

Non-probability-based Surveys in Practice

Matthew Salganik & Cambria Naslund *

Summer Institute in Computational Social Science 2019

How accurate are estimates from non-probability-based online surveys? Can results from online surveys administered to non-probability-based samples generalize to the broader population? In this activity, we will conduct a survey on MTurk workers and try a variety of estimation techniques. Then, we will compare our estimates to those that come from high-quality probability-based samples used by the Pew Research Center.

In the process, we will gain experience with writing survey questionnaires, creating surveys on Google Forms, publishing surveys on MTurk, data wrangling, implementing various weighting methods, documenting surveys for dissemination, and thinking about broader issues related to survey errors and data quality. Please read the entire activity before beginning.

Note: This assignment uses Google Forms as the survey tool because it is free. See appendix for comments about using Qualtrics, a more sophisticated platform that requires payment or institutional support.

Before group activity

- Read Chapter 3 in *Bit by Bit*
- Read the paper that motivated this activity: *Online, Opt-in Surveys: Fast and Cheap, but are they Accurate?* by Goel et al. In this paper, the researchers conduct an opt-in, non-probability survey of people on MTurk. They compared the raw sample estimates and those made after various weighting adjustments to estimates that come from surveys using probability-based sampling methods.
- We have already conducted a survey with a large MTurk sample with questions similar to the ones used by Goel et al. You are welcome to use [this data](#) to try more advanced estimation estimates. Please review [the survey](#) we used to collect these responses. You will be added as collaborators on Google Forms and can use our survey as a template for yours. We administered a selection of questions from various *2019 Pew Surveys*, but you are welcome to administer and analyze questions from the Pew survey that we did not ask.

During group activity

- Create a survey on Google Forms. You can see our template [here](#).
 - Don't forget to include demographic questions you will need later for post-stratification. You are provided with [information about the sizes of different demographic groups in the United States](#). These population estimates are collected with questions from the [American Community Survey](#).
 - Include a question asking for workers' MTurk Worker ID. This will help you validate that a worker actually completed the survey and enable you to remove people who took the survey more than once.
 - In designing your survey, think of ways to ensure the quality of the data collected. For example, use attention screeners, offer the "don't know" options judiciously, think about when to allow respondents to "skip" a question and when to require a response, etc. This should go into your documentation.
 - Make sure you preview the survey on Google Forms before you publish it on MTurk.
- Deploy your survey to MTurk. We estimate that the survey will take about 6 minutes, and we would like to pay an hourly wage of \$15 per hour so you should pay \$1.50 per completed survey. ([Getting great survey results from MTurk and Google Forms](#))

*based on the activities from SICSS 2017 (created by Matthew Salganik and Yo-Yo Chen) and SICSS 2018 (created by Matthew Salganik and Janet Xu).

- For participants at Princeton location: We have created MTurk accounts for you to use. Please create your tasks from these accounts with the pre-loaded funds. If you use your own account for data collection, we won't be able to reimburse you.
- After the data has been collected, validate and remove all personally-identifying information:
 - Download the CSV of responses from Google Forms.
 - If you used an attention screener, delete responses that did not meet the criteria.
 - Check to make sure that all of your MTurk workers actually took the survey by comparing the list of Worker IDs provided in the survey data against the Worker IDs recorded by the MTurk platform.¹
 - Remove redundant entries, if any.
 - After you have used the Worker ID data to validate answers and to remove redundant entries, delete it from your dataset. The Worker ID is a unique string that can be used to personally identify people.
- Analyze your data. Compare your raw and weighted estimates, respectively, with what's published by the existing survey. Use [this template](#) as you work through these steps.
 - If you use the [survey data we already collected on MTurk](#), you can see the results for those questions published by Pew [here](#).²
 - Replicate Figure 1 from [Goel et al.](#)
 - Next, try different adjustment methods, and for each method you use, replicate Figure 2 from [Goel et al.](#)
 - When doing the adjustment methods, you should start with the simplest thing first. Do not jump to the most fancy technique right away. You may use post-stratification packages, but only after you have coded up at least one technique by hand (if you have coded one of these techniques as part of your earlier research that counts).
 - Start with cell-based post-stratification using a small number of cells. *Hint: think about what may happen if you have empty cells.*
- Prepare your data for dissemination and open sourcing so that it can be used by others. ([Some best practices](#) for preparing survey metadata). We will talk about this in the afternoon using [these slides](#).

Tips and additional resources

- You might want to divide the work among group members, including creating the survey on Google Forms, reshaping data files, breaking up the analysis into modular functions, etc.
- [Here](#) are some more resources on web surveys and [here](#) is Amazon's best practices for MTurk Requesters.
- [Here](#) is a paper about practical tips for ethical data sharing.
- [Here](#) is a case study of some of the issues that come up with data sharing.

Further reading

- Dutwin and Buskirk. “[Apples to Oranges or Gala versus Golden Delicious?: Comparing Data Quality of Nonprobability Internet Samples to Low Response Rate Probability Samples](#)”
- Baker et al. “[Summary Report of the AAPOR Task Force on Non-probability Sampling](#)”
- Kalton and Flores-Cervantes “[Weighting methods](#)”
- Berinsky, Margolis, and Sances “[Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-Administered Surveys](#)”

¹Tip: to validate WorkerID matches, you can download a CSV of WorkerIDs from your MTurk results page and match that with your survey results data.

²*Technical note:* Most survey questions have a “don't know/refuse” residual category, but predicting the percentage of people who refused to answer may not always be substantively relevant. To omit this category from the analysis, we normalized the existing survey results by dividing by the percentage of people who responded to that question.

Appendix: Qualtrics-specific instructions

- If you are using Qualtrics instead of Google Forms, you do not have to include a question on the survey asking for MTurk Worker ID. Instead, you can verify survey completion by generating a unique validation code for each survey respondent, and have them enter that on MTurk instead.
- Here are [detailed instructions](#) for generating survey completion codes.
- Qualtrics has more features and customizable options – for example, you can randomize the response options to a survey question or enable more sophisticated skip logic. Try playing around with these!