

NON-RESPONSE BIAS IN SURVEY SAMPLING

White Paper Series



Introduction

As market researchers, we use data gathered from surveys to make informed decisions and give recommendations to clients on ways to improve their sales and standing in the marketplace. Whether these recommendations are to invest resources to increase satisfaction with customer support or to introduce a new product to the market, we need to feel confident that the results from many crosstabs and multivariate analyses are speaking the truth. The confidence we have in our results stems from the quality of our data. In today's market research industry, we do a lot to help ensure our data meet certain standards: screener questions target the specific audience we want, online panels take many steps to ensure their samples contain the target we need, we weight respondents to match specific population demographics, etc. [Refer to *Situational Use of Data Weighting* for more details] However, one of the most over-looked problems is that of non-response bias.

Non-Response Bias

In data collection, there are two types of non-response: item and unit non-response. Item non-response occurs when certain questions in a survey are not answered by a respondent. Unit non-response takes place when a randomly sampled individual cannot be contacted or refuses to participate in a survey. The bias occurs when answers to questions differ among the observed and non-respondent items or units. A general formula for measuring bias is:

$$\text{Bias} = P (O - N)$$

Where

- P** is the proportion of non-respondents from the targeted sample (i.e. non-response rate)
- O** is the answer based on observed responses
- N** is the answer based on non-respondents only

Bias is calculated as the product of two components: non-response rate and the difference between the observed and non-respondent answers. Increasing either of the two components will lead to an increase in bias. Since it's quite difficult and often impractical to design a survey to impact the difference between the observed and non-respondent answers, researchers will often focus their attention on reducing the non-response rate in order to reduce bias.

The first (and possibly most important) step in reducing non-response bias is to create a properly designed survey. Whether it be online or by phone, the design of the survey can have a large impact on whether a respondent chooses to partake in the survey, and to what extent they complete the survey. Having a personable yet professional introduction, interesting survey content, short survey length, clear and concise wording, practical and appealing incentives, placing multiple follow-up calls or email reminders on non-respondents, and being mindful of the time, day, or season that the survey is fielded all can impact the non-response rate. Even after designing a great survey, both item and unit non-response are likely to exist.

There are many ways to deal with item non-response, with case deletion and mean replacement being the most popular. Due to researchers' familiarity with item non-response, the remainder

of this paper focuses on the less talked about subject of unit non-response bias. Our next steps are to measure and then adjust for any bias in the data.

Identifying Unit Non-Response Bias

Comparing Initial and Late Respondents

As mentioned earlier, following up on non-respondents is an excellent way to reduce the non-response rate, but it also gives us some additional information. Because late-respondents, or those that respond after several attempts, are theorized to have some similarities with non-respondents, one approach would be to compare scores on key metrics from both the initial respondents and the late respondents. Any differences would be considered an estimate of non-response bias. Of course, we need to keep in mind that these similarities with the non-respondents and the differences from the initial responders do not always pan out.

Comparing Survey Results to Known Population Parameters

Instead of comparing initial and late responders, we can compare the demographic profiles (e.g. age, gender, race, and income) of our respondents to some reliable external source. One option for this would be U.S. Census data for our intended target population. If the comparison results in clear differences, we conclude that these differences indicate that we may have non-response bias in our data. While this method allows for the comparison of data gathered on respondents to population totals, it lacks the ability to measure differences on key variables of interest. In order to evaluate key variables, a popular technique is to weight the survey respondents based

on census totals and compare the weighted and un-weighted results. [Refer to *Situational Use of Data Weighting* for more details] If they differ, we may conclude non-response bias is present, assuming that the demographic or database variables have an association with the response rate.

Using Known Database Variables to Identify Bias

Suppose that along with contact information for the entire sampled group, we have some additional demographic and database (e.g. tenure with company, purchase quantity) variables. One option utilizing this extra information is to examine non-response rates over different sub-groups of the population using the demographic variables. Differences indicate the possible presence of bias in the data. Another option is to use the database variables and compare statistics among the responders and non-responders where any differences give evidence of non-response bias. Both of these methods fail to focus on key survey variables and assume that the demographic or database variables are correlated with non-response bias.

Adjusting for Unit Non-Response Bias

The above methods give the researcher a sense of whether bias exists, but do not provide a way to deal with it. The following procedures both measure and adjust for non-response bias.

Weighting-Class Adjustments

Suppose again that additional demographic or database variables are available for all members of the targeted sample group. These variables are used to create sub-groups containing respondents and non-respondents. Weights are then calculated based on the proportions in each sub-group and applied to the respondents to reflect the total sample population. Comparisons on key variables are then observed between the unadjusted and weighting-class adjusted respondents. If clear differences are detected, then non-response bias is assumed to be at fault and the weighting-class adjustments are used as they provide results with less bias. Poststratification is another technique similar to weighting-class adjustment, except that the procedure uses population counts instead of the total sample counts. The downside to these techniques is that they assume that the differences between respondents and non-respondents are captured in the subgroups, and that there is no rule of thumb for comparing adjustments to determine which to use.

Other Adjustment Techniques

A couple of other techniques exist to adjust for non-response bias such as propensity models, which require some information (e.g. demographics) be known for the entire sampled group, or calibration methods, which make the use of auxiliary population data like from

a census. Both of these methods are extensions to techniques previously discussed, and the interested reader is encouraged to research more on these topics in detail.

Conclusion

The purpose of this paper was to make the reader aware of non-response bias, describe ways to reduce its effects in the design stage before fielding a survey, and explain the ways to measure and adjust datasets for non-response bias. Because post-survey adjustments are merely estimated “fixes” to the problem, the most effective way to reduce non-response bias is to reduce non-response rates through properly designed studies. Then the adjustments on the back-end will help reduce, but not eliminate, non-response bias.