USING PARADATA TO EXAMINE THE EFFECTS OF INTERVIEWER CHARACTERISTICS ON SURVEY RESPONSE AND DATA QUALITY

Adam Safir, Tamara Black, and Rebecca Steinbach, The Urban Institute
Adam Safir, UI, 2100 M Street NW, Washington, DC 20037

Key Words: Interviewer effects, error, bias

1. INTRODUCTION

This research began with the premise that both survey organizations and data users can benefit from the release of paradata. The release of paradata, or administrative data, to accompany other public use survey data make possible a more comprehensive and independent evaluation of non-sampling error in survey estimates of interest, and is in keeping with Deming's idea of a "full-systems thinking" approach to data quality (Scheuren 2001). We attempt to demonstrate the feasibility of this approach through a practical example using data from the National Survey of America's Families (NSAF), with a particular emphasis on interviewer-related error.

While macro paradata represent global process summaries (e.g., sampling error or response rates), micro paradata (hereafter to be referred to as simply paradata) provide process details about specific interviews on a case-by-case basis, such as how many attempts were required to contact a sampled unit or the tenure of the interviewer who completed the case. In general, survey organizations do not release an exhaustive set of paradata items. This may be for any number of reasons, including legitimate confidentiality concerns, lack of researcher interest, or inability by data producers to see any clear value added (Scheuren 2001).

The two complimentary research objectives, therefore, were: (1) to better understand the costs and benefits of releasing paradata along with other survey data on public use files, and, endemic to this process, (2) to examine the relationship between known interviewer characteristics and data quality. Using administrative data from the NSAF, the analysis described in this paper grouped telephone interviewers according to predetermined characteristics and compared measurements on survey items of interest.

2. SURVEY BACKGROUND

The NSAF is a survey of the economic, health, and social characteristics of children, adults under the age of 65, and their families. Two rounds of interviews have been conducted. The first round was conducted from February 1997 through November 1997, and the second round from February 1999 through October 1999. Each round of interviews yielded information on over 40,000 households and 100,000 persons. Westat conducted the data collection for the NSAF.

The survey is representative of the civilian, non-institutionalized population under age 65, and features an over-sample of low-income households with children at the state level. Large representative samples of households were taken in each of 13 targeted states and the balance of the nation. The 13 states were Alabama, California, Colorado, Florida, Massachusetts, Michigan, Minnesota, Mississippi, New Jersey, New York, Texas, Washington, and Wisconsin. These 13 states represent over half of the U.S. population and reflect a broad array of government programs, fiscal capacity, and approaches to child well being. Collectively, the survey estimates form what we believe to be a sound baseline from which many of the changes brought about during the period of devolution can be measured and assessed (Converse, et al 2001).

3. INTERVIEWER-RELATED ERROR

In general, error can be defined as the difference between a survey answer and the true value of what the researcher is interested in measuring. The degree of interviewer-related error (only one component of total survey error) in the measurement process is related to the extent to which interviewers can be associated with the answers they obtain (Fowler and Mangione 1990).

Unique interviewer characteristics may affect survey responses for a number of reasons. For example, interviewers have a range of skill sets — some excel at gaining cooperation and overcoming the objections of

reluctant respondents, while others are better able to administer a standardized interview. Second, certain interviewer characteristics may alter the context or meaning of questions. Finally, interviewer characteristics may affect the quality of the relationship between the interviewer and respondent (Fowler and Mangione 1990). In other words, interviewers may alter the delivery of the survey script depending on factors related to tenure, length of shift, and/or perception of the survey.

Interviewer error is particularly important in telephone surveys, where a small number of interviewers may complete a large number of interviews (Singer, et al 1983). To the extent that there exists a systematic variation in responses directly attributable to the interviewer, paradata represent a useful tool for ascertaining the degree, and in some cases, the ignorability, of interview-related error.

4. METHODS

4.1 Analytic Objectives

Our intention was to accomplish the first research objective, assessing the utility of paradata, through the pursuit of the second objective: analyzing the relationship between known interviewer characteristics culled from paradata on the public use files and items reflecting various components of survey response. The three hypotheses stemming from this second objective can be described as follows:

Skill. It was hypothesized that the relative skill of the interviewer might influence the interviewer-respondent relationship in such a way as to produce a noticeable effect on survey statistics.

Tenure. Second, it was hypothesized that interviewers who had worked on the survey in the past, and were therefore more familiar with the questionnaire and subject matter, might have developed habits that would result in a perceptible and identifiable impact on the interview.

Experience. Finally, it was hypothesized that the effects of current accumulated experience on the survey would be evidenced in the survey statistics.

4.2 Variable Descriptions

Independent Variables. Two broad categories of interviewing skill are (1) effectiveness in gaining cooperation and (2) ability in asking survey items and recording responses accurately. While a more complete analysis of interviewing skill would include paradata variables reflecting as many of these measures as possible (e.g., cooperation rate, percent of monitored questions asked exactly as worded, accuracy in data entry, etc.), the sole measure of skill contained in the administrative data was cooperation rate quartile. Therefore, two paradata variables were selected of this type: SCCOCOOP, which grouped interviewers into screener cooperation rate quartiles, and EXCOCOOP, which grouped interviewers into extended interview cooperation rate quartiles.

For tenure, the variable EXCOWRKN, which indicated whether the interviewer had worked on the survey in the first round of data collection, was selected. This variable was meant to provide a measure of the interviewer's overall familiarity with survey.

Finally, to measure current accumulated experience, the variables INTCNT and TOTINT were selected. For each record, INTCNT indicated the number of cases completed by the interviewer who completed that particular interview. The value of TOTINT reflected the total number of interviews completed by the interviewer who completed a specific case.

Dependent Variables. The two types of dependent variables selected for the analysis were specific questionnaire items and more general (or global) survey response measures.

To analyze the effect of interviewer characteristics on response to questionnaire items, questions were classified into four groups: factual/non-sensitive, factual/sensitive, subjective/non-sensitive, and subjective/sensitive. In the context of the NSAF, the term "sensitive" refers mainly to questions that the respondent might consider sensitive relative to the other questions in the survey. The term "subjective" is used to distinguish between personal items, such as opinion questions, and more factual items, such as questions about welfare receipiency or family income.

Using this classification, we expected to see a loose hierarchy of effects. That is, the subjective/sensitive measures were expected to be the most likely to be influenced by characteristics of the interviewer, and the factual/sensitive and subjective/non-sensitive variables to a lesser extent. The factual/non-sensitive variables were selected primarily for control purposes.

The analysis also examined three more global measures of survey response and data quality. These included

interviewer rates of income switching, average interview length, and item-nonresponse. Rates of income switching, along with average interview length and item non-response rates on selected variables were determined by interviewer for all cases included in the final analysis file.

4.3 Analysis

Logistic regression was employed to test the hypotheses regarding skill, tenure, and response to groups of questionnaire items (figure 1). The logit model included the screener cooperation rate quartile of the interviewer who completed the case (SCQ), the extended cooperation rate quartile of the interviewer who completed the case (ECQ), and the tenure of the interviewer who completed the case (WR1). Dummy variables were created for the cooperation rate quartile variables in the model. Respondent race (RR) and metropolitan status (MET) were entered into the model as a way of controlling for effects related to sample composition.

Figure 1. Logit Model

$$\text{logit } (Y) = \alpha + \beta(SCQ) + \beta(ECQ) + \beta(WR1) + \beta(RR) + \beta(MET)$$

Chi-square tests of independence were used to test the survey measures hypothesis on experience. Linear regression was employed to relate skill, tenure and experience (total interviews completed, or "TI") to the more global measures (figure 2).

Figure 2. OLS Model

$$Y = \alpha + \beta(SCQ) + \beta(ECQ) + \beta(WR1) + \beta(TI)$$

Although we expected the screener and extended cooperation rate quartile variables to be highly correlated, in fact they were not. And as can be seen in figure 3, the distribution of interviewers across screener and extended cooperation rate quartile does not display the dominant clustering along the diagonal that one might expect.

Figure 3. Cooperation Rate Quartile Comparison

Screener	Extended				
	1st	2nd	3rd	4th	Total
1st	**28.3**	32.1	21.7	17.9	100
2nd	23.7	**30.7**	24.6	21.1	100
3rd	22.2	27.8	**34.1**	15.9	100
4th	7.5	25.3	18.1	**48.9**	100

While close to 50% of the 4th screener cooperation rate quartile interviewers can be found in the comparable extended cooperation rate quartile, just 28% of the 1st screener cooperation rate quartile interviewers are in the 1st extended cooperation rate quartile.

Finally, it is important to note that in analyzing the regression results, there was less interest in the summary statistics for the model's explanatory power, as there was in examining the behavior of the coefficients to shed light on whether there was a non-random, statistically significant difference across interviewer classifications that might contribute bias to the sample results.

4.4 Controls

Under ideal survey conditions, interviewer assignment is totally random and interpenetrated; however, in practice, this is rarely the case. Some interviewers work exclusively the day shift, others the evening shift, and still others, designated as refusal conversion specialists, may be assigned only the most difficult cases. To account for the non-random assignment of cases to interviewers, the analytic data set was limited to those completed cases which had never refused and which had been started and completed by the same interviewer. This latter component was intended to control for completed break-offs, in which a different interviewer completed a case that had been started by another interviewer. To control for additional possible differences related to sample composition, the data set was also limited to households with children. The final sample size for the analysis file was 12,711 cases.

5. FINDINGS

5.1 Questionnaire Items (Skill and Tenure)

Subjective/non-sensitive. The first group of variables tested in the logistic regression model were the subjective/non-sensitive variables. These were mainly variables measuring opinions about issues such as welfare, parenthood, and childbearing. As can be seen in table 1, the final analysis revealed very little association between characteristics of the interviewer and question response. Although some of the regression coefficients showed significance, no clear patterns emerged.

Subjective/sensitive. The subjective/sensitive variables included indices measuring parental aggravation, behavioral problems of children, children's school engagement, and mental health of parents. For these variables, the regression model was predicting that the

response would be what one might consider "non-sensitive," that is, no parental aggravation, no behavioral problems, etc.

The regression coefficients for these variables revealed more interesting trends. Table 1 displays the observed patterns in the screener cooperation rate coefficients. Although the results for some of the items suggested that interviewers in higher cooperation rate quartiles obtained more sensitive responses, the pattern was considered inclusive because of the absence of significant results. On other items, somewhat more curious patterns were observed. For example, in some cases the first and fourth cooperation rate quartile interviewers appeared more similar in relative size of the coefficient, or the coefficients for the second and fourth cooperation rate quartile interviewers appeared more similar, but a stepped, ordered relationship between cooperation rate quartile and item response was not evident. The regression coefficients for the extended cooperation rate mirrored these results.

Factual/sensitive. The factual/sensitive measures were comprised of items such as citizenship, health insurance, high school education, family poverty, marital status, and interruption in telephone service. In the analysis of the factual/sensitive measures, the logistic regression was constructed to predict a "sensitive" response, such as non-citizenship, lack of insurance, no high school education, etc.

For the factual/sensitive measures, a stepped, ordered pattern was seen on citizenship, education, and poverty. In other words, as the interviewer moved into higher cooperation rate quartiles, he/she elicited more sensitive responses. However, the health insurance, marital status, and telephone service interruption variables seemed to move in the opposite direction and also displayed non-ordered effects.

The results became even more curious when the regression coefficients of the extended cooperation rate quartile were examined. Here, the ordered effects were evident, but moving in the opposite direction. That is, the positive impact of the coefficient was weaker for interviewers in higher extended cooperation rate quartiles (i.e., fewer sensitive responses were obtained). Again, for these variables, the coefficients for citizenship, education, and poverty were significant.

5.2 Questionnaire Items (Experience)

The analysis also examined the accumulated survey experience of interviewers in the current round of data collection. For this effort, which used contingency table analysis, cases were grouped according to the number of interviews completed by the interviewer before completing that particular case. To examine learning effects early in the experience curve, the first 20 interviews completed by the interviewer were classified into quartiles and aggregated across all cases. To examine long-term learning effects, the first 200 interviews completed by each interviewer were grouped into deciles and aggregated across all cases.

The results of this analysis showed no significant differences in the distribution of the variable groupings for the early-stage learning effect quartiles. The long-term learning effect groupings also did not display significant differences among the variable groups by interviewer classification.

5.3 Global Measures

Income Switching. In an effort to improve the precision of estimates of low-income families, the NSAF sample design included an over-sample of families below 200% of the federal poverty level. The survey sub-sampled non-low-income families using a single question on the screener to determine income, and then used a series of questions on the extended interview to generate a more comprehensive estimate of family income.

Income switching occurred in the survey when a sampled unit screened in at one income level, but was determined to be at a different income level during later in the interview. A "false negative" occurred when a household incorrectly screened in as high income, but was determined on the extended interview to be low income. Similarly, a "false positive" occurred when a household reported being low income on the screener but was revealed on the extended interview to be non-low-income.

Due to the differential probabilities of selection specified by the sample design, false negatives were assigned larger weights relative to true positives (households which correctly screened in as low-income). This results in an increase in the variance of survey estimates for the low-income sample. Alternatively, false positives create sampling inefficiencies and lead to an increased cost of survey administration. For these reasons it is advantageous to minimize both the false negative and false positive rates on the survey.

In examining income-switching rates by interviewer, the regressions showed a few minor trends but no definitive patterns. False positive rates tended to increase for interviewers in higher screener cooperation rate quartiles, but paradoxically tended to decrease for the higher extended cooperation rate interviewers. Both false positive and false negative rates tended to decrease for interviewers who had worked in the first round of the survey, although these results were not significant. Experience, as measured by total number of interviews, had no discernable effect on income switching rates.

Interview Length. In computing mean interview length by interviewer, the analysis controlled for income, presence of children, number of child interviews, and presence of spouse/partner in the household, all factors that influence the number of items asked during the survey. The regressions showed mixed results and few trends. Interviewers in higher screener cooperation rate quartiles tended to have longer average interviews; however, the opposite proved true for extended cooperation rate. While tenure appeared to be associated with conducting shorter interviews, experience seemed to have no effect.

Item Non-Response. Finally, the analysis examined interviewer item non-response rates. Here again, the regressions produced few significant results, although the analysis did reveal a significant relationship between screener cooperation rates and item non-response on the race question. As interviewers move into higher screener cooperation rate quartiles their item non-response rate on race tended to increase. Additionally, there was some indication that working in round 1 was associated with higher imputation rates across all the variables tested.

6. DISCUSSION

6.1 Paradata

The first objective of this research was to gain knowledge on effective use of paradata, including a better understanding of the costs/benefits paradata represent to both the data producer and data user. It is clear that paradata extend a useful tool to researchers who are interested in examining the quality of data for themselves, beyond that which is communicated through response rates and sampling error.

However, the utility for the data user is tempered by the complexity of the data. Learning to use the paradata variables efficiently in order to examine potential bias of interest is a non-trivial matter and may represent a significant barrier to some researchers. Alternatively, the survey system itself may choose to shoulder the burden of providing more user-friendly summary variables based on paradata source variables, although this clearly presents an added cost to the organization, both in terms of anticipating items of interest, as well as in increased programming and documentation costs. However, there is no question that as response rates continue to decline on a national level, additional measures of survey quality are of increased importance.

6.2 Interviewer Effects

After investigating the relationship between survey response and interviewer characteristics such as skill, tenure, and experience, the analysis found the patterns of response to be fairly similar across interviewer classifications. In the few cases in which significant differences were evident, the trends were inconsistent and inconclusive. Therefore, it was surmised that interviewer effects stemming from tenure and experience were close to undetectable, using the variables employed in this analysis. However, the analysis of effects related to skill does merit further examination, particularly with regard to the observed phenomena of screener and extended interview cooperation rates having seeming opposing effects.

This analysis was limited by a number of factors. Because actual cooperation rate values are not provided, the range, or spread, both within and across cooperation rate quartiles was lost. In addition, some critical dimensions of interviewer skill were omitted. These were mostly variables that might measure the actual administration of the survey, such as percent of questions monitored not read exactly as worded, or instances of directive probing. These dimensions would have been very attractive to analyze, but were simply unavailable in the administrative data. Finally, the restrictive controls used to build the analysis sets probably also served to reduce the variability of the estimates, but this was believed to have been a critical component of facilitating a clean look at the relationship between interviewer characteristics and characteristics of the interview.

7. REFERENCES

Fowler, Floyd J. Jr. and Thomas W. Mangione. 1990. *Standardized Survey Interviewing*. Newbury Park: Sage Publications.

Converse, Nathan, A. Safir, and F. Scheuren. 2001. "No. 11: 1999 Public Use File Data Documentation." http://newfederalism.urban.org/nsaf/methodology1999.html. (Accessed November 14, 2001)

Scheuren, Fritz. 2001. "Macro and Micro Paradata for Survey Assessment" *1999 NSAF Collection of Papers* Washington, D.C.: The Urban Institute. *Assessing the New Federalism* Methodology Report No. 7.

Singer, Eleanor, M. Frankel, and M. Glassman. 1983. "The Effect of Interviewer Characteristics and Expectations on Response*." Public Opinion Quarterly* 47:68-83

Table 1. Logit Model Results[1]

Variable	Intercept B	SE	1st SCQ B	SE	2nd SCQ B	SE	3rd SCQ B	SE	1st ECQ B	SE	2nd ECQ B	SE	3rd ECQ B	SE	Worked C1 B	SE	Race (Black) B	SE	Race (Other) B	SE	Metro Status B	SE
Subjective/non-sensitive																						
Welfare Helps People	-1.39	0.23	0.18	0.16	0.21	0.15	0.27	0.12	-0.18	0.19	-0.22	0.15	0.08	0.14	**-0.42**	0.18	0.26	0.16	-0.52	0.33	0.00	0.13
Single Parents Effective	-0.06	0.17	-0.15	0.13	-0.07	0.11	0.00	0.10	0.07	0.11	0.02	0.10	0.00	0.08	0.05	0.13	-1.06	0.14	0.15	0.16	-0.06	0.10
Want Kids Should Marry	-2.64	0.26	-0.21	0.19	0.04	0.21	0.22	0.21	0.29	0.28	0.09	0.22	0.02	0.19	0.14	0.25	-0.34	0.27	0.35	0.33	-0.35	0.23
Working Moms Effective	-1.03	0.18	-0.07	0.14	-0.17	0.13	0.03	0.12	0.02	0.13	-0.06	0.13	0.19	0.12	-0.05	0.14	-0.44	0.14	-0.55	0.20	0.00	0.10
Welfare Works Less	-0.52	0.18	0.12	0.11	0.10	0.12	0.02	0.11	-0.30	0.17	-0.10	0.16	**-0.34**	0.12	-0.38	0.14	0.39	0.12	-0.06	0.23	0.02	0.10
Fair/Poor Health	-3.57	0.42	-0.33	0.27	**-0.42**	0.21	**-0.48**	0.23	0.45	0.30	0.35	0.27	0.37	0.26	0.20	0.32	0.96	0.22	0.33	0.50	0.29	0.21
Confidence in Health Care	-2.57	0.39	-0.31	0.24	-0.26	0.19	-0.21	0.19	0.60	0.34	0.20	0.25	0.04	0.25	-0.06	0.30	0.50	0.20	0.35	0.32	-0.19	0.17
Subject/sensitive																						
Parental Aggravation	1.92	0.31	0.13	0.17	0.12	0.19	0.10	0.18	0.22	0.27	0.21	0.18	0.09	0.17	0.29	0.26	-0.83	0.19	-0.47	0.28	0.09	0.16
Behavioral Problems (A)	2.18	0.53	0.47	0.30	0.14	0.29	-0.03	0.30	-0.44	0.38	-0.31	0.29	-0.46	0.32	0.73	0.49	-0.24	0.38	-0.68	0.71	0.18	0.27
Behavioral Problems (B)	3.24	0.49	0.35	0.37	**0.71**	0.32	0.40	0.30	-0.58	0.48	-0.75	0.44	-0.68	0.37	-0.48	0.38	-0.38	0.31	0.53	0.53	-0.24	0.26
School Engagement	1.12	0.22	**0.60**	0.16	0.32	0.16	**0.44**	0.14	**-0.39**	0.17	-0.23	0.15	**-0.23**	0.12	0.34	0.19	-0.29	0.16	0.20	0.36	-0.05	0.11
Negative Mental Health	1.94	0.27	**0.40**	0.14	0.18	0.15	0.23	0.13	**-0.38**	0.18	-0.34	0.13	**-0.45**	0.11	0.10	0.21	-0.36	0.14	-0.13	0.29	-0.28	0.14
Factual/sensitive																						
Citizenship	-2.96	0.35	**-1.53**	0.31	**-1.51**	0.20	**-0.66**	0.17	**1.24**	0.26	**1.06**	0.23	**0.90**	0.25	0.39	0.26	-0.59	0.21	0.79	0.27	-1.06	0.29
Health Insurance	-2.36	0.37	-0.25	0.20	**-0.34**	0.16	**-0.33**	0.14	0.36	0.22	0.18	0.17	0.29	0.18	0.06	0.31	0.23	0.17	-0.32	0.30	0.21	0.15
HS Education	-3.10	0.28	**-0.64**	0.23	**-0.60**	0.20	**-0.42**	0.16	**1.09**	0.22	**0.93**	0.17	**0.94**	0.17	0.46	0.24	0.24	0.17	0.07	0.27	0.17	0.16
Poverty	-0.87	0.18	**-0.44**	0.12	**-0.27**	0.13	**-0.27**	0.10	**0.58**	0.15	**0.27**	0.13	**0.24**	0.12	-0.04	0.15	1.09	0.11	-0.12	0.20	0.54	0.08
Marital Status	-1.49	0.21	-0.03	0.12	-0.03	0.10	-0.18	0.11	0.16	0.15	0.14	0.11	0.03	0.11	0.27	0.17	1.70	0.11	0.24	0.24	-0.01	0.10
Telephone Interruption	-2.94	0.35	-0.08	0.23	-0.16	0.21	-0.17	0.21	0.32	0.18	0.10	0.20	0.25	0.16	-0.03	0.29	1.33	0.16	0.22	0.36	0.65	0.18
Child Working	0.65	0.32	0.11	0.19	-0.15	0.19	0.26	0.19	0.07	0.24	0.15	0.19	0.09	0.19	-0.01	0.27	1.22	0.27	1.09	0.42	-0.17	0.14
Suspended/expelled	-2.09	0.41	0.05	0.31	-0.37	0.33	-0.03	0.21	0.17	0.34	0.14	0.27	0.44	0.30	-0.25	0.32	0.69	0.27	0.66	0.47	0.26	0.19

Table 2. OLS Model Results

Variable	Intercept B	SE	1st SCQ B	SE	2nd SCQ B	SE	3rd SCQ B	SE	1st ECQ B	SE	2nd ECQ B	SE	3rd ECQ B	SE	Worked C1 B	SE	Total Int B	SE
Income Switching																		
Total False Negative Rate	12.54	1.26	-1.06	1.42	-1.40	1.45	0.49	1.40	0.50	1.53	0.04	1.38	-0.08	1.39	-0.36	1.49	-0.01	0.01
Total False Positive Rate	28.74	1.67	1.02	1.91	1.47	1.93	1.93	1.87	1.20	2.04	-0.13	1.86	-0.34	1.86	-1.25	1.99	0.00	0.01
False Negative Rate Same Resp	11.16	1.34	-1.64	1.51	-2.53	1.55	-0.46	1.49	1.41	1.62	-0.92	1.47	-0.39	1.47	0.31	1.59	-0.01	0.01
False Positive Rate Same Resp.	25.44	1.76	1.33	1.99	0.51	2.02	2.16	1.95	-0.31	2.14	-0.26	1.95	-1.05	1.94	-1.55	2.07	0.01	0.01
False Negative Rate Diff. Resp.	15.30	2.13	-1.42	2.37	0.24	2.43	1.59	2.33	0.56	2.54	3.67	2.31	1.60	2.31	-0.15	2.47	0.00	0.01
False Positive Rate Diff. Resp.	34.25	3.04	-1.29	3.32	2.30	3.41	1.50	3.26	7.02	3.60	3.88	3.27	2.68	3.27	-2.70	3.42	-0.01	0.02
Average Interview Length																		
Avg. Int. Length High Income	41.57	0.82	**-1.83**	0.89	**-2.51**	0.92	-1.66	0.88	-0.10	0.96	1.33	0.88	0.48	0.88	-0.30	0.92	-0.01	0.00
Avg. Int. Length Low Income	48.65	1.55	-3.11	1.67	-2.70	1.71	-2.20	1.66	2.99	1.84	2.99	1.67	1.90	1.68	**-3.47**	1.73	**-0.02**	0.01
Item Non-Response Rates																		
UBRACE4	9.26	1.47	**-5.60**	1.67	**-5.08**	1.70	**-3.97**	1.64	1.52	1.79	1.94	1.62	1.26	1.62	0.62	1.76	0.00	0.01
IHRAMT	6.97	0.64	0.77	0.73	0.95	0.75	0.08	0.72	-0.58	0.79	-0.70	0.71	-0.96	0.71	0.41	0.77	0.00	0.00
IPAYAMT	6.69	0.65	0.88	0.74	0.52	0.76	0.77	0.73	-0.50	0.80	-0.12	0.72	0.40	0.72	0.84	0.78	0.00	0.00
ISETOT	2.62	0.38	-0.42	0.43	-0.29	0.44	-0.23	0.43	0.45	0.47	0.36	0.42	0.12	0.42	0.34	0.46	0.00	0.00

[1] Figures in bold are statistically significant at $p < .05$. Significance not shown for control variables (race and metro status).