

Non-response Bias in Surveys

Yian Wang

Nov 1, 2020

1. Introduction

This report summarizes the process in beginning to understand non-response bias in unemployment surveys. The Gelman paper “The Mythical Swing Voter” analyzes how non-response bias drives polling differences in politics. Further, we look into how unemployment surveys are conducted in Canada, and analyze relevant studies regarding our question about non-response survey bias.

The goal is to eventually see whether the unemployment numbers given COVID-19 are artificially low due to non-response bias.

2. Background: How Unemployment Surveys are Conducted

According to “The Evolution of Rotation Group Bias”, unemployment surveys in Canada are conducted by a Labour Force Survey (LFS) every month (Krueger et al.). These are sent out at random to avoid bias (“Surveys and Statistical Programs”). The majority of participants are contacted and interviewed during a week in the middle of the month, and thus the comparative time frame is similar each month (Krueger et al.).

Statistics Canada states that some of their non-response rates are reconciled by imputation and adjusting weights for households in the same area, under the assumption that the characteristics of nearby households are similar to that of the non-responding household. The average non-response rate on the LFS is about 10% (“The Daily”).

3. Relevant Literature

3.1 The Mythical Swing Voter

Published Date: April 18, 2016

Authors: Andrew Gelman, Sharad Goel, Douglas Rivers, and David Rothschild

Affiliation: Columbia University, Stanford University, Microsoft Research

URL: <http://www.stat.columbia.edu/~gelman/research/published/swingers.pdf>

3.1.1 Overview

“The Mythical Swing Voter” analyzes the plausibility of voters constantly changing their opinion on their initial voting intent by drawing on three studies MRP on the 2012 election and applying MRP to correct for biases to conclude that the recorded vote swings in official polls were largely due to non-response and sample selection bias.

The first study drawing on traditional cross-sectional surveys gives two potential explanations for seeing Obama’s support drop from 55% to 48%: (1) individuals were genuinely influenced by debates and changed their voting intentions based on that alone, or (2) the two samples of polls taken before and after the debate were different. The conclusion of this study is that cross-sectional data does not make it possible to determine which explanation is true since the respondents in the sample do not overlap.

The second study observes an Xbox panel survey. This was done by surveying 345,858 unique individuals during the 45 days before the election about their voting intentions. Then, by adjusting for factors and post-stratifying, this experiment's results the day prior to the election closely paralleled the election outcomes estimated from exit poll data. The conclusion of the Xbox panel is that there is insufficient evidence that Romney gained significant support from Obama's side following the first debate.

The third study looks at the RAND American Life Panel, which has a higher response rate alongside a more stable sample composition. Changes in this panel are compared to that of the first study in order to approximate the selection bias in the recorded drop in Obama's support after the first debate. The RAND panel shows much smaller swings than those recorded in the first study, and estimates Obama's lowest support rate at 51%. This is similar to the estimate in the second study with 50%.

3.1.2 Limitations

- This paper studies the non-response bias on political opinion polls, whereas the purpose of our investigation is to see the effects of non-response bias on an objective measure not based on opinion.

3.2 What Can Instrumental Variables Tell Us About Nonresponse in Household Surveys and Political Polls?

Published Date: Jan 29, 2019

Authors: Coady Wing

Affiliation: Indiana University

URL: <https://pdfs.semanticscholar.org/053d/7832f6850a1e9030076a46e8041861ad5fb7.pdf>

3.2.1 Overview

This paper outlines how understanding non-respondents in surveys is becoming more critical with the declining response rate in political polls and household surveys. By analyzing how exogenous increases in survey participation affects sample composition, Wing connects questions regarding survey outcomes to a causal model of survey participation.

Additionally, he introduces a data transformation and instrumental variable estimator that when applied to data with missing responses, gives rise to a new summary statistic, Compiler Average Survey Response (CASR), that measures average response from respondents who join due to the exogenous increase in response rate.

By applying these above methods to the 2011 Swiss Electoral Study data, Wing finds that cash incentive definitively induced survey participation, while altering the bias of the weighted results. Thus it is concluded that in the original study, the non-response bias was not fully explained by the observed covariates.

3.2.2 Limitations

1. The CASR statistic is only valid under 4 assumptions:
 - Independence of instrumental variable
 - Exclusion: No causal effect is exerted on the outcome variable by the instrument or participant variable
 - First Stage: The instrument affects some of the population's participation
 - Monotonicity: Participation is affected by the instrument in the same direction for every individual
2. This method does not determine whether there exists non-response bias, nor does it correct it; it only shows how an exogenous change affects the sample composition and allows the researcher to understand these insights

3.2.3 Other Notes

- Design-based model so it can be applied to a number of areas with weaker assumptions
- Cites "The Mythical Swing Voter"

3.3 Difficulty to Reach Respondents and Nonresponse Bias: Evidence from Large Government Surveys

Published Date: December 20, 2017

Authors: Ori Heffetz and Daniel B. Reeves

Affiliation: Cornell University

URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2758787

3.3.1 Overview

This paper looks at the difference in responses between respondents that are easy to reach or hard to reach, depending on how often they were contacted to be surveyed. Analyzing three popular government surveys that report attempts made to contact, the findings help to assess sensitivity in estimates to response rates and assumptions about non-respondents.

Based on the number of times contacted, respondents are divided into 3-4 groups of reach-difficulty and compared each group's outcomes to see cross-demographic group differences. The Census/BLS's Current Population Survey (CPS) is used to analyze labour force participation rate and unemployment rate.

Even following a control for demographic differences, unemployment rate consistently decreases 1.5% from easy-to-reach to hard-to-reach, and thus this difficulty of reaching is found to be heavily correlated with the outcome of interest. This is important because someone considered difficult-to-reach in one survey would be a non-respondent in a different one since they could be so hard to reach that they become a non-respondent.

3.3.2 Limitations

- Though there is indirect evidence, this data cannot be used to directly see if within-sample trends extend to non-respondents
- Difficulty-of-reaching measure is noisy
- The results are not reported by an automatic system; the interviewer reports may be (un)intentionally misreported

3.4 Telephone Presurveys, Self-Selection, and Nonresponse Bias to Mail and Internet Surveys in Economic Research

Published Date: Aug 21, 2006

Authors: Darren Hudson, Lee-Hong Seah, Diane Hite, and Tim Haab

Affiliation: Mississippi State University, Auburn University, Ohio State University

URL: <https://www.tandfonline.com/doi/full/10.1080/13504850410001674876>

3.4.1 Overview

This paper outlines the limitations of conducting telephone, mail, and internet surveys. While mail surveys are low-cost, the largest issue surrounding these is the non-respondents as their characteristics are not observed regardless of sample size. Not everyone has a phone, or is comfortable with giving personal information to someone asking on the other end. Similarly with internet surveys, accessibility is a concern, along with the fact that people may not be willing to share their personal details over a random online survey. Thus non-response is an issue for all of these surveys.

By performing a telephone pre-survey to see demographics, then a followup survey over the internet or mail on American households, the researchers were able to directly see non-response bias and its demographics from the pre-survey. The findings are that the internet surveys exhibit no evidence of greater non-response than mail surveys, though it did exhibit much lower response rates.

3.4.2 Limitations

- Self-selection bias

- In the pre-survey they were asked if they wanted to do a follow-up and only contacted those who said yes; this result should be analyzed more

3.5 The Evolution of Rotation Group Bias: Will the Real Unemployment Rate Please Stand Up?

Published Date: September 2014

Authors: Alan B. Krueger, Alexandre Mas, and Xiaotong Niu

Affiliation: Princeton University, Congressional Budget Office, IZA

URL: <http://ftp.iza.org/dp8512.pdf>

3.5.1 Overview

This paper discusses rotation group bias, rather than non-response bias in unemployment surveys. However, it investigates Canada's Labour Force Survey (LFS) in regards to discussion group bias and how non-response bias contributes to it. There is evidence of non-response contributing significantly to rotational group bias.

In the Canadian monthly Labour Force Survey, there exist 6 rotation groups that are interviewed for 6 consecutive months, with a low response rate at around 5% in 2005. They found no systematic tendency for rotation group bias in the Canadian LFS.

Though they found that non-response does bias the estimates of unemployment duration. There are two reasons for the bias: (1) non-respondents have different answers which are not recorded, and (2) finding a job prior to the survey date increases probability of becoming a non-respondent. The latter bias always exists.

3.5.2 Limitations

- Doesn't provide solutions to mitigate non-response bias, despite evidence suggesting we do need to find ways to reduce it.
- Data is from January 1976 to December 2010, might be slightly out-dated now.

4. Conclusion

There is no literature about economics and unemployment in particular that cites the Gelman paper directly, though there are a number of relevant literature surrounding the idea of non-response bias in economic and household surveys.

Wing analyzes exogenous increases in survey participation to show how non-respondents can indeed affect sample composition. He develops a sample statistic, Compiler Average Survey Response, to measure these average exogenous increases (Wing).

Additionally, "Difficulty to Reach Respondents and Nonresponse Bias" finds that as a respondent becomes harder to reach, the more the outcome of interest is affected. Since a hard-to-reach respondent may be a non-respondent in a different survey, there is indirect evidence showing that this may be extendable to non-respondents in surveys (Heffetz and Reeves).

In "Nonresponse Bias to Mail and Internet Surveys", limitations of conducting telephone, mail, and internet surveys are outlined, showing why non-response may be an issue. Accessibility and privacy are two main factors (Hudson et al.).

Finally, "The Evolution of Rotation Group Bias" analyzes rotational group bias and how that is affected by non-response bias. There is evidence of increased non-response contributing to increased rotational group bias. In Canada's LFS, they find no tendency to rotational group bias, which may be due to non-response rate being low (Krueger et al.).

Most of the literature was regarding non-response bias in large government surveys, but not specifically those of unemployment. This is likely due to the fact that the Labour Force Survey is mandatory and there are

repercussions in place to failure to discourage non-response. Further, non-response rates are reconciled by imputation and weight-adjustment (“Surveys and Statistical Programs”). While then it may be possible that remaining non-response bias is negligible, that is not to say there does not exist additional bias in these unemployment surveys as in “The Evolution of Rotation Group Bias” (Krueger et al.) or accessibility concerns that can contribute to bias (Hudson et al.).

Still, in 2005, the non-response rate for the LFS was about 5%, whereas it has averaged 10% more recently (“The Daily”). I investigated whether this increase in non-response could have been due to COVID; however, upon looking at previous years leading up to 2020, there was no evidence supporting this hypothesis as this rate has been consistently averaging 10% for the past few years (“The Daily”). Thus, we cannot tell whether there is a significant difference in non-response in relation to the pandemic.

5. Sources Cited

- Heffetz, Ori and Daniel Reeves. “Difficulty to Reach Respondents and Nonresponse Bias: Evidence from Large Government Surveys.” *SSRN Electronic Journal*, 2016, 10.2139/ssrn.2758787. Accessed Oct 30, 2020.
- Hudson, Darren, et al. “Telephone Presurveys, Self-Selection, and Non-Response Bias to Mail and Internet Surveys in Economic Research.” *Applied Economics Letters*, vol. 11, no. 4, 21 Aug. 2006, pp. 237–240, 10.1080/13504850410001674876. Accessed 30 Oct. 2020.
- Krueger, Alan B., et al. “The Evolution of Rotation Group Bias: Will the Real Unemployment Rate Please Stand Up?” *The Review of Economics and Statistics*, vol. 99, no. 2, Sept. 2014, pp. 258–264, 10.1162/rest_a_00630. Accessed 30 Oct. 2020.
- “Surveys and Statistical Programs - Labour Force Survey (LFS).” *Statcan.Gc.Ca*, Government of Canada, Statistics Canada, 19 Oct. 2020a, www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=1288011. Accessed 27 Oct. 2020.
- “The Daily — Labour Force Survey, August 2020.” *Statcan.Gc.Ca*, Government of Canada, Statistics Canada, 4 Aug. 2020, www150.statcan.gc.ca/n1/daily-quotidien/200904/dq200904a-eng.htm. Accessed 27 Oct. 2020.
- Wing, Coady. “What Can Instrumental Variables Tell Us About Nonresponse In Household Surveys and Political Polls?” *Political Analysis*, vol. 27(3), 29 Jan. 2019, pp. 1–19, 10.1017/pan.2018.58. Accessed Oct 29, 2020.