Haoran Yang

Stacey Suver

CSE 300.03

November 17, 2017

Research Proposal on Mixed Mode Data Collection

Data collection is a crucial step in many researches and decisions since the data collected is used to predict the major trend of sample's opinions. To obtain high quality data, various ways of data collection methods were developed. These data collection methods are generally categorized into two categories, where one is single mode data collection and the other is mixed mode data collection. The two ways of data collection has its own strength and weaknesses considering collected data's accuracy, reliability and efficiency. These two ways can be decomposed into more specific methods that differ in data quality as well as collection costs. However, based on previous studies, mixed mode data collection are proven to be more adaptable than traditional ways. Since mixed mode data collection involves multiple mode of collection methods, different combinations would result in different data qualities. This research proposal hopes to analyze previous developed mixed mode methodologies and explore an optimized mode combination that can obtain most accurate, reliable and efficient data with the least collection cost. In addition, the research hopes to lead future related researchers that investigate the possibilities of different combinations in mix mode data collections.

Although scholars in previous researches regarding mixed mode data collections question some aspects of mixed mode strategy such as the potential of bias in collected data,

they all agree that mixed mode data collection is more flexible in respect to single mode data collection method. EKOS Research Associates Inc. listed some primary reasons using mixed-mode data collection, which includes cost, interaction with sample region, response rate, consistency, and better visual depictions. In another report where Daniel Tobin and other researchers conducted a research, which concludes that mixed mode data collection by emails and postal emails can not only increase response rates but also reduce cost. Daniel's conclusion is similar to the first of primary benefits mentioned in the EKOS's research. In addition, B. Janssen's research also pointed out that mixed mode data collection is essential when performing her experiment on westerners with low income since interviewees lack online access.

Granted the advantages mentioned above, some researchers also suggest that the challenges of using the mixed mode data collection. Data Recognition Corporation pointed out that most of the researchers did not mention the bias on the different modes for gathering the data, and most of the literatures only consider paper/pencil versus the web. The corporation also pointed out the potential difference between conducting a mixed-mode data collection and different data gathering for each. Similarly, in Edith D. de Leeuw's research, he also pointed out that it is hard to design the questionnaire for the mixed mode data collection with considering the response rate, bias reduction, and the cost.

However, previous researches merely consider the difference between paper-based data collection and web-based data gathering. We need to know if mixed-mode data collection is quality and cost optimized compared to other mixed-mode ways such as the telephone-based data collection with paper-based or with web-based. To determine the most efficient mode

combinations, I would conduct the survey in 3 counties on long island with about 5,000 populations, and distribute the surveys with three distinct mode combinations. Then the experiment would collect responses from the three counties with the consistent distribution method. All three towns would receive the questionnaire with exact same content. We will conduct mixed-mode survey by paper-based and web-based survey for the first village. In order to improve the accuracy of the survey, we would ask the villagers to fill in the address, and compare them later on. Also, we would not ask personal questions in the questionnaire since responders would rather not to answer the question which would decrease the accuracy of the data collection. With the consideration of the population of the village, it would take us 1 month to collect the data. After the first month, we will conduct the same survey in the second county by paper-based and telephone-based. Since part of them is done by telephone-based data collection, we can ask the respondents with a scale of 10 and we would ask them to give a rate on a certain event. In order to increase the response rate, we would not tell the respondents the choice of not to answer. However, when respondents express wishes to not answer the question, they are allowed to do so. Estimated time of completely collect county two's data would require about 4 months. Lastly, the third data gathering survey will involve a combination of telephone-based and by web-based. It would also take about 4 months to gather the information from the sample group. For accuracy purposes, repetition of the experiment is crucial. The experiment should be conducted again with similar survey after a year of the first round with exact settings. Since people would change their preference in that year. We need to set the error rate of changing their choices would be in a reasonable rate (i.e. 10 percent).

The expected combination with the least cost, highest response rate, and the highest efficiency to be the third mixed-mode way, which is the combination of telephone-based and web-based data collection. Our hypothesis reasons as follows: the appointment with the respondents for the telephone interview is flexible, which would likely to increase the response rate; the cost of sending mails to the respondents is not under consideration since it's web and telephone based; the convenience for the residents to respond by filling out the form automatically online. With above considerations, we would expect the optimized strategy combination would be the combination of the telephone-based interview and web-based form.

Although the experiment hopes to determine the most efficient data collection mode combinations, data collection method would still need flexibilities depending on the respondents. However, even though our research does not analyze fully the potential probabilities of mixed mode data collection methodologies, our experiment is significant since it explores the specifics of the combinations whereas previous researches solely compares single mode versus mixed mode data collection. More experiments should be performed in the future that takes the respondent's situations into account in the future. In conclusion, this research hopes to provide in depth some possible combinations in mixed mode data collection strategy and a comparison among these possible combinations.

Works Cited

Baum, Herbert M., et al. "Mixed-Mode Methods for Conducting Survey Research." *Encyclopedia of Survey Research Methods*, 2012, doi:10.4135/9781412963947.n300.

De Leeuw, Edith D. "To Mix or Not to Mix Data Collection Modes in Surveys." *Journal of Official Statistics*, vol. 21, no. 2, June 2005, pp. 233–255.

Government of Canada, Public Works and Government Services Canada, Integrated Services Branch, Government Information Services. "Mixed-Mode Public Opinion Surveys and Mode Transition - Standards and Best Practices - Public Opinion Research in the Government of Canada - PWGSC." *Government of Canada, Public Works and Government Services Canada, Integrated Services Branch, Government Information Services*, 27 July 2017, www.tpsgc-pwgsc.gc.ca/rop-por/rspe-srbp-eng.html.

Janssen, B. "Web data collection in a mixed mode approach: an experiment ." *European Conference on Quality in Survey Statistics*, 2016.

"Mixed-Mode Surveys: A Strategy to Reduce Costs and Enhance Response Rates." *The Journal of Extension (JOE)*, joe.org/joe/2012december/tt8.php.