

ARTICLES

Benefits—and Costs—of a Multi-Mode Survey of Recent College Graduates

David Crow¹, Martin Johnson², Robert Hanneman²

¹ Centro de Investigación y Docencia Económicas (CIDE, Mexico City), ² University of California, Riverside

Keywords: survey practice

<https://doi.org/10.29115/SP-2011-0024>

Survey Practice

Vol. 4, Issue 5, 2011

Benefits—and Costs—of a Multi-Mode Survey of Recent College Graduates

Multi-mode surveys can enhance the representativeness of respondents by reducing coverage and non-response errors, and by increasing response rates and the overall number of respondents (Dillman, Smyth, and Christian 2009, 302–4; Weisberg 2005, 278–96)—particularly since the causes of unit non-response and, consequently, respondents' profiles vary across modes (Groves et al. 2002). Multi-mode surveys are increasingly the standard practice among researchers (Biemer and Lyberg 2003). However, they impose additional costs and administrative burdens (Dillman, Smyth, and Christian 2009, 301). But how much does the use of multiple modes of contact improve the representativeness of respondents relative to the sample? Does this improvement, if any, come at an acceptable cost? We analyze the effects of using three modes of contact sequentially in a survey of recent graduates of University of California, Riverside, and assess the fieldwork costs associated with the use of multiple contact modes.

While the use of multiple contact modes improved the gender and ethnic representativeness of the respondents, the modes varied in their cost effectiveness. In this sample of recent college graduates, e-mail proved to be the most cost effective mode of contact, representing eight percent of our fieldwork budget while improving the representativeness of the core telephone portion of the sample and generating about one-third of our completed survey responses.

Sampling Frame and Survey Administration

The UC Riverside Survey Research Center (SRC) conducted a survey of alumni for the university's Institutional Research Coordinating Group May through August of 2009. The target population comprised all 2,469 students awarded bachelor's degrees in 2002–2003. Contact information provided by the university Registrar included telephone numbers for 947 graduates, e-mail addresses for 1,495, and mailing addresses for 2,232. There were significant coverage gaps. For example, we had both telephone numbers and e-mails for only 617 graduates. We had only mailing addresses for 528 graduates, and no

information at all for 115. We refer to the 2,354 graduates for whom we had any contact information as the *sample*, albeit a nonprobability one, covering 95% of the population of these recent graduates.

The SRC administered the multi-mode survey sequentially. We first called all graduates for whom we had telephone numbers, averaging 3.3 attempts per sample record, during three weeks in May, 2009. After administering the survey by telephone, we sent up to four e-mail invitations to everyone for whom we had e-mail addresses (provided they had not already responded by telephone) for six weeks between June and August, 2009. Simultaneous to the telephone and Web portions of the survey, we mailed up to two self-answered questionnaires (SAQs) to everyone for whom we had a mailing address but neither a telephone number nor an e-mail address, on May 26 and June 22, 2009. Consequently, respondents we contacted by different modes do not represent independent samples, limiting our inferences about the differences among them. However, we are able to examine how this practical effort to expand our sample affected its overall representativeness as we added contact modes and data. We cautiously make limited comparisons across the subsamples by mode of contact. The registrar-supplied file also included auxiliary variables (sex, ethnicity, citizenship, college, and GPA) that allow for comparison of respondents to non-respondents on known demographic quantities.

Response Rates and Sample Characteristics by Contact Mode

Our efforts netted 330 interviews, 125 by telephone, 111 by e-mail, and 94 by mail. The overall response rate was 14.0% (330 respondents divided by 2,354 sample units). This overall response rate belies dramatically different response rates by mode of attempted contact. Table 1 divides the sample into four groups by the type of contact we attempted: 1.) sample members we only tried to call by telephone (Column 2), 2.) those we only attempted to contact by e-mail (Column 3), 3.) those to whom we sent e-mail invitations after trying to reach by telephone (Column 4), and 4.) those to whom we only mailed invitations by post. The subsamples vary somewhat in their demographic characteristics. For example, the persons in our sample for whom we only had e-mail contact information contained a higher proportion of women (59.2%) than the people for whom we only had physical addresses (51.3%). In the overall sample, 56.4% were female (Column 5). These differences were significant (ANOVA F-value of 3.1, $p=.028$). The last line of the table summarizes the different response rates across these modes of attempted contact: 29.7% for telephone only, 8.1% for e-mail only, 7.9% for e-mail after telephone, and 17.2% for mail only (differences significant at $F = 47.0$, $p < .001$).

Table 1 also demonstrates that the composition of the sample varies significantly by attempted contact mode for the characteristic of ethnicity. Specifically, proportions of Asian Americans and Whites in the sample also

Table 1 Selected sample characteristics by mode of attempted contact and response rates.

Variable	Phone Only (N=425)	E-mail Only (N=879)	Phone, then E-mail (N=522)	Mail Only (N=528)	Sample (N=2,354)	F (df=3)	p
Female	58.4%	59.2%	55.4%	51.3%	56.4%	3.1	.028
Ethnicity							
Asian Am.	33.9%	40.4%	44.3%	33.0%	38.4%	6.5	.000
Latino	21.4%	22.9%	18.6%	29.7%	26.0%	1.2	.306
White	29.7%	23.8%	23.0%	29.7%	26.0%	3.8	.009
Response Rate	29.7%	8.1%	7.9%	17.2%	14.0%	47.0	.000

differed significantly by mode of attempted contact. Analysis of the remaining registrar-supplied background variables (not shown) revealed that the distribution of enrollment in UCR's undergraduate colleges differed by mode of attempted contact, but within-degree GPA and citizenship status did *not*. The differences we observe among contact mode groups—known *ex ante*—justify pursuing a multi-mode design, rather than jettisoning any particular contact mode.

Response Modes and Sample Representativeness

Successive addition of contact modes affords mixed evidence of increased representativeness, moving the aggregate demographic characteristics of respondents closer to the characteristics of non-respondents in some cases, but not others. Table 2 shows that for the variable of sex, 63.2% of telephone respondents were female, compared to 56.5% of non-respondents (though given the small sample size, the difference is insignificant, $p = .149$). E-mail responses brought the proportion of females down to 58.9%, and mail, to 56.1%. Generally, multiple modes made the sample more representative of the population for Asian Americans (though, as the p -values show, in no case did respondents approach non-respondents) and Latinos, though not for Whites. Analysis of other background variables in the registrar-supplied file (not shown) revealed modest, but statistically insignificant, representativeness gains for GPA. For citizenship status and college affiliation, adding modes actually reduced the representativeness of respondent demographic characteristics, compared to non-respondents.

Mode Effects on Substantive Variables

Furthermore, there were significant differences across contact modes on substantive questions included on the survey. The survey's sponsors were interested in graduates' employment rate and evaluations of their UC Riverside experience, as measured by agreement (on a six-point scale) with the statement, "Knowing what I know now, I would still choose to enroll at UCR." As the ANOVA F- and p-values in Table 3 show, respondents varied significantly on these characteristics according to contact mode. Across the contact mode groups, the percentage of graduates employed full or part-time ranges from

Table 2 Respondent representativeness by contact mode for selected characteristics.

Variable	Non-Respondents (N=2,024)	Phone (N=125)	t	P	Phone+ E-mail (N=236)	t	P	Phone+ E-mail+Mail (N=330)	t	p
Female	56.5%	63.2%	-1.44	.149	58.9%	-0.67	.498	56.1%	0.16	.874
Ethnicity										
Asian Am.	41.1%	20.8%	4.49	.000	25.0%	4.74	.000	23.9%	5.85	.000
Latino	21.7%	25.6%	-1.00	.317	22.0%	-0.10	.917	20.6%	0.44	.654
White	23.2%	39.2%	-3.98	.000	38.1%	-4.88	.000	40.3%	-6.38	.000

Table 3 Selected survey variables by contact mode.

Variable	Phone Only (N=126)	E-mail Only (N=71)	Phone, then E-mail (N=41)	Mail Only (N=92)	Total (N=330)	F (df=3)	P
Employment	71.4%	81.7%	90.2%	87.0%	80.3%	3.9	.009
Would Choose UCR Again (1-6)	5.09	4.59	4.76	4.47	4.77	4.5	.004

Table 4 Selected survey variables by contact mode, weighted by race and gender.

Variable	Phone Only (N=126)	E-mail Only (N=71)	Phone, then E-mail (N=41)	Mail Only (N=92)	Total (N=330)	F (df=3)	P
Employment	73.2%	79.4%	89.3%	85.0%	79.9%	2.5	.061
Would Choose UCR Again (1-6)	5.04	4.51	4.82	4.42	4.73	4.6	.004

71.4% (telephone only) to 90.2% (telephone, then e-mail), with an overall mean of 80.3% (F-value of 3.9, $p = .009$). Mean satisfaction ratings of UCR also vary, with a minimum of 4.47 (mail only), a maximum of 5.09 (telephone only), and overall mean of 4.77 (F-value of 4.5, $p = .004$). These mode effects militate, *post hoc*, in favor of a multi-mode design.

We also compare aggregate responses on these substantive items by contact mode, weighting in group of respondents to reflect the demographic composition of the sample, using post-stratification weights defined by gender and ethnicity, the characteristics explored in Table 2. The weighted contact mode group data is reported in Table 4. The results are similar to those in Table 3. The weighted percentage of graduates in the groups employed full or part-time ranges from 73.2% (telephone only) to 89.3% (telephone, then e-mail), with an overall mean of 79.9% (F-value of 2.5, $p = .061$). Satisfaction with UCR also varies, ranging from 4.42 (mail only) to 5.04 (telephone only), and overall mean of 4.73 (F-value of 4.6, $p = .004$).

Comparing these tables provides another perspective on the quality of the respondent pool that resulted from the multi-mode strategy. When the pool of respondents differs markedly from known demographics of the sample, analysts are tempted to post-weight the data. Because the multi-mode strategy was able to produce demographics (sex and ethnicity) among respondents that were fairly close to the sample, difference on key variables between raw data (Table 3) and post-weighted data (Table 4) are not striking.

Cost

Was the improved sample coverage and added precision worth the extra cost incurred by multiple modes? This is necessarily a subjective judgment. Variable costs—that is, hourly wages associated with contacting and interviewing that increase according to the number of attempted contacts—totaled about \$8,963 for the telephone portion of the survey (\$9.47 per sample unit we attempted to contact by telephone and \$71.78 per completed telephone interview), \$1,088 for the e-mail portion (\$0.78 per sample unit we attempted to contact by e-mail, \$9.81 per completed e-mail interview), and \$3,272 for the mail portion (\$6.19 per sample unit we attempted to contact by post, \$34.80 per completed mail interview). Telephone contacts represented 67 percent of the fieldwork budget, mail 25 percent, and email 8 percent.

In some cases, the added expense associated with the use of the additional contact modes bought important gains in representativeness. For example, with regard to gender, each extra \$1,000 spent on e-mail brought respondents nearly four percentage points closer to non-respondents and each extra \$1,000 spent on mail, 0.61 percentage points closer. For several variables, however, there was a *negative* return on investment: Survey respondents became less representative of non-respondents, and thus the sample, on citizenship and academic college as we included additional response modes. The value of multiple modes lay in increased precision: the margin of error went from $\hat{A} \pm 8.8\%$ for telephone to $\hat{A} \pm 6.4\%$ with e-mail, to $\hat{A} \pm 5.4\%$ with mail. Thus, each extra \$1,000 redounded in a reduction of 0.78 percentage points in the margin of error.

Conclusion

This case study of the 2009 UC-Riverside Alumni Survey suggests the pursuit of multiple contact modes is a mixed bag in terms of implications for sample representativeness and cost. Survey researchers may find appealing the use of all available tools to recruit participants into a study. In our case, the use of these multiple modes simultaneously improved and weakened the demographic representativeness of our respondents. Increases in sample precision mitigate the rather heterogenous results in terms of the demographic representativeness of respondents. Of course, our inferences are limited by the nature of the underlying population (young college graduates), the non-random assignment of study participants to mode of contact, and our small sample size.

Did the costs justify the fieldwork strategy? As a practical matter, we think so. The largest portion of our respondents came from the subsample for whom we had telephone numbers. This mode of contact was essential. The use of e-mail in this sample of young college graduates was also particularly valuable, as both the least expensive and most efficient means of contact, supplementing the core telephone mode of the study. While mail was a less efficient means of data collection, costlier and less effective than e-mail contact, we are hesitant to suggest that we or other researchers disregard potential mail respondents given their importance to our overall study, yielding 28 percent of our respondents. Further, the significant differences we observe in the characteristics of respondents across modes offer strong support for pursuing this multi-mode design.

REFERENCES

- Biemer, P.P., and L.E. Lyberg. 2003. *Introduction to Survey Quality*. Wiley Series in Survey Methodology. Hoboken, NJ: John Wiley and Sons.
- Dillman, D.A., J.D. Smyth, and L.M. Christian. 2009. *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. 3rd ed. Hoboken, NJ: John Wiley and Sons.
- Groves, R.M., D.A. Dillman, J.L. Eltinge, and R.J.A. Little. 2002. "Survey Nonresponse in Design, Data Collection, and Analysis." In *Survey Nonresponse*, edited by R.M. Groves, D.A. Dillman, J.L. Eltinge, and R.J.A. Little. New York: John Wiley and Sons.
- Weisberg, H.F. 2005. *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. Chicago: The University of Chicago Press.