**Using process data to study interviewer effects on nonresponse in the Consumer Expenditure Survey.**

**John Dixon1, Erica Yu1, and John Gloster2**

***1U.S. Bureau of Labor Statistics, 2U.S. Census Bureau***

**Abstract**

This study estimates interviewer effects on nonresponse in the Consumer Expenditure Survey (CE). The study uses data collected during the CE interview process through the CAPI instrument software, which provides interview audit trails, and the Contact History Instrument (CHI), which provides interviewer contact attempts and concerns expressed by respondents. Multilevel models are used to investigate the relationship between interview processes (e.g., length of interview, mode of collection), and interviewer characteristics (e.g., workload) with nonresponse.

**Key** **Words:** paradata, survey quality, multilevel model.

**Introduction**

Interviewer characteristics and interview processes are commonly thought to impact response rates. For example, larger interviewer workloads can reduce the ability of interviewers to make callbacks (Botman and Thornberry 1992). Jackle and colleagues (2013) found that interviewer experience was related to higher interview rates over a number of European surveys. Groves and Couper (1998) described the potential for interviewer characteristics and the survey collection environment to impact nonresponse. For a review of the interviewer effects literature, see West and Blom (2017).

Measures of interview process characteristics and interviewer characteristics included in this study and their hypothesized relationship to nonresponse are:

 *Interview process characteristics*

* + Number of contacts; with more difficulty contacting a household the less time is available to persuade them.
	+ Date of first attempted contact; Attempted contacts later in the interview period allows less time for making contact, scheduling an interview, and persuasion.
	+ Length of interview; Many long interviews leave less time for the remaining interviews. The lore is that long interviews are tiring for interviewers, which might reduce the time or enthusiasm to complete their interview schedule.

*Interviewer characteristics*

* + High workload limits the amount of time an interviewer has to contact and convince a respondent to participate.
	+ Total survey workload; Working on other surveys can compete for time from the survey we are studying.
	+ Less experience may produce lower response rates.

**Sources of data**

Interviewer characteristics came from administrative data managed by the Census Bureau, including the number of interviews that each interviewer administers and for how many other surveys that interviewer collected data. The Contact History Instrument (CHI) provides contact outcomes and respondent concerns for each attempted contact. The interviewers check all that apply. While the consistency of use of the categories may vary between interviewers, it is our best window into the contacts leading up to a final outcome – response or nonresponse.

Measures of item edits and changes, as proxies for measurement error, came from audit trails generated by the CAPI instrument.

The survey households are in the sample for four quarters. This study uses information from 2017, for which 97,317 households were used for the analysis, with 617 interviewers. Interviewers who were assigned fewer than five households or didn’t provide CHI data for any contact attempts were excluded.

**Results**

*Concern*

The most common concern expressed by respondents was “too busy”, followed by “scheduling difficulties”, and “not interested/does not want to be bothered”, which was also most predictive of a refusal outcome. Other frequent concerns were “interview takes too much time” and “privacy concerns”.

*Caseload*

Analysis of caseload suggests that interviewer-level response rates are not related to the number of interviews an interviewer completes. In both panels of Figure 1, each point represents an interviewer. The Y-axis is the response rate for an interviewer. The X-axis (caseload), shows the number of interviews each FR has. In the left panel, the caseload refers to the CE caseload. Low caseloads have as many 0% response rates as 100% response rates. In the right panel, where the caseload refers to total caseload across all surveys worked, more successful interviewers conduct more interviews.

Figure 1: **Response rates and CE caseload, total caseload**

** **

*Start date of contact attempts*

The lore about interviewers is that those who attempt to interview their households early in the month are more likely to succeed. However, these data show there is no relationship. Figure 2 shows this, as well as the effect of interviewer changes (due to scheduling, caseload balancing, or refusal conversion) on response rate. There is some evidence that interviewer changes are associated with higher response rates, possibly due to refusal conversion experts “saving” interviews, but there is no relationship to first date of attempted contact.

Figure 2: **Response rates and first interview date**

****

*Interview length*

The lore is that long interviews are tiring for interviewers, which might reduce the time or enthusiasm to complete their interview schedule. However, these data suggest there is no relationship (Figure 3).

Figure 3: **Response rates and interview length**

****

*Number of Additional Surveys Worked*

The lore suggests that interviewers who work more surveys have less time for and loyalty to any one survey, leading to lower response rates. Figure 4 indicates a slight trend for the opposite; the more successful interviewers work multiple surveys.

Figure 4: **Response rates and number of additional surveys worked**

****

*Experience conducting the CE survey*

Early studies of experience showed a learning curve, where the more experience an interviewer gets, the better they get at getting interviews (Olson and Bilgen 2011). For the CE, however, there doesn’t seem to be a relationship between experience and interview rate. This null relationship is likely due to the effectiveness of the Census training program.

Figure 7: **Response rates and months working on CE**

****

*Multilevel model*

A multilevel model (R lme4, logistic link) was be used to examine if interviewer effects on nonresponse varied by the predictors explored so far. The slopes (random effects) of interviewers were used to identify those interviewers who had higher response rates after adjusting for factors like caseload and contacts. While some of the predictors predicted nonresponse, the random effect was nonsignificant, indicating the differences between interviewers was small (an ICC of .14 would indicate the small effect). In Table 1, the predictors are grouped by the direction of effect, with significant positive effects in blue and significant negative effects in orange.

Table 1: **Multilevel Model for Nonresponse**

|  |  |  |
| --- | --- | --- |
|  | **Model 1** | **Model 2** |
| **Intercept** | 0.072(0.049)  | 1.971(0.338)  |
| **Concerns** |  | 0.096(0.025) (Pr. <.0001) |
| **Total workload** |  | 0.053(0.015) (Pr. =.0003) |
| **Months of experience** |  | 0.122(0.063) (Pr.=0.052) |
| **Length of interview** |  |  0.113(0.060) (Pr.=0.058) |
| **Date of first attempted contact** |  | -0.087(0.005) (Pr. <.0001) |
| **Number of contacts** |  | -0.101(0.012) (Pr. <.0001) |
| **CE caseload** |  | -0.022(0.011) (Pr. =0.045) |
| **Total surveys** |  | -0.044(0.065) (Pr.=0.505) |
|  |  |  |
| ***Random*** | *0.535(0.732*  | *0.590(0.768)* |
| **ICC** | 0.140 | 0.152 |

**Discussion**

This analysis is one of the first to utilize interview process data to study interviewer effects on nonresponse in the CE survey. These new models of nonresponse provide some evidence that the mythologies around interviewers are not always true. For example, reducing workloads likely would not reduce nonresponse given that larger workloads were associated with higher interviewer-level response rates. However, since this study is not experimental, it is difficult to draw clear conclusions. Other survey management practices or interviewer characteristics not measured here could have an impact. Case assignment practices could confound interviewer practices, e.g., newer interviewers could be given interviews thought to be easier. The nature of the relationship between interviewer and response rate doesn’t tell us what changes could be made to survey practice or interviewer training.

**Future Research**

We need more data on interviewer contact strategies and interviewer behavior (particularly measures which might relate to time management), as well as respondent characteristics. Models that consider interactions could show more complex relationships. Tree models might also be useful for understanding interactions.

**References**

Botman, S.L. and O.T. Thornberry (1992), "Survey Design Features Correlates of Nonresponse," ASA Proceedings of the Section on Survey Research Methods, 309-314.

Brunton-Smith, I., Sturgis, P., and Leckie, G., (2017), “Detecting and understanding interviewer effects on survey data using a cross-classified mixed-effects location scale model”, National Centre for Research Methods Working Paper (2/16). and Journal of the Royal Statistical Society: Series A (Statistics in Society), Vol. 180.10.

Groves, RM. and Couper, MP. (1998) Nonresponse in Household Interview Surveys. New York: John Wiley & Sons, Inc.

Jäckle, A., Lynn, P., Sinibaldi, J. and Tipping, S. (2013). The Effect of Interviewer Experience,

Attitudes, Personality and Skills on Respondent Cooperation with Face-to-Face Surveys.

Survey Research Methods, 7, 1-152.

Loosveldt, G., and Beullens, K., (2014), “A Procedure to Assess Interviewer Effects on Nonresponse Bias”, SAGE Open January-March 2014: 1–12

Olson, K, and Bilgen, I, (2011), Public Opinion Quarterly, Vol. 75, No. 1, Spring 2011, pp. 99–114

West, B, and Blom, A., (2017), “Explaining interviewer effects: A research synthesis’, Journal of Survey Statistics and Methodology (2017) 5, 175–211.