
Record/Receipt Use			1,221	1,135
Age of Householder	-15	-14	-18	-18
Family Income	0.07	0.07	0.07	0.07
Family Size	932	945	937	950
Home Owner	2,102	2,113	1,948	1,976
Family Income Less Than Poverty Threshold	-1,889	-1,898	-1,789	-1,812
Hispanic Origin of Householder	-790	-808	-718	-741
College Educated Householder	2,934	2,934	2,813	2,828
Female Householder	-436	-445	-476	-480

All coefficient estimates are statistically significant at $p < .05$ except for those enclosed by []