

RE: Reviews of Cameron and DeShazo “Comprehensive selectivity assessment for a major consumer panel: attitudes toward government regulation of environment, health and safety risks” submitted to SSCORE

REVIEW #1

The intended contribution of the paper is to determine "whether representativeness of [Knowledge Networks] is adequately maintained [...] - so that models based on only an estimating sample can be safely assumed to produce inferences that can be considered valid for the entire US population." (p. 5)

The idea is very interesting and the article is sound and without major flaws. Yet, I have one major concern that relates to the contribution of the paper:

If we assume that RDD produces random and representative samples, why should an online panel like KN - which recruits respondents through RDD and provides panelists without computers or Internet access with Web-TV hardware and Internet access- not be representative?

The list of households that KN *attempts* to recruit will indeed be very close to representative. However, as our paper explores, not all recruitment attempts are successful (only 35% of the initial RDD contacts agree to participate). There is more attrition when it comes time to collect “profile” data from these individuals (only 18% of the initial RDD contacts actually get “profiled”). Even among those who get through the profiling stage, there is a drop-out rate over time (i.e. only 7.3% of the initial RDD contacts were still in the panel at the time we “drew” our intended sample).

Our actual estimating sample is a mere half of one percent of the initial over-half-million RDD contacts. Even if the initially RDD contact list was perfectly representative, everything that happens between those initial contacts and the remaining pool at the time our sample is drawn could produce a highly non-representative sample. Most people assume representativeness of the current panel members (in our sample, this would be the 7.3% of surviving RDD contacts at the time our intended sample is drawn) and only model, at best, the attrition between a sample drawn from the current panel and actual survey respondents. KN has come under fire from the OMB in the last couple of years for downplaying the potentially non-representative nature of the active panel, even though the initial recruiting contacts are as random as they come.

The authors do not provide any theoretical reason for this fundamental question. I believe that this is an issue of major importance. If they would have examined Harris or any other panel with different, non-random recruitment methods, their investigation would make sense to me. But in case of KN, it is neither logical nor well supported in the article. Also, to my knowledge, RDD recruitment is the exception rather than the norm for online panels worldwide. Therefore, the findings are limited in the sense that they do not allow generalization to most

other panels with different recruitment methods. Given the lack of theoretical underpinning, the paper seems like a methodology exercise.

It is true that KN is the current “gold standard” in terms of panel composition. However, given the concerns about the processes driving attrition from the KN panel, we sought to examine that attrition phase. The validity of our own KN panel research depends upon the representativeness of the active KN panel, not just its recruitment process, so we were highly motivated to get to the bottom of this.

It perplexes me that this referee feels we need a “theoretical” reason why the KN panel may not be representative. We had hoped that the attrition data in Table 1 would make that very clear, but this referee’s confusion certainly points to our need to be much more careful to emphasize this point in the paper. We can do this.

The implication section of the article needs more elaboration. Can KN really be used for predicting election results or policy making? On page 4 you mention that a “sample that mimics the population in terms of the marginal distributions of a few observable variables may still be non-representative if the sample and the population differ in terms of unmeasured or unobserved characteristics”. You address this point by matching panel information with census tract factors and presidential voting patterns. Yet, panel members could still be more opinionated, motivated, interested in surveys, or differ on other unobserved characteristics which could affect results.

That’s the whole point of Heckman-type sample selection correction measures—to allow the unobserved characteristics to be explicitly correlated, so that they do not bias the implications of the estimated “outcome” equation. In the process of modeling the selection process, however, one needs to control for as many observable factors as possible in modeling the selection process itself. That is where we need to match the half-million RDD contacts with census tract factors and presidential voting patterns. These are the “observable” variables upon which we rely, in modeling the selection process.

This comment suggests that the referee believes we are merely comparing census tract factors and voting patterns for our sample of 2911 actual respondents to the same variables in the population. This is the sort of thing that has traditionally been done to assess representativeness, but it is not what we are doing.

To reiterate a discussion of the type that we can feature more prominently in the paper:

Typical assessments of sample representativeness focus only on “observables,” namely variables that can be measured for both the population and the sample (such as age, gender, income). The subtlety in formal sample selection models is that even if the sample “looks like the population” on these measured attributes, it could differ according to unmeasured (unobservable) attributes, such as “avidity,” or “concern about the topic” and so on. Modern selection-correction methods pay close attention to apparent correlations in the unobservables (the

error terms) between the process that leads to sample membership and the magnitudes of the variable that is being studied in the sample. For instance, if liberals are more likely to be willing to spend their valuable time completing a survey, then the opinions expressed in the survey will have a “liberal bias.” Concern has been apparently expressed that the KN panel has a “liberal bias” because some conservatives have been unhappy with the results of some KN surveys. We were curious about whether there was any evidence of such a bias, in spite of the apparent representativeness of the panel in terms of observable variables.

Evidence indicates that on average, online panels face an attrition rate of 13%. In how far would that affect the representativeness of KN over time?

I'm not sure how this referee is counting his or her measure of “attrition”. As our Table 1 indicates, attrition in the KN panel amounts to 93% by the time the active panel is sampled for our survey. It matters *greatly* what is taken as the starting point in measuring attrition. For KN, that starting point is the random (RDD) contact list. For other online panels, perhaps the implicit starting point is merely those people who originally enroll in the panel, which might correspond to the KN “profiled” group, where 82% of the original contacts have already dropped out of the “random” group. A small attrition rate for internet panels which do not recruit by RDD is pretty meaningless. We can make this comparison more prominent in the paper.

REVIEW #2

This manuscript models the selection process for a major consumer panel (Knowledge Networks, Inc.). The authors evaluate all possible instances at which selection bias could occur, starting with a population of over 500,000 random-digit-dialed contacts and ending with a final sample of 2,911 survey respondents. Then, the effect of the selection process on respondents' answers to a question about the role of government in environmental, health and safety regulation is assessed using two distinct methods. This is a clearly written, well-organized manuscript and has particular salience for researchers and policy-makers who access survey data from Knowledge Networks, Inc.

Reviewer's Recommendation-Accept with minor revisions.

Suggestions to Enhance the Manuscript:

Introduction

The first paragraph seems unnecessary and detracts from the real focus of the paper. Why introduce online survey methods when the focus of the paper is on selection bias and traditional telephone surveys (i.e., RDD)? The manuscript could begin with the fourth paragraph (on page 4) that begins with the discussion of the OMB's standards for survey research. Then, KN, Inc could be introduced.

We missed the boat with this referee, on this point. The Knowledge Networks panel IS an online survey method. The focus of the paper is indeed on selection bias, but NOT about traditional telephone surveys. The RDD element of the story is that the members of the online survey panel are initially *recruited* via RDD telephone contacts, but they are not *surveyed* by telephone. So this referee is mistaken (undoubtedly our fault). We do need to discuss online survey methods, since KN's panel is an online panel. It is just their recruiting method that is light-years ahead of other online surveys when it comes to potential representativeness. We can make a greater effort to make this clear in the paper.

The authors introduce a specific auxiliary question about the proper role of government in regulating environmental, health, and safety risks to their survey respondents in order to examine bias in the political preferences of the KN sample. Given that single-item measures have poor psychometric properties, to what extent can the authors be sure that their question reliably and validly measures political preference? The authors should consider addressing this limitation in their Discussion section?

This qualification would be easy to make. We have access only to our own surveys, not to the many others that KN has done. We needed one illustrative example of a question that might speak to the accusations of "liberal bias." This is the best that is available.

Section 2.2 Associating Census Tract Factors and 2000 Presidential Voting Patterns

The authors should provide a reference for their discussion of "estimated regressors" on page 11.

Adrian Pagan's 1984 article in the International Economic Review is the standard reference, although perhaps a textbook reference from Greene's Econometric Analysis would be most helpful to the average reader. We can easily include both.

Section 4. Sensitivity Analysis Using Fitted Propensities/Probabilities

Page 18, second paragraph, line 7: Make the word "result" plural.

Section 5. Conclusions and Caveats

The authors should be explicit that their results, in the context of the limitations of their methods (i.e., use of single-item measure to assess political preference), do not support the conclusion that the KN panel is biased according to political preference.

We can modify the conclusions to make it clear that the point of the paper is to lay out HOW one should approach the question of ideological biases in the results of a survey using the KN panel. The main point is that one cannot address non-random selection on UNOBSERVABLES without formally modeling the selection process itself and allowing for correlated residuals across the

selection process and the outcome equation of interest. This emphasis should be easy to incorporate.

Add a period to the very last sentence of the manuscript.

Easy.