

GOV 1010 AT HARVARD UNIVERSITY – AN MSG CASE STUDY

Background:

GOV 1010 is an annual introductory undergraduate survey research class taught at Harvard University. The course draws an interdisciplinary mix of students, with a heavy concentration of Statistics and Government concentrators. The Fall 2014 Semester Syllabus included a class project designed to enhance understanding of survey research. As part of the project, students are asked to work in a small group to conduct an original survey on a topic of their own choosing. In service to an excellent university-client relationship with research academia at Harvard, MSG donated ABS sample in 2014-2015 to help a group of four students conduct a rigorous and scientifically representative study.

The Survey:

The survey was designed to measure the social, financial, and community dimensions of well-being, with a focus on differences between cities and suburbs. The study was focused on the greater Boston metro area, and was designed to obtain approximately equal numbers of urban and suburban respondents.

The Target Population:

The students wanted a sample that was statistically representative to residents of households in the greater Boston area as designated by the (Massachusetts) Metropolitan Area Planning Council (MAOC). Since the students were interested in differences between urban and suburban residents, the students designed the sample to target equal numbers of urban and suburban households, effectively oversampling the MCPA's urban core and adjoining (streetcar) suburbs.

Survey Mode:

Based on class readings, the students decided to use a multi-mode approach in order to increase response rate. They employed an Address-Based-Sample (ABS), with mail as a primary form of data collection, and included a telephone -sample with a sub-sample of respondents as a follow-up. The students used further research to design an optimal survey to encourage response, including a short questionnaire, clear formatting, a separate cover-letter, and first-class postage for both the mailing and the return.

Sample Frame:

The household sample was randomly generated from the MSG ABS sample frame. MSG's address frame is based on the Computerized Delivery Sequence File (CDSF) provided by the U.S. Postal Service. With more than 135 million delivery points on file, the latest generation of the CDSF is the most complete address database available. As such, it is safe to assume that if an address cannot be matched against the CDSF, it is most likely an undeliverable address. In as much, MSG's ABS frame essentially eliminates any coverage issues associated with other sample designs such as RDD (Random Digit Dialing). What is more, by providing validation services for both correctness and completeness of addresses, the CDFS can significantly enhance the address hygiene. Consequently, this system helps reduce the number of undeliverable-as-addressed mailings, increase the speed of delivery, and reduce cost. Also, with daily feedback from tens of thousands of letter carriers the database is updated on a nearly continuous basis.

Sample Selection:

The sample stratification was defined by the students using census block-groups. The equal probability-based sample excluded non-deliverable addresses types (listed on the CDSF as vacant, seasonal, and similar characteristics), effectively improving sample efficiency with a slight (1%) reduction in population coverage. MSG also significantly reduced Frame Multiplicity by selecting PO Boxes that were flagged as OWGM (Only Way to Get Mail) and Throwbacks (mail addressed to a street address is delivered to a PO Box) while excluding Traditional PO Boxes. MSG also appended telephone numbers in cases where a listed telephone number was available. The cover-letter asked that the survey be completed by the person in the household with the most-recent birthday.

Effectiveness of Sample:

In total, MSG provided 2,000 sample records to the students – divided equally among the two strata., although the students were only able to mail surveys to half of the records due to field costs. Despite the short time-frame the students had to complete the survey (less than two weeks) and the fact that the survey was mailed only a day before Thanksgiving, the students received a total of 119 returned mail surveys within ten days of the mailing. In addition, the students conducted a telephone follow-up with a random sub-sample of 200 households (among those with telephone numbers), yielding an additional 21 responses. The overall response rate for the survey (AAPOR RR2) was 14.27%.

Comments:

Course instructor Chase H. Harrison commented “MSG’s willingness to donate the sample for this project is just incredible. I always encourage my students to use the best tools available, although I don’t always have the resources to help them use those tools, especially for things like general-population samples. Thanks to MSG, this group was able to use a sample that otherwise would have been beyond their reach, and were able to experience what it is like to work on real-time survey with professional tools.”