A STUDY OF PANEL EFFECTS IN RANDOM DIGIT DIALING (RDD) STUDIES

M. Fahimi, D. Maklan, D. Morganstein, and M. Wilson, Westat, Inc.
Mansour Fahimi, Westat, Inc., 1650 Research Boulevard, Rockville, Maryland 20850

Introduction

In repeated surveys designed to measure change, there are a number of appealing reasons for designing subsequent samples that include a subset of units from previous surveys. Aside from cost reductions that result from sampling already identified units, there are also statistical gains to be made from partially overlapping designs; e.g., due to the inherent correlation among the common units, more efficient estimates of change are likely. More detail on this gain is provided in the appendix.

On the negative side, surveys that do include panel components (including the Current Population Survey, National Crime Survey, and Consumer Expenditure Interview Survey) have consistently been haunted by a sizable difference in responses between the overlapping and the cross-sectional units. This phenomena known as *Panel Conditioning* is the reactive effect of prior interviews on current responses (O'Muircheartaigh 1989).

It has been argued that panel conditioning could happen because the baseline interview changes the follow-up responses by providing some kind of retrospection. Mooney (1962) has found that older respondents tend to complain more about illnesses during their first time in the panel than on subsequent times. Ferber (1964) claims that the quality of reporting on savings holdings improved with length of time in the panel. Neter and Waksberg (1964 & 1965) indicate that resident owners report less expenditure on housing repair on the second than on the third time in the panel. Traugott and Katosh (1979) report that longer-term members of a panel give more accurate responses on voting behavior than newer members. Lievesley and Waterton (1985) have found that panel respondents give fewer *don't know* and "*socially desirable*" answers a year later than the fresh cross-sectional respondents. On the other hand, Kalton (1989) indicates that respondents become less willing to repeatedly make the effort to report accurately.

This work presents an examination of this phenomenon for two large RDD surveys, hereafter referred to as the *Attitude Survey* and the *Smoking Survey*. The baselines for both surveys were administered in the Fall 1990, while the respective follow-up studies were conducted about a year later.

1.0 Overview of the Sample Designs

The two surveys examined in this study are both telephone surveys that have employed the so-called Modified Waksberg (1978) method of RDD. Before describing the specifics of the sampling designs for these surveys, a brief account of the standard and the modified Waksberg RDD techniques will be presented.

1.1 The Waksberg Methods of RDD

The standard Waksberg telephone sampling method was designed to reduce the number of nonproductive (e.g., nonresidential) calls by taking advantage of the fact that a high proportion of nonworking and commercial numbers occurs in consecutive sequences. The procedure is a two-stage cluster design, where in the first stage residential clusters of 100 telephone numbers with an identical first eight digits are selected. (A cluster is residential if a randomly dialed number within that cluster turns out to be residential.) In the second stage, a fixed number of telephone households, within each identified residential cluster, are selected for interview.

In contrast, the modified Waksberg procedure calls for a designated number of telephone numbers (instead of households) per cluster. Consequently, it is not necessary to wait until the original sample of clusters has been worked to completion to determine whether the desired number of households within clusters has been achieved. A slight tradeoff is made in the choice of the modified versus the standard Waksberg method. The standard method, while producing a self-weighting sample, is operationally more cumbersome and time consuming. With the modified method, on the other hand, the gain in sampling efficiency and reduction in resource expenditure may compensate for the slight increase in variances compared to the standard method.

1.2 Sample Design for the Attitude Survey

The baseline sample universe for this survey consisted of all young Americans between the ages of 16 and 24. From this population, a sample was drawn to ensure at least 10,000 complete telephone interviews (6,600 males and 3,400 females). These interviews were obtained by calling 17 telephone numbers randomly selected from each of the identified 7,458 residential telephone clusters. The resulting sample yielded complete interviews with 6,398 eligible males and 3,399 eligible females.

For the follow-up administration, the sample consisted of a cross-sectional component, as well as a panel component. The resulting 4,893 cross-sectional respondents were identified through release of 10 additional numbers within each of the existing 7,458 residential telephone clusters. The 5,498 panel respondents were sampled from the pool of 9,797 respondents interviewed during the baseline. Table 1 summarizes the follow-up sample composition.

Table 1. Sample composition for the follow-up of the Attitude Survey

Sample Component	*Size*	*Respondents*
Cross-Sectional (CS)	74,580 Numbers	3,187 Males 1,706 Females
Panel (PNL)	5,555 Males 3,118 Females	3,543 Males 1,955 Females
Total		6,730 Males 3,674 Females

1.3 Sample Design for the Smoking Survey

For this survey too, the general approach employed a two-stage sampling design in which a sample of 32,135 households was initially screened to obtain information about the smoking status of all the household members. At the second stage, all resulting 85,379 screened household members were stratified and then subsampled at designated rates to yield the desired sample sizes. In particular, current smokers and former smokers who had quit in the past five years were sampled at about three times the rate of nonsmokers. The distribution of the resulting 24,296 adult respondents (at least 18 years of age) by sex and smoking status is shown in Table 2.

Table 2. Number of complete interviews with adults in the 32,135 household sample by sex and smoking status

Smoking Status	*Gender*		*Total*
	Male	*Female*	
Nonsmokers	4,062	5,341	9,403
Rest	7,418	7,475	14,893
Total	11,480	12,816	24,296

Similar to the Attitude Survey, the follow-up sample for this survey consisted of two components as well: a cross-sectional sample of 10,774 households from all telephone households, and a panel sample of 9,317 adults selected from the pool of 24,296 adults interviewed during the baseline survey. The following table displays the distribution of the responding adults for each of the two components by gender and smoking status.

Table 3. Responding adults in the cross-sectional and the panel samples of the follow-up survey by gender and smoking status

Smoking Status	*Sample*	*Male*	*Female*	*Total*
Nonsmokers	*CS*	1,545	1,888	3,433
	PNL	844	1,006	1,850
Rest	*CS*	1,926	1,904	3,830
	PNL	1,369	1,423	2,792
Total	*CS*	3,471	3,792	7,263
	PNL	2,213	2,429	4,642

2.0 Results

In order to examine panel effects, it has been necessary to obtain estimates of some key population parameters for both baseline and follow-up surveys for each of the two surveys. Moreover, these estimates had to be obtained and compared across the cross-sectional and panel components. All estimates were based on weighted data, where the weighting procedure involved raking using population estimates obtained from the CPS. The associated standard errors were computed using replications based on a Jackknife technique.

2.1 The Attitude Survey

One of the primary objectives of this survey has been to measure year-to-year change in the youth attitude as measured by the proportion of persons responding *definitely* or *probably* to the questions regarding their attitude.

Table 4. Estimates (and sampling errors) of youth attitude by sample component

Sample component	Baseline survey	Follow-up survey	Difference
CS	16.00 (0.42)	16.74 (0.77)	-0.74
PNL	N/A	12.47 (0.64)	N/A
Overall	16.00 (0.42)	14.48 (0.35)	1.52

Table 4 shows that the attitude estimate obtained from the baseline survey lags that of the cross-sectional component of the follow-up survey by 0.74 = 16.74 - 16.00 percent, which is not statistically significant. On the other hand, when the this comparison is made using the overall estimates, the baseline estimate exceeds that of the follow-up by 1.52 = 16.00 - 14.48 percent. This difference is statistically significant. Clearly, this discrepancy is caused by the statistically significant decline in the attitude estimate resulting from the panel component (4.27 = 16.74 - 12.47 percent). As summarized in Table 5, similar results were found when analogous comparisons were made for various age/gender and race/ethnicity groups. In every cell, the panel estimate from the follow-up survey is less that the overall baseline estimate. The new cross-sectional estimates, on the other hand, are more often higher than those of the baseline.

Table 5. Estimates (and sampling errors) of youth attitude by demographic categories and sample component

Demographic category	*Sample component*	*Baseline survey*	*Follow-up survey*
Males	*CS*	23.55 (0.55)	23.94 (1.19)
	PNL	N/A	17.64 (0.75)
	Overall	23.55 (0.55)	20.59 (0.69)
Females	*CS*	8.99 (0.49)	10.16 (1.01)
	PNL	N/A	7.70 (0.93)
	Overall	8.99 (0.49)	8.86 (0.54)
Males 17 to 21	*CS*	26.77 (0.74)	29.13 (1.72)
	PNL	N/A	20.80 (0.92)
	Overall	26.77 (0.74)	24.72 (0.98)
Males 22 to 24	*CS*	18.01 (0.87)	14.80 (1.39)
	PNL	N/A	12.15 (1.20)
	Overall	18.01 (0.87)	13.39 (0.98)
Females 17 to 21	*CS*	10.20 (0.62)	11.69 (1.21)
	PNL	N/A	9.53 (1.08)
	Overall	10.20 (0.62)	10.52 (0.65)
Females 22 to 24	*CS*	7.12 (0.92)	7.89 (1.57)
	PNL	N/A	4.64 (1.87)
	Overall	7.12 (0.92)	6.23 (1.16)
White	*CS*	13.80 (0.42)	13.00 (0.77)
	PNL	N/A	10.07 (0.52)
	Overall	13.80 (0.42)	11.91(0.46)
Black	*CS*	26.71 (1.63)	29.59 (3.07)
	PNL	N/A	24.93 (3.02)
	Overall	26.71 (1.63)	27.14 (1.96)
Other	*CS*	22.70 (1.55)	25.80 (2.42)
	PNL	N/A	17.88 (2.23)
	Overall	22.70 (1.55)	21.72 (1.69)

2.2 The Smoking Survey

The main statistic of interest for this study was estimates of current smoking prevalence among adults obtained for both baseline and follow-up surveys. As summarized in Table 6, these estimates are compared across the cross-sectional and the panel components for both occasions.

Based on the overall estimates, smoking prevalence among adults has decreased by 0.91 = 22.28 - 21.37 percent, which is not statistically significant. Although there is a larger decline of 2.27 = 22.28 - 20.01 based on the panel estimate, this difference is not statistically significant either. Moreover, comparing estimates of smoking prevalence across various demographic indices reveals similar results. As summarized in Table 7, no uniform panel effect could be detected from these comparisons.

Table 6. Estimates (and sampling errors) of smoking prevalence by sample component

Sample component	*Baseline survey*	*Follow-up survey*	*Difference*
CS	22.28 (0.29)	21.84 (0.63)	0.44
PNL	N/A	20.01 (1.17)	N/A
Overall	22.28 (0.29)	21.37 (0.57)	0.91

Table 7. Estimates (and sampling errors) of smoking prevalence by demographic categories and sample component

Demographic category	*Sample component*	*Baseline survey*	*Follow-up survey*
Males	*CS*	25.33 (0.49)	22.94 (0.78)
	PNL	N/A	22.78 (1.43)
	Overall	25.33 (0.49)	22.90 (0.70)
Females	*CS*	19.29 (0.32)	20.76 (0.81)
	PNL	N/A	17.29 (1.33)
	Overall	19.29 (0.32)	19.87 (0.72)
18 - 24	*CS*	22.41 (0.87)	22.69 (1.56)
	PNL	N/A	21.84 (2.61)
	Overall	22.41 (0.87)	22.47 (1.42)
25 - 44	*CS*	24.37 (0.44)	25.01 (0.87)
	PNL	N/A	19.97 (1.40)
	Overall	24.37 (0.44)	23.71 (0.76)
Over 44	*CS*	19.60 (0.48)	17.52 (0.92)
	PNL	N/A	19.31 (1.79)
	Overall	19.60 (0.48)	17.98 (0.88)
White	*CS*	22.84 (0.46)	23.62 (0.76)
	PNL	N/A	21.96 (1.09)
	Overall	22.84 (0.46)	23.19 (0.62)
Black	*CS*	29.84 (2.15)	23.71 (2.72)
	PNL	N/A	23.98 (4.93)
	Overall	29.84 (2.15)	23.78 (2.40)
Hispanic	*CS*	19.45 (1.14)	18.19 (1.33)
	PNL	N/A	13.84 (1.95)
	Overall	19.45 (1.14)	17.07 (1.02)
Other	*CS*	20.32 (1.65)	17.70 (1.75)
	PNL	N/A	19.37 (4.90)
	Overall	20.32 (1.65)	18.13 (1.71)

3.0 Concluding Remarks

As stated in the beginning, there are a number of compelling reasons for designing repeated surveys that do include panel components -- cost savings and reducing sampling errors for estimating change due to the existing correlation among common units, for example. Despite these tangible rewards, however, survey designers must consider the possible effect of the panel conditioning. As seen above, this effect can become more unpredictable when overall estimates are obtained by blending the results from the panel and cross-sectional components.

The Attitude survey indicates that there are some latent differences among the responses obtained from the cross-sectional and panel respondents. According to this survey, the estimate of attitude is significantly lower among the panel than the cross-sectional respondents. The Smoking Survey, on the other hand, seems to indicate that panel respondents do not yield a significantly different estimate of smoking prevalence.

Additional research is planned for these two surveys. First, it would be valuable to compare the follow-up responses of the panel members with their respective baseline responses. By eliminating potential noise in the overall baseline estimates, this comparison should reveal a more efficient estimate of change. Second, using logistic regression techniques, more scrupulous comparisons could be made among various subgroups with respect to the panel effect.

It has been suggested that in health surveys, some respondents exaggerate their illnesses to a lesser degree the longer they remain in the panel. With attitudinal surveys, panel respondents are thought to provide more specific answers. Then again, panel responses can be contaminated with confounding effects such as *Recall, Changes in the Population,* and *Aging* (Holt, 1989). In light of these concerns, cost savings and statistical gains should be judged carefully against the possibility of encountering panel conditioning.

Appendix (Estimating Change)

To measure change in some statistic obtained from two partially overlapping surveys, the pool of respondents from the first survey (baseline) and the subsequent survey (follow-up) can be categorized as follows.

G_1: Survey 1 specific respondents, those baseline respondents who will not be followed up;

G_2: PNL respondents, consisting of those respondents to be followed from the pool of baseline respondents;

G_3: Survey 2 specific respondents, consisting of the new cross-sectional respondents.

In light of the above, an estimate of change in a proportion P, say, would be of the following form.

$$\Delta_P = \omega \times \left(P_{G_2}^{\text{follow-up}} - P_{G_2}^{\text{baseline}}\right) + (1 - \omega) \times \left(P_{G_1} - P_{G_3}\right)$$

where $P_{G_2}^{\text{baseline}}$ and $P_{G2}^{\text{follow-up}}$ represent estimates of the desired proportion based on the paneled group G_2 obtained from the baseline and the follow-up surveys, respectively, and P_{G_1} and P_{G_2} represent estimates of P based on the nonpaneled groups G_1 and G_3. An optimal value of ω should then be estimated as a function of the variance/covariance of P obtained via the baseline and follow-up results. The following derivation illustrates the dependence of the optimal value of ω on the variance and covariance components.

$$\text{Let: } \begin{cases} \alpha = \text{Var}\left(P_{G_2}^{\text{follow-up}} - P_{G_2}^{\text{baseline}}\right), \\ \beta = \text{Var}\left(P_{G_1} - P_{G_3}\right), \text{ and} \\ \gamma = \text{Cov}\left[\left(P_G^{\text{follow-up}} - P_G^{\text{baseline}}\right), \left(P_G - P_G\right)\right] \cong 0, \end{cases}$$

Then:

$$\begin{aligned} \text{Var}\left(\Delta_P\right) &= \omega^2\alpha + (1 - \omega)^2\beta + 2\omega(1 - \omega)\gamma \\ &= (\alpha + \beta - 2\gamma)\omega^2 + (-2\beta + 2\gamma)\omega + \beta. \end{aligned}$$

Consequently:

$$\text{Min}\left\{\text{Var}\left(\Delta_P\right)\right\} \Leftrightarrow \omega = \frac{-(-2\beta + 2\gamma)}{2(\alpha + \beta - 2\gamma)}.$$

But:

$$\begin{aligned} \omega &= \frac{\beta - \gamma}{\alpha + \beta - 2\gamma} \\ &= \frac{\beta}{\alpha + \beta - 2\gamma} - \frac{\gamma}{\alpha + \beta - 2\gamma} \rightarrow \frac{\beta}{\alpha + \beta} \text{ as } \gamma \rightarrow 0. \end{aligned}$$

Hence:

$$\omega_{\text{Optimal}} = \frac{\text{Var}\left(P_{G_1} - P_{G_3}\right)}{\text{Var}\left(P_{G_2}^{\text{follow-up}} - P_{G_2}^{\text{baseline}}\right) + \text{Var}\left(P_{G_1} - P_{G_3}\right)}.$$

References

Ferber, R. (1964), *Does a Panel Operation Increase the Reliability of Survey Data: The Case of Consumer Savings*, Proceedings of the Social Statistics Section, American Statistical Association, 210-216.

Holt D. (1989), *Panel Conditioning: Discussion*, Panel Surveys, Edited by D. Kasprzyk, G. Duncan, G. Kalton, and M.P. Singh, Wiley, New York.

Kalton, G., Kasprzyk, D., & McMillen, D.B. (1989), *Nonsampling Errors in Panel Surveys*, Panel Surveys, Edited by D. Kasprzyk, G. Duncan, G. Kalton, and M.P. Singh, Wiley, New York.

Lievesley, D., & Waterton, J. (1985), *Measuring Individual Attitude Change in British Social Attitude: The 1985 Report*, Gower Publishing Company, Aldershot, England, 177-194.

Mooney, H.W. (1962), *Methodology in two California Health Surveys (Public Health Monograph No. 70)*, U.S. Government Printing Office, Washington, DC.

Neter, J., & Waksberg, J. (1964), *A Study of Response Errors in Expenditures Data from Household Surveys*, Journal of the American Statistical Association, 59, 18-55.

Neter, J., & Waksberg, J (1965), *Response Errors in Collection of Expenditures Data by Household Interviews: An Expenditure Study*, Bureau of the Census Technical Report No. 11, U.S. Government Printing Office, Washington, DC.

O'Muircheartaigh C. (1989), *Sources of Nonsampling Errors: Discussion*, Panel Surveys, Edited by D. Kasprzyk, G. Duncan, G. Kalton, and M.P. Singh, Wiley, New York.

Traugott, M.W., & Katosh, J.P. (1979), *Response Validity in Surveys of Voting Behavior*, Public Quarterly, 43, 359-377.

Waksberg, J. (1978), *Sampling Methods for Random Digit Dialing*, Journal of American Statistical Association, 73, 40-46.