Joint Statistical Meetings, 2004 Toronto
Section: Creative Approaches to Combat the Brutal Environment of Telephone Surveys
Roberta L. Sangster, Organizer

# *Predicting Wave Nonresponse from Prior Wave Data Quality*

Brian J. Meekins and Roberta L. Sangster

Keywords: Data quality, panel attrition

This paper examines the effect of data quality items on subsequent wave nonresponse in a large, national telephone panel survey. The goal of this article is to identify covariates that predict panel nonresponse and generate logistic regression coefficients that could be used in propensity score modeling or nonresponse adjustment. Other covariates are considered, including call history variables and demographics.

Previous research modeling nonresponse in longitudinal surveys for the purposes of weighting or adjustment uses few data quality measures. One notable exception is Rizzo, Kalton, and Brick (2001). These authors found that the number of imputed items on the prior wave was useful for adjustment of nonresponse in the Survey of Income and Program Participation (SIPP). Recent research examining the relationship between item nonresponse and subsequent wave nonresponse has found a fairly strong positive relationship between item and subsequent unit nonresponse (Loosveldt, Pickery, and Billiet 2002). Other research on the Current Population Survey mirrors these results (Dixon 2002).

## Data

The data are from the Telephone Point of Purchase Survey (TPOPS). The TPOPS is conducted by the U.S. Census Bureau on behalf of the Bureau of Labor Statistics. The TPOPS is a rotating panel where respondents are interviewed on four occurrences (quarterly) over the course of one year. It is a nationally representative RDD survey with an eight week calling period. Approximately 35% of the sample in a given quarter is “new” RDD recruited sample. Once completed respondents can return at any other wave – in other words they can re-enter the panel. These data come from quarters beginning in January 2001 through March 2002. “Soft refusals” and 50% of non-contacts in wave one are reintroduced into wave two.

## Methodology

Logistic regression is employed to develop models at each of the three subsequent waves of the survey (waves 2, 3, & 4). At each wave the transition from completion on the previous wave to nonresponse, noncontact, refusal, and attrition on subsequent waves are regressed on a number of covariates that include: data quality measures from the prior wave, call history variables from the prior wave, and a number of demographics. The best models for nonresponse, noncontact, refusal, and attrition are selected at each wave and presented here (for a total of 12). In addition to direct effects, first and second order interaction effects are evaluated, but only a rare few improve model fit.

### *Data Quality Measures*

The TPOPS asks respondents to provide outlet (business) information for purchases of a certain commodity (e.g. men’s outerwear, cruises). Specifically, respondents are asked for the name and location of the outlet and the address of that outlet. If the respondent is unable to provide the full address of the outlet then the respondent is asked to provide the nearest intersecting street. In addition, the respondent is asked the amount of the purchase for that specific commodity at that outlet.

The quality of these responses can be easily measured and coded. Because the requested information is wholly volunteered and somewhat difficult to provide, these questions provide a wider range of data quality measure than close-ended questions. Indeed, a respondent can quickly learn that claiming no purchases of a commodity will eliminate the address and amount queries and greatly reduce their burden.

A number of data quality measures were attempted, including the total number of outlets reported, the total number of unique outlets reported, the number of usable outlets[1], the average completeness of address information about the outlets, the number of expenditures

[1] A “usable” outlet is one where the name and address information is sufficient, in itself, to locate the outlet.

reported, and the average amount of rounding for these expenditures. Two measures were constructed that combine information about the number of outlets/purchases claimed the quality of the responses: 1. the number outlets adjusted for the completeness of the address; 2. the number of expenditures reported adjusted for rounding. Many of these measures are, of course, highly correlated, and in some cases directly dependent, so that only a subset could be entered into a model at a given time.

A more commonly employed measure of data quality is the incidence of item missing. We considered two variables, the first indicating if the respondent refused to answer the race question, the second indicating the respondent refused (or didn't know) any of the other demographic questions. Demographic questions are asked only on wave one of the TPOPS, so that if missing on race or any other demographic was used in the model, then that model could not, of course, also include the missing demographic. Here again the authors made a choice between the missingness indicator and the demographic based on model fit.

***Call History Variables***

In addition to demographic and data quality measures, call history variables from the prior wave were entered into the models. Two variables were used primarily. The first indicates whether a respondent was "reluctant" on the prior wave. A reluctant respondent was defined as having a soft refusal on the prior wave or five consecutive callbacks prior to completion. A second indicator was constructed to identify respondents that were "hard to reach" on the prior wave. These respondents were defined as having 10 or more attempts of any combination of ring no answer, busy, or answering machine. These variables were introduced in the model together with the total number of attempts and the interaction of each with the total number of attempts. Finally, the proportion of attempts (or total number in some cases) in the wave in question that occurred at the same time of day on the same day of the week (+- 30 minutes) as the completion of the previous wave was also tested.

***Demographics and Data Quality by Wave***

*Table 1* shows the final disposition of respondents by interview wave. Note the large percentage of sample units that are determined to be ineligible in wave one. This is because wave one is the RDD wave. This number drops off drastically, in subsequent waves, although climbing throughout the course of the panel. Indeed, the increase in ineligible sample units over the waves (calendar year) is large enough to be suspect and it is believed by the authors that , in addition to "bad numbers" as a result of respondents' moving for losing service, some households remove themselves from the sample using the eligibility criteria. Refusals actually decrease through the waves, especially after wave two, where the soft refusals are finally removed.

**Table 1: Panel Characteristics Across Waves**

| *Percent of Sample* | **Wave one** | **Wave two** | **Wave three** | **Wave four** |
|---|---|---|---|---|
| Interview/partial | 25.9 | 51.1 | 52.5 | 52.0 |
| Refusal | 13.1 | 20.2 | 12.6 | 10.6 |
| Non-contact | 3.6 | 8.7 | 9.3 | 9.2 |
| Unknown eligibility | 10.1 | 10.2 | 11.1 | 11.0 |
| Not eligible | 47.3 | 9.7 | 14.5 | 17.2 |
| | | | | |
| % of Wave one Completes | 100.0 | 75.9 | 68.4 | 68.0 |
| % Attrition from previous | 0.0 | 11.3 | 9.4 | 12.9 |
| | | | | |
| *Demographics of Completions* | | | | |
| Number in CU over 62 (Mean) | 0.3 | 0.4 | 0.4 | 0.4 |
| Percent owning home | 69.5 | 72.5 | 74.3 | 75.2 |
| Years in area | 2.2 | 2.4 | 2.5 | 2.8 |
| Mean age of ref person | 47.0 | 48.1 | 48.7 | 49.1 |
| Percent with male ref person | 53.7 | 54.5 | 54.3 | 54.9 |
| Percent with ref person of Hispanic origin | 11.0 | 10.4 | 10.1 | 10.1 |
| | | | | |
| *Data Quality Measures (Completions)* | | | | |
| Number of outlets – gross | 5.2 | 4.6 | 4.5 | 4.5 |
| Adjusted number of outlets | 4.1 | 3.7 | 3.6 | 3.6 |
| Adjusted number of expenditures | 3.3 | 3.0 | 3.0 | 3.0 |

Also shown in *Table 1* is the percentage of wave one completions that are retained in each subsequent wave. Drop off is steep from wave one to wave two and wave two to wave three. The drop-off from wave three to wave four is almost nonexistent. Note that this is not the same as attrition. Because respondents can re-enter the panel after they have completed an interview in wave one, the 68.4% of wave one completions completed in wave three are not necessarily wave two completions.

*Table 1* also displays the percent attrition from the previous wave. Attriters are defined as those leaving the panel and not returning for any subsequent interview. Therefore, 11.3% of wave one completions are not interviewed in wave two through 4, while 9.4% of wave two completions are not interviewed in waves 3 and 4.

*Table 1* also shows the differences in selected the demographic composition of completions on each wave. There is surprisingly little change in many demographic variables. Reference persons tend to be a bit older, households have older respondents and are more likely to be owners who have lived in the area for a bit longer. In addition, respondents after the first wave are less likely to be Hispanic. These differences are not large and there are some glaring omissions. Race did not change, number of people in CU did not change. Sex of reference person changed nominally to male – but reference person is not necessarily the same as respondent gender. Region of the country and marital status did not change.

All measures of data quality were reduced very modestly over time. *Table 1* shows the mean number of outlets, mean adjusted number of outlets, and the mean number of expenditures adjusted by wave. Missingness on race or other demographics could not be ascertained because this information was only collected at wave one.

Such modest changes in both the demographic and data quality measures may lead us to believe that they would have relatively little explanatory power, and indeed, the portion of variance explained by models incorporating these variables is not large. However, some of the effect sizes are quite large and this may help us to understand the relationship between these variables and nonresponse.

**Results**

Models are built in attempt to explain nonresponse, which is then further divided into noncontact, refusal, and attrition. These four dependent variables are regressed on a number of demographic variables, call history variables, and data quality measures. The optimal models for each dependent variable is presented here.

***Nonresponse***

From wave one to wave two a number of covariates were statistically significant, including data quality measures and call history variables. A respondent's refusal to answer the race question in wave one was a fairly strong indicator of nonresponse, decreasing the likelihood of nonresponse by over 50%. The number of expenditures reported, adjusted for

**Table 2: Logistic Regression Nonresponse**

| | Wave two | | Wave three | | Wave four | |
|---|---|---|---|---|---|---|
| | Exp(b) | sig. | Exp(b) | sig. | Exp(b) | sig. |
| CU tenure (Rents) | 1.417 | 0.000 | 1.479 | 0.000 | 1.583 | 0.000 |
| Region (Midwest) | 0.885 | 0.028 | 0.932 | 0.286 | 1.049 | 0.566 |
| Region (South) | 1.171 | 0.003 | 1.143 | 0.029 | 1.258 | 0.000 |
| Region (Northeast) | 1.007 | 0.904 | 0.995 | 0.942 | 1.099 | 0.275 |
| Married | 0.886 | 0.005 | 0.894 | 0.034 | 0.754 | 0.000 |
| Age of reference person | 0.985 | 0.000 | 0.983 | 0.000 | 0.986 | 0.000 |
| PSU size (Large metro) | 1.371 | 0.000 | --- | --- | --- | --- |
| PSU size (Small metro) | 1.063 | 0.440 | --- | --- | --- | --- |
| Number in CU over 62 | 1.138 | 0.001 | --- | --- | --- | --- |
| Hispanic origin | --- | --- | 0.749 | 0.000 | --- | --- |
| CU size | --- | --- | 0.952 | 0.001 | --- | --- |
| Race of ref (Black) | --- | --- | 1.219 | 0.052 | 1.386 | 0.000 |
| Race of ref (Other) | --- | --- | 0.984 | 0.834 | 1.410 | 0.000 |
| Season (Jan-March) | --- | --- | --- | --- | 1.005 | 0.947 |
| Season (April-June) | --- | --- | --- | --- | 0.853 | 0.049 |
| Season (July-Sept) | --- | --- | --- | --- | 0.857 | 0.046 |
| Prop atts at same time/day as prior wave completion | 0.683 | 0.052 | --- | --- | --- | --- |
| Missing on race | 1.564 | 0.016 | --- | --- | --- | --- |
| Adjusted expenditures | 0.947 | 0.000 | 0.948 | 0.000 | --- | --- |
| Reluctant on prior wave | 2.284 | 0.000 | 1.989 | 0.000 | 1.962 | 0.000 |
| Hard to reach on prior wave | 1.356 | 0.000 | 2.179 | 0.000 | 2.153 | 0.000 |
| Average time b/n call attempts | --- | --- | 1.025 | 0.000 | 1.059 | 0.000 |

rounding, is also a significant predictor, although its effect is quite small. Every additional expenditure reported (without rounding) in wave one only decreased the odds of refusal by 5%. Being reluctant or hard to reach on wave one also increased the odds of nonresponse. The effect was quite large for reluctance [exp(B) = 2.284], and somewhat smaller for being hard to reach [exp(B) = 1.356]. The proportion of attempts made at the same time of day and day of week as wave one completion may have decreased the odds of nonresponse but the effect is not statistically significant (at $p <= .05$). Whether a respondent is married, owns or rents their home, their age, and their location in the U.S. are all significant predictors of nonresponse for all waves. Married respondents, those that own their own home, older respondents and those residing in the Midwest (compared to the West) were more likely to respond.

As stated prior, race is only measured at wave one, so that we might not expect refusing to answer the race question to be a significant predictor of nonresponse on subsequent waves. Indeed, this indicator drops out after the wave1 to wave2 transition. However, the race and Hispanic origin of the respondent is a significant predictor of nonresponse from wave two to wave three and race remains a significant predictor of nonresponse from wave three to wave four. The number of adjusted expenditures in wave two remains a weak but statistically significant predictor of nonresponse in wave three but drops out in wave four. Reluctance on prior wave is somewhat less strong in predicting wave three and wave four nonresponse, and being hard to reach on the prior wave surpasses it in strength. This may be expected, that as a panel is extended the uncooperative are eliminated rather quickly, while those that do not respond because they are difficult to contact remain. Breaking out refusals from noncontacts illuminates this relationship.

***Refusal***

*Table 3* shows the results from logistic regression models where odds of refusal are regressed on the same list of covariates. As we might expect reluctance on prior wave is a strong predictor of wave two refusal, and in fact, becomes a much stronger predictor for wave three and wave four refusal. Surprisingly, missing on race in wave one is not a significant predictor of wave two refusal and did not improve the overall fit of the model when it was included. The number of expenditures adjusted for rounding reported in the prior wave is a significant predictor of refusals in every wave, although the size of the effect remains modest. The number of attempts made in the current wave of interviewing is a significant but weak predictor of refusal on each wave. No bivariate interaction effects between call history variables and data quality variables are statistically significant. In addition, no interactions between these variables and any demographics are statistically significant and did not improve model fit.

| **Table 3: Logistic Regression Refusal** | Wave two | | Wave three | | Wave four | |
|---|---|---|---|---|---|---|
| | Exp(b) | sig. | Exp(b) | sig. | Exp(b) | sig. |
| Region (Midwest) | 0.789 | 0.001 | 0.837 | 0.029 | --- | --- |
| Region (South) | 0.870 | 0.033 | 0.970 | 0.769 | --- | --- |
| Region (Northeast) | 1.031 | 0.654 | 0.952 | 0.744 | --- | --- |
| PSU size (Large metro) | 1.328 | 0.003 | 1.325 | 0.008 | --- | --- |
| PSU size (Small metro) | 1.084 | 0.420 | 1.164 | 0.089 | --- | --- |
| CU tenure (Rents) | 0.804 | 0.000 | --- | --- | --- | --- |
| Married | 1.236 | 0.000 | --- | --- | --- | --- |
| Hispanic origin | 1.211 | 0.015 | --- | --- | --- | --- |
| Race of ref (Black) | --- | --- | --- | --- | 1.492 | 0.000 |
| Race of ref (Other) | --- | --- | --- | --- | 1.142 | 0.353 |
| Season (Jan-March) | --- | --- | 0.877 | 0.021 | 0.869 | 0.256 |
| Season (April-June) | --- | --- | 0.840 | 0.009 | 0.709 | 0.008 |
| Season (July-Sept) | --- | --- | 0.847 | 0.014 | 0.638 | 0.000 |
| Adjusted expenditures | 0.932 | 0.000 | 0.911 | 0.000 | 0.936 | 0.006 |
| Reluctant on prior wave | 1.846 | 0.000 | 3.218 | 0.000 | 3.052 | 0.000 |
| Number of attempts | 1.012 | 0.016 | 1.009 | 0.049 | 1.023 | 0.018 |

The size of the primary sampling unit in which the respondent is located was a significant predictor of refusal in waves 2 and 3, as is the region of the country. Large PSUs and the Western region (reference) of the country were more likely to refuse. For wave one, renters were actually less likely to refuse, although this effect disappears in subsequent waves. Another somewhat curious finding is that marital status increases the odds of refusal on wave two. A possible explanation is that being married does increase the burden of the interview – where more purchases are typically reported for married couples and some have to be reported by proxy. This effect is also not present for wave three and four. Race and season of interview do emerge in later waves as significant predictors of refusal, with Black respondents being more likely to refuse. While it is possible that race acts to mediate the effects of PSU size and region, the interactions were tested and found to be insignificant.

***Noncontact***

As shown in *Table* 4, and not surprisingly, the largest predictor of noncontact in any wave is being hard to reach on prior waves. This relationship holds even when controlling for the number of attempts made in prior wave. Being reluctant on a prior wave actually decreased the odds of having a noncontact, probably because repeated contact with the respondent is a stipulation of being coded as reluctant. Interestingly, the number of usable outlets reported in wave one, that is, outlets with fairly complete address information, is positively related to the odds of a noncontact in wave two, while the number of expenditures reported is negatively related. The difference in these effects may be a result of the number of usable outlets being a better indicator of shopping behavior than data quality.

Region is an important predictor of noncontact, with the Northeast and South being the most difficult to contact. Older respondents, respondents of Hispanic origin, and those living in larger Consumer Units (households) are more likely to be contacted. Contact is also related to the season of the interview, with noncontact being more likely in winter and spring. In later waves, marital status, race, and whether a respondent owns or rents their home are significant predictors of noncontact, where renters, unmarried respondents, and non-whites are less likely to be contacted.

***Attition***

Up to this point we have considered nonresponse at each wave, dependent only upon completion in the previous wave and other characteristics in the models. It may be useful to examine true attrition, or those who never re-enter the panel. Thus, we will examine wave two nonrespondents

**Table 4 Logistic Regression Noncontact**

| | Wave two | | Wave three | | Wave four | |
|---|---|---|---|---|---|---|
| | Exp(b) | sig. | Exp(b) | sig. | Exp(b) | sig. |
| Region (Northeast) | 1.393 | 0.000 | --- | --- | 1.382 | 0.024 |
| Region (Midwest) | 1.137 | 0.130 | --- | --- | 1.247 | 0.116 |
| Region (South) | 1.348 | 0.000 | --- | --- | 1.318 | 0.036 |
| Age of reference person | 0.987 | 0.000 | --- | --- | 0.984 | 0.000 |
| Hispanic origin (No) | 0.803 | 0.008 | --- | --- | --- | --- |
| CU size | 0.869 | 0.000 | --- | --- | --- | --- |
| Season (Jan-March) | 1.394 | 0.000 | --- | --- | --- | --- |
| Season (April-June) | 1.173 | 0.065 | --- | --- | --- | --- |
| Season (July-Sept) | 0.580 | 0.945 | --- | --- | --- | --- |
| CU tenure (Rents) | --- | --- | 1.289 | 0.001 | --- | --- |
| Race of ref (Black) | --- | --- | 1.003 | 0.983 | 1.380 | 0.017 |
| Race of ref (Other) | --- | --- | 0.773 | 0.014 | 1.448 | 0.006 |
| Married | --- | --- | --- | --- | 0.557 | 0.000 |
| Usable outlets | 1.057 | 0.001 | --- | --- | --- | --- |
| Adjusted expenditures | 0.936 | 0.003 | --- | --- | --- | --- |
| Reluctant on prior wave | 0.437 | 0.000 | --- | --- | --- | --- |
| Hard to reach on prior wave | 3.709 | 0.000 | 3.355 | 0.000 | 4.344 | 0.001 |
| Number of attempts in prior wave | 1.139 | 0.000 | 1.113 | 0.000 | 1.117 | 0.000 |
| Hard to reach (Yes) *Number of attempts | 0.910 | 0.000 | 0.933 | 0.001 | 0.909 | 0.000 |
| Prop atts at same time/day as prior wave completion | 0.508 | 0.043 | 0.446 | 0.026 | 0.325 | 0.028 |
| Average time b/n call attempts | 1.038 | 0.000 | 1.023 | 0.035 | 1.031 | 0.044 |

that do not respond in wave three or four and wave three nonrespondents who also do not respond in wave four.

**Table 5 Logistic Regression Attrition**

| | Wave two | | Wave three | |
|---|---|---|---|---|
| | Exp(b) | sig. | Exp(b) | sig. |
| CU tenure (Rents) | 1.667 | 0.000 | 1.662 | 0.000 |
| Married | 0.812 | 0.000 | 0.837 | 0.006 |
| Age of reference person | 0.978 | 0.000 | 0.981 | 0.000 |
| Region (Midwest) | 0.937 | 0.337 | 0.964 | 0.654 |
| Region (South) | 1.275 | 0.000 | 1.246 | 0.003 |
| Region (Northeast) | 1.031 | 0.656 | 1.011 | 0.896 |
| CU size | 0.914 | 0.000 | 0.948 | 0.020 |
| PSU size (Large metro) | 1.262 | 0.012 | --- | --- |
| PSU size (Small metro) | 0.971 | 0.756 | --- | --- |
| Hispanic origin | --- | --- | 0.820 | 0.034 |
| Race of ref (Black) | --- | --- | 1.289 | 0.035 |
| Race of ref (Other) | --- | --- | 0.988 | 0.898 |
| Prop atts at same time/day as prior wave completion | 0.553 | 0.022 | --- | --- |
| Adjusted expenditures | 0.969 | 0.004 | 0.959 | 0.003 |
| Reluctant on prior wave | 0.810 | 0.003 | 1.184 | 0.092 |
| Number of attempts | 1.062 | 0.000 | 1.059 | 0.000 |

*Table 5* shows the results of logistic regression of attrition status on the same list of covariates used in other models. In both wave two and wave three, respondents who rent and respondents from the South are more likely to attrite. Married respondents, older respondents, and respondents from larger CU's are less likely to attrite. The size of the PSU in which the household is located is a significant predictor of attrition at wave two, but not wave three, where race and Hispanic origin are significant predictors at wave three but not wave two. This may indicate that loss of the ability to contact the respondent is somewhat more of a factor in wave two attrition, where refusal to complete the interview is more of a factor in predicting wave three attrition.

This idea is supported by the significance of the proportion of attempts made at the same time of day on the same day of the week as the prior wave's completion for attrition in wave two model, but not the wave three model. The relationship of reluctance is in further support of this theory. Note that reluctance is negatively related to attrition, on wave two similar to the noncontact model, but positively related to attrition on wave three (although not statistically significant). Although the effect size is small, the number of expenditures reported in the prior wave is a significant predictor of attrition, where an increase in data quality, slightly decreases the odds of attrition.

**Discussion**

Prior wave data quality is only a modest predictor of wave nonresponse. The effect of missing race on subsequent nonresponse is rather large, but completely disappears from the more specific models for noncontact and refusal. Missing on other demographic questions (combined) has no effect on nonresponse. Adjusted expenditures, which is practically interchangeable with adjusted outlets in these models has consistent but small effects on nonresponse. These effects are only marginally stronger for refusal versus nonresponse.

Call history variables have greater predictive power. Being hard to reach on the previous wave greatly increases the odds of nonresponse on following waves. This effects is especially large for noncontacts, even controlling for the number of attempts. Being hard to reach has no implications for a refusal on subsequent waves. Being a reluctant respondent greatly increases the odds of refusal and nonresponse on all waves. When examining noncontact however, we find that being reluctant in wave one has a fairly large negative effect in predicting wave two noncontact. Reluctant respondents are much less likely to have a noncontact in wave two. This might indicate that reluctant respondents are not extensively screening their calls as we might expect.

The fit of these logistic models is considerably better for noncontact than refusal and for that reason may be more useful for establishing calling rules and refusal conversion guidelines rather than adjustment. However, while the models may fit relatively poorly they clearly show some strong relationships that merit further investigation. Among these is the relationship between respondent burden versus reluctance due to privacy or other reasons. The somewhat different findings for missingness on race and other data quality variables, might suggest a qualitative difference between refusals due to respondent effort and those originating from privacy or other concerns.

In addition, these findings are certainly strong enough to encourage others to investigate using data quality measures and call history variables in their adjustment schemes in their own surveys.

**References**

Dixon, John and Clyde Tucker. 2002. “Modeling Household and Inteviewer Nonresponse Rates from Household and Regional Characteristics.” Federal Committee on Survey Methodology Working Paper.

Loosveldt, Geert, Jan Pickery, and Jaak Billiet. 2002. “Item Nonresponse as a Predictor of Unit Nonresponse in a Panel Survey.” Journal of Offical Statistics. 18: 545-557.

Rizzo, L, G. Kalton, and Michael Brick. 1994. “Weighting Adjustments for Panel Nonresponse in the SIPP.” *JSM Proceedings*