

Surveying Rare Populations Using a Probability-based Online Panel

Graham Wright
Brandeis University

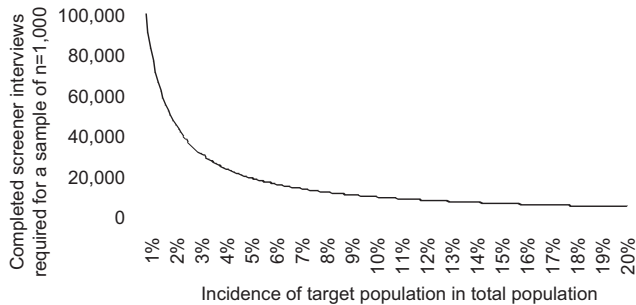
Jordon Peugh
GfK Custom Research

Obtaining adequate and representative samples of rare populations, such as certain religious or ethnic minorities, is a perennial challenge in survey research and is complicated by high costs and a strong potential for bias (Tighe et al. 2010). In the case of American Jews, a probability-based online panel can provide a representative sample for a reasonable cost. However, a key issue for surveys of rare populations is that baseline data about the size and characteristics of these populations often do not exist, making it difficult to assess the level of bias in any given sample and adjust for it with weighting. This paper illustrates this difficulty by comparing the results from a probability-based online panel to other national surveys of American Jews. In each case, methodological issues prevent a definitive assessment of bias.

Since the late 1970s, the standard for achieving a representative sample of respondents has been to use Random Digit Dialing (RDD) telephone methods, which have evolved to adapt to a changing technology environment (Tucker and Lepkowski 2008). However, surveys of rare populations such as Jews must first identify such individuals through screening interviews. Figure 1 shows the number of interviews required to achieve a sample of 1,000 respondents of a rare population by incidence of the population. The elbow of the curve is around 4–5 percent, beyond which the number of interviews (and thus the cost) increases exponentially. The U.S. Census does not ask questions about religious preference, so the exact proportion of Jews in the United States is questionable although most researchers put the figure at less than 2 percent (Saxe et al. in press). With such a low incidence, relying on a pure RDD sample is infeasible, particularly if one wants to ensure an adequate response rate. As a result of these issues, previous efforts to use RDD-based methods to achieve a national sample of American Jews have been both extremely costly and plagued by concerns about sampling bias and methodological errors.

To identify a more efficient alternative, Brandeis University's Cohen Center for Modern Jewish Studies (CMJS) collaborated with Knowledge Networks

Figure 1 Screener interviews required for a sample of n=1,000.



(KN)¹, to use their nationally representative online panel, KnowledgePanel to administer surveys to a sample of American Jews in 2010 and 2011. This approach potentially has powerful implications for the study of American Jewry and other rare populations.

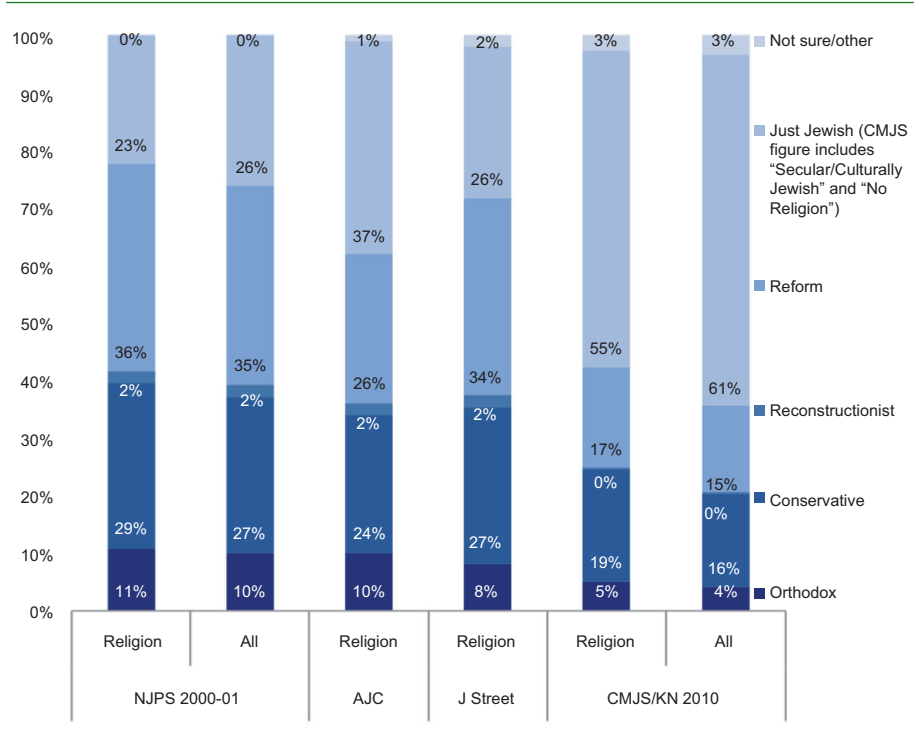
KnowledgePanel is an online non-volunteer panel, in which potential panel members are chosen via a statistically valid address-based sampling (ABS) method. ABS involves probability-based sampling of addresses from the U.S. Postal Service's Computerized Delivery Sequence File, which covers 97 percent of the residential addresses in the U.S. population. Recruited households without Internet access are provided a computer and free Internet service in exchange for participation. KnowledgePanel consists of about 50,000 U.S. residents age 18 and older, including cell phone-only and Spanish-speaking households. Three percent of the adult members of KnowledgePanel identify as Jewish by religion or background.

The use of such a panel to understand a rare population begs the question of how representative the panel is of the population as a whole. However, the lack of authoritative population data needed for such a comparison is what suggests the use of such a panel in the first place. This is the case for the American Jewish population. To illustrate this problem, we compare the CMJS/KN results to three prior surveys of American Jews: the 2000-01 National Jewish Population Study (NJPS), an RDD-based survey that was the subject of much methodological criticism due to its apparent bias; the 2001 Survey of American Jewish Opinions conducted by the American Jewish Committee using an opt-in volunteer sample (AJC); and a 2009 survey conducted for J Street, an Israel advocacy group, by Gerstein/Agne Strategic Communications also using an opt-in sample (J Street). The two opt-in surveys appear to be unweighted and provide no methodological assurance of their representativeness.

For this paper, we highlight differences in denominational affiliation across the surveys. Figure 2 shows estimates of the proportion of Jews (either Jews by religion or all Jews, including those who are Jewish by background but not by religion) who identify with each of the major Jewish denominations for the

¹ GfK Custom Research.

Figure 2 Jewish denominational affiliation by survey (for Jews by religion and all Jews).



2010 CMJS/KN survey, as well as the three other national surveys.

The most striking discrepancy is the higher proportion of Jews without any particular denominational affiliation in the CMJS/KN survey compared to the other surveys. A number of factors could have produced this discrepancy, which we will examine in turn: differences in question wording between the surveys, actual changes in Jewish denominational affiliation during the past 10 years (in the case of comparisons between CMJS/KN estimates and NJPS and AJC estimates), bias in the estimates of the three other surveys, and bias in the CMJS/KN panel.

Question wording has been shown to have significant effects on reported Jewish denomination. The standard CMJS denomination question prompts respondents with options of “Secular/Culturally Jewish,” “Just Jewish,” and “No Religion” in addition to the traditional denomination options. This is in contrast with the three comparison studies, which only include a category for “Just Jewish.” CMJS conducted an experiment using the KN panel members, asking the same sample about denomination in both formats a few months apart. When additional prompts for “Secular/Culturally Jewish” and “No Religion” were excluded, the proportion not affiliated with any denomination dropped from 62 percent to 48 percent, implying that question wording can have a significant effect on choice of denomination, at least among those of low denominational affiliation. This effect likely contributes to the discrepancy

recorded above.

Lower levels of affiliation in the CMJS/KN study compared to the AJC and NJPS 2001 estimates might also be due to respondents actually changing denomination in the 10 years between the surveys. To assess this, Jewish KN panel members were asked what denomination they identified with currently and 10 years previous. Table 1 shows that all three of the mainstream denominations (Orthodox, Conservative and Reform) have lost significant numbers of members in the previous 10 years, mainly to the “Just Jewish” category.

Thus, the discrepancy between the 2001 era NJPS and AJC estimates and the 2010 CMJS/KN estimates seem, in part, due to actual increases in the proportion of unaffiliated Jews in the intervening years. If this is accurate, then the similarity between the 2009 J Street estimates and earlier NJPS figures is cause for skepticism rather than confidence in the J Street figures.

Even after accounting for the above trends, it is likely that the lower rates of affiliation in the CMJS/KN study are, in part, due to methodological differences between the studies or bias in their respective samples. The NJPS 2001 found almost double the proportion of adult Orthodox respondents compared to a comparable study done in 1990. Neither rates of denominational switching, immigration, nor age cohort effects seemed able to explain such growth (Saxe et al. 2007). This suggests that the low response rate of the survey led to bias due to Orthodox respondents being more likely to respond than the non-Orthodox.

The 2001 AJC and 2009 J Street surveys are derived from Jewish subsamples of opt-in market research panels assembled by Market Facts Inc (now Synovate) and YouGovPolymetrix respectively. Because these samples are non-random (they are made up only of individuals who volunteer to be included), standard statistical tests cannot be used with them; thus, results derived from them cannot be said to be reflective of the actual population with any degree of confidence. In addition, neither study reported information on weighting, response rate, or other efforts to alleviate bias. Thus, it is distinctly possible that bias in the panels is the cause of the discrepancy with CMJS/KN data (Chang and Krosnick 2009; Paeski and Krosnick 2010; Yaeger et al. 2009).

Table 1 Adult denominational change from 2000 to 2010.

Denomination in 2010	Denomination in 2000 (recollection)				
	Orthodox	Conservative	Reconstructionist	Reform	Just Jewish
Orthodox	68%	1%	0%	1%	2%
Conservative	2%	71%	0%	4%	2%
Reconstructionist	0%	2%	38%	1%	0%
Reform	5%	15%	15%	70%	2%
Just Jewish	25%	11%	47%	24%	95%
Total	100%	100%	100%	100%	100%
	<i>n</i> =33	<i>n</i> =246	<i>n</i> =12	<i>n</i> =296	<i>n</i> =338

The final potential explanation for the disparate estimates is bias in the KN Jewish sample itself. While KN uses probability-based sampling and makes substantial efforts to ensure the representativeness and randomness of KnowledgePanel, it is possible, especially for small, distinctive populations such as Jews, that a sub-group may be over- or underestimated. Specifically, the small proportion of Orthodox identified in the panel (around 5 percent for Jews by religion) is much lower than in the other studies. With no definitive national benchmark, however, this discrepancy cannot be quantified with any specific degree of certainty.

Our conclusion is that an online, probability-based panel can be a cost-effective source for achieving a representative sample for a rare population. It is likely that estimates that result cannot definitively be proven to be unbiased. However, the extreme difficulty in surveying such populations means that these estimates will likely still be among the most accurate available.

References

- Chang, L.C. and J.A. Krosnick. 2009. National surveys via RDD telephone interviewing vs. the internet: comparing sample representativeness and response quality. *Public Opinion Quarterly* 73(4): 641–678.
- Pasek, J., and J.A. Krosnick. 2010. Measuring intent to participate and participation in the 2010 census and their correlates and trends: comparisons of RDD telephone and non-probability sample internet survey data. Statistical Research Division Study Series: Survey Methodology, (#2010–15). U.S. Census Bureau, Washington. Available at <http://www.census.gov/srd/papers/pdf/ssm2010-15.pdf>.
- Saxe, L., E. Tighe, B. Phillips and C. Kadushin. 2007. Reconsidering the size and characteristics of the American Jewish population. Cohen Center for Modern Jewish Studies, Brandeis University, Waltham.
- Saxe, L., E. Tighe and M. Boxer. In press. Measuring the size and characteristics of American Jewry: a new paradigm to understand an ancient people. In: (U. Rebhun, ed.) *Studies in contemporary Jewry* (Vol. XVIII). Oxford University Press, New York, Oxford.
- Tighe, E., D. Livert, M. Barnett and L. Saxe. 2010. Cross-survey analysis to estimate low-incidence religious groups. *Sociological Methods & Research* 39: 56–82.
- Tucker, C. and J.M. Lepkowski. 2008. Telephone survey methods: adapting to change. In: (J.M. Lepkowski, C.M. Tucker, J.M. Brick, E.D. de Leeuw, L. Japiec, P.J. Lavarakas, M.W. Link and R.L. Sangster, eds.) *Advances in telephone survey research methodology*. John Wiley & Sons, Inc., New Jersey.
- Yeager, D.S., J.A. Krosnick, C. LinChiat, H.S. Javitz, M.S. Levendusky, A. Simpser and R. Wang. 2011. Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. *Public Opinion Quarterly*.