Validating Nonresponse

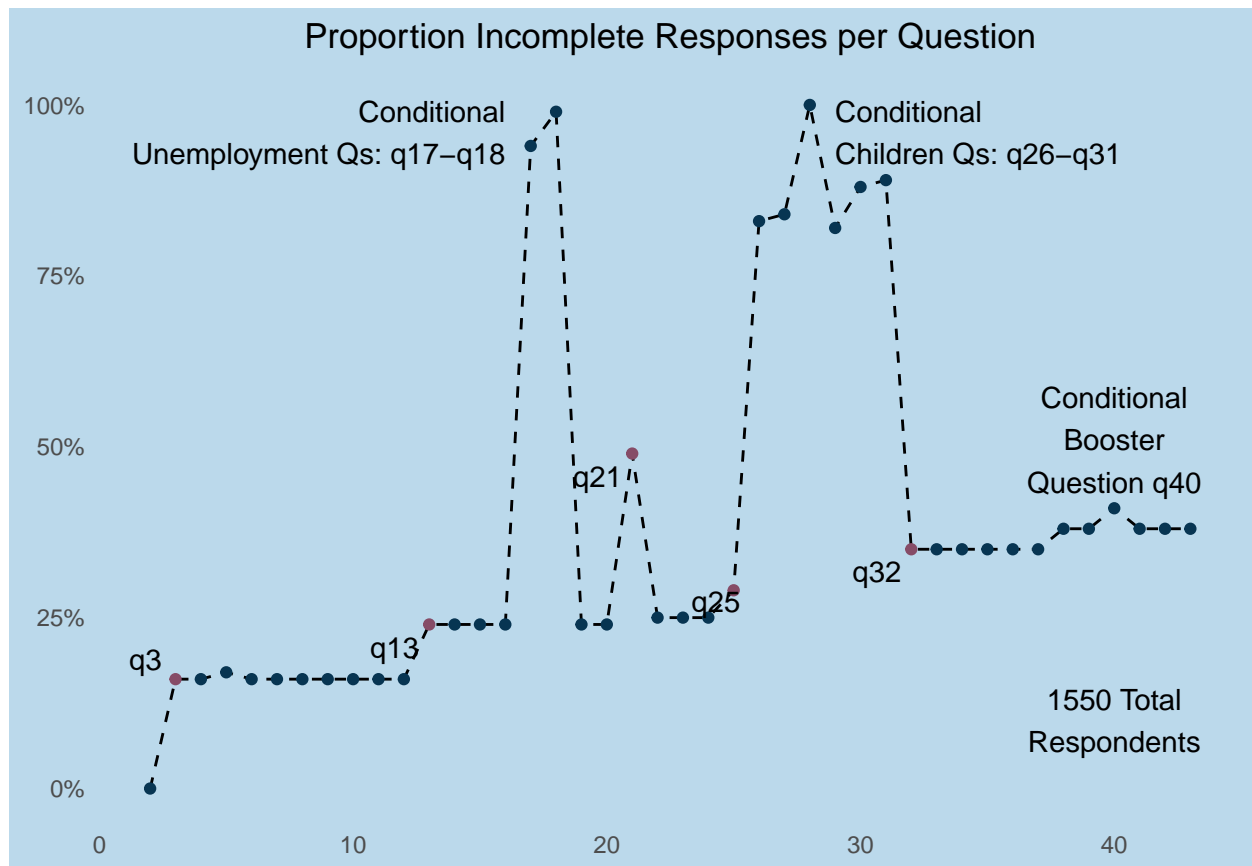Arielle Herman

2/21/2022

When do we lose respondents?

Analyzing when we lose respondents will help us shape the structure and questions of the next survey to better retain respondents throughout the survey. The below graphs and analysis are based on the newly compiled data (as of Feb 22).

validate disqualification criteria

```
by_reqs <- final_clean %>% group_by(
  q2 == 1,
  completion >= 0.5,
  duration >= 100
) %>% count

kableExtra::kable(by_reqs)
```

q2 == 1	completion >= 0.5	duration >= 100	n
FALSE	TRUE	FALSE	63
FALSE	TRUE	TRUE	6
TRUE	FALSE	FALSE	243
TRUE	FALSE	TRUE	138
TRUE	TRUE	FALSE	1
TRUE	TRUE	TRUE	1168
NA	FALSE	FALSE	110
NA	FALSE	TRUE	14

Out of **1743** total responses, **124** respondents didn't fill any responses at all and **439** respondents only completed the first question. We can see the major drop offs in respondents in the below graph:

The major drop-offs come at:

1. second question (q3), asks about zipcode.
2. twelth question (q13), the first income question
3. twentieth question (q21), that asks about experienced difficulties
4. thirtieth question (q32), that asks about government information

Second Question (q3)

This drop-off charts respondents who marked that they were NYC residents but did not complete the next question. Possible reasons for quitting the survey include:

- unwillingness to complete the survey
- said they were a NYC resident but their permanent zipcode is not in NYC
- lack of a permnanet address / zipcode

Regarding the last expanation, we may want to include a `don't know` category to encourage any respondents that didn't continue due to lack of a permanent address.

Twelth Question (q13)

Respondents also drop off right before the first income bracket question. Possible explanations include:

- difficulty getting this information

This may be especially difficult for low-income or hourly wage workers, who may not know their income bracket without significant thought. If this is the case, this question may cause selection bias toward higher income brackets.

Twentieth Question (q21)

Respondents avoid this question, but appear to continue with the question following. It is so far unclear why this is the case, and the NA responses seem pretty evenly distributed across partners. I have yet to test it on other demographics. Some possible explanations that require more thought include:

- question wasn't required in qualtrics survey
- respondents were discouraged from completing this question

Thirtieth Question (q32)

Respondents with and without children are not responding to q32:

has_child	answered_q32	n	prop
household with children	answered question q32	265	86%
household with children	did not answer q32	44	14%
household without children	answered question q32	730	91%
household without children	did not answer q32	70	9%

We lose the largest number of respondents from the demographic of no household children, but the larger proportion of respondents from households with children. If respondents from households without children are quitting the survey because of the skip logic, we may want to encourage them to complete it by moving the children questions later in the survey. However, it is unclear why there is a drop off between these two questions, and more thought is required.